Spring 4-22-2013

Three Essays on Return Predictability and Decentralized Investment Management

Dashan Huang

Washington University in St. Louis

Washington University in St. Louis
Olin Business School

Dissertation Examination Committee:
Guofu Zhou, Chair
Hung-Gay Fung
Ohad Kadan
Hong Liu
Rodolfo Manuelli
Ngoc-Khanh Tran

Three Essays on Return Predictability and Decentralized Investment Management

by

Dashan Huang

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2013

St. Louis, Missouri

Contents

Acknowledgments

I wish to sincerely thank my advisor Guofu Zhou for his constant encouragement and guidance during the past five years, and this thesis is just a celebration of my cordial relationship with him as an advisor and a friend. I am grateful to Ohad Kadan, Hong Liu, and Ngoc-Khanh Tran for their constant support and valuable comments. I would also like to thank Hung-Gay Fung and Rodolfo Manuelli for serving on my dissertation committee and for providing helpful suggestions. Special thanks go to the members of finance group for many discussions and feedback on my work.

I am immensely indebted to my parents for their nurture, upbringing, and incredible love. I would like to express my heartfelt gratitude, love, and admiration to my wife, Jiexun, the most beautiful woman in my eyes. Without her unconditional support, encouragement, patience, and contentment, I would not have made my PhD program go so smoothly. Finally, this thesis is dedicated to my daughter, Sophie.

Chapter 1

What Is The Maximum Return Predictability?

This paper investigates whether return predictability can be explained by existing asset pricing models. Using different assumptions, we develop two theoretical upper bounds on the R-square of the regression of stock returns on predictive variables. Empirically, we find that the predictive R-square is significantly larger than the upper bounds, implying that extant asset pricing models are incapable of explaining the degree of return predictability. The reason for this inconsistency is the low correlation between the excess returns and the state variables used in the discount factor. The finding of this paper suggests the development of new asset pricing models with new state variables that are highly correlated with stock returns.

1.1 Introduction

In the past three decades, financial economists and investors have found numerous economic variables that can be identified as predictors of stock returns.[1] The evidence on return predictability has led to the development of new asset pricing models, such as the habit formation model (Campbell and Cochrane 1999), the long-run risk model (Bansal and Yaron, 2004), and the rare disaster model (Barro, 2006; Gabaix,

[1]Examples include the short-term interest rate (Fama and Schwert, 1977; Breen, Glosten, and Jagannathan, 1989; Ang and Bekaert, 2007), the dividend yield (Fama and French, 1988; Campbell and Yogo, 2006; Ang and Bekaert, 2007), the earnings-price ratio (Campbell and Shiller, 1988), term spreads (Campbell, 1987; Fama and French, 1988), the book-to-market ratio (Kothari and Shanken, 1997), inflation (Campbell and Vuolteenaho, 2004), corporate issuing activity (Baker and Wurgler, 2000), the consumption-wealth ratio (Lettau and Ludvigson, 2001), stock volatility (French, Schwert, and Stambaugh, 1987; Guo, 2006), output (Rangvid, 2006), oil price (Driesprong, Jacobsen, and Maat, 2008), output gap (Cooper and Prestley, 2009), and open interest (Hong and Yogo, 2012).

2012; Gourio, 2012; Wachter, 2012). While many asset pricing models can generate time-varying expected returns, it is unclear whether they allow the same degree of predictability as observed in the data.

This paper asks whether predictability can be *fully* explained by a general asset pricing model, of which the above three models are special cases. To answer this question, I develop two theoretical upper bounds on the R^2 of the regression of stock returns on any predictive variable. If the predictive R^2 is less than the bounds, return predictability is consistent with asset pricing models. Otherwise, the models can be rejected. In this sense, the proposed bounds provide a new way to diagnose asset pricing models.

With the assumptions that the stochastic discount factor (SDF) is a function of a set of known state variables and investors' risk aversions have an upper bound (maximum risk aversion), the first bound in this paper depends on three key parameters: the multiple correlation between the excess return and the state variables of the SDF, the maximum risk aversion, and the volatility of the marginal investor's optimal wealth. The rationale of the maximum risk aversion is from Ross (2005) who shows that the volatility of the SDF is positively related to risk aversion and that any upper bound on the SDF volatility is directly related to the upper bound on the marginal investor's risk aversion.

Instead of the maximum risk aversion, the second bound assumes that the volatility of the SDF is bounded above by the market Sharpe ratio and also depends on three important parameters: the multiple correlation (as used in the bound with maximum risk aversion), the market Sharpe ratio, and a parameter chosen by end-users that excludes arbitrage opportunities or "good-deals" in the sense of Cochrane and Saá-Requejo (2000). This bound is in the spirit of Ross (1976) and Cochrane and Saá-Requejo (2000) who advocate using the market Sharpe ratio to restrict the SDF volatility. The intuition is that extremely high Sharpe ratios cannot persistently exist in the market and the volatility of the SDF is intimately linked to the market Sharpe ratio. Hence, excluding extremely high Sharpe ratios is equivalent to imposing an upper bound on the SDF volatility.

In the applications, I consider ten widely explored variables utilized by Goyal and Welch (2008) to predict the excess returns of the market portfolio and cross-sectional portfolios, such as portfolios formed based on size, book-to-market ratio, momentum, and industry. For the state variables in the SDF, I first consider the consumption growth rate and the three factors used by Fama and French (1993). The results show that the predictive R^2s are almost always larger than the proposed upper bounds. When the consumption growth rate is used as the state variable in the SDF, the two proposed bounds are approximately zero regardless of any of the ten predictors is used. When the state variables are the Fama-French three factors, out of ten predictors, six predictors generate larger R^2s than the bounds with the maximum risk aversion and seven are larger than the bounds with the market Sharpe ratio. Cross-sectionally, when any one of the ten variables is used as a predictor, with several exceptions, all the predictive R^2s violate the upper bounds, no matter whether the state variables of the SDF are the consumption growth rate or the Fama-French three factors.[2].

I then consider the market portfolio forecast in the case when the state variables are those used in the habit formation model, the long-run risk model, or the rare disaster model. The state variables in the habit formation model are the consumption growth rate and the surplus consumption ratio. All the ten predictors generate larger R^2s than the two bounds. For example, when the dividend-price ratio is the predictor, the predictive R^2 is 0.27% while the upper bound is 0.03% with the maximum risk aversion and 0.02% with the market Sharpe ratio. Constantinides and Ghosh (2011) show that the state variables in the SDF of the long-run risk model can be the consumption growth rate, the risk-free rate, and the dividend-price ratio.

[2]The results are robust when the momentum factor is included.

Nine predictive R^2s violate the two bounds. With respect to the rare disaster model, Wachter (2012) shows that the state variables can be the consumption growth rate and the dividend-price ratio. In this case, both bounds are similar to that in the habit formation and long-run risk models, and all ten predictive R^2s exceed the two proposed bounds significantly. In summary, one can conclude with a high degree of confidence that the above three models explain only a fraction of predictability.

What happens when market frictions are introduced into the bounds? It may be the case that the profits documented in the literature are not attainable for investors due to the presence of market frictions. I follow Nagel (2012) by augmenting the SDF with a factor that captures different notions of transaction cost, such as the marginal value of liquidity services of tradeable assets in Holmström and Tirole (2001), the transaction costs in Acharya and Pedersen (2005), or the funding liquidity in Brunnermeier and Petersen (2009). When the liquidity factor in Pátor and Stambaugh (2003) is used as a proxy of transaction cost, the proposed bounds are improved but still less than the predictive R^2s significantly. In this sense, transaction cost or market friction is not a key source to explain return predictability.

Since the bounds are robust to any specification of investors' preference, the incapability of extant asset pricing models in explaining return predictability is mainly due to the low contemporaneous correlation between the excess return and the state variables. This explanation is supported by the fact that the upper bounds are higher when the state variables are the Fama-French three factors than the consumption growth rate, because the Fama-French three factors have a higher contemporaneous correlation with the excess return. Therefore, the finding of this paper suggests the development of new asset pricing models with new state variables that are highly correlated with stock returns. This is consistent with Cochrane and Hansen (1992) and Campbell and Cochrane (1999) who find that the low correlation exacerbates a lot of asset pricing puzzles. More recently, Albuquerque, Eichenbaum, and Rebelo

(2012) introduce a demand shock to a representative agent's rate of time preference to account for the equity premium, bond term premia, and the correlation puzzle.

In the literature, most studies focus on the qualitative property of predictability, and only a few studies explicitly explore the quantitative magnitude allowed by asset pricing models. Hansen and Singleton (1983) seem to be the first to consider this problem exclusively and find that the predictability of stock returns are proportional to the predictability of the consumption growth rate. The weak predictability of the consumption growth rate implies that stock returns are almost unpredictable. Ferson and Harvey (1991) and Ferson and Korajczyk (1995) find that the multi-beta model explain a large fraction of return predictability. Kirby (1998) develops a formal test and finds that none of the recognized models can deliver sufficient predictability to accommodate the empirical pattern. Bansal, Kiku, and Yaron (2012) show that the dividend-price ratio can only generate a marginal degree of predictability with the long-run risk model. de Roon and Szymanowska (2012) show that transaction costs rather than short sale constraint can reconcile Kirby (1998). All these papers assume specific utility functions and so the results vary with different models and parameter specifications.

Ross (2005) proposes an upper bound on the predictive R^2 and finds that predictability is consistent with asset pricing models. Zhou (2010) proposes a tighter bound and shows that most predictors generate larger predictive R^2s than his bound if the SDF is driven by the consumption growth rate. This paper is closely related to Ross (2005) and Zhou (2010) but departs from them in four aspects. First, I propose a new bound with the market Sharpe ratio rather than the maximum risk aversion, giving a new choice to those who are uncertain about risk aversion. Second, Ross (2005) implicitly assumes that the correlation between the forecasted excess return and the state variables is 1, making it a special case of my bounds. Third, Zhou

(2010) uses the correlation between the state variables and the default SDF,[3] while my bounds use the correlation between the excess return and the state variables, thereby providing some insights on cross-sectional predictability as to why some assets are more predictable than others. Fourth and more important, my bounds use conditional information explicitly and are much tighter than Ross (2005) and Zhou (2010). When the market portfolio is included in the state variables of the SDF, the bounds in Ross (2005) and Zhou (2010) lose the power to bind the predictive R^2 while my bounds still work well.

The rest of the paper is organized as follows. Section 2 shows how the predictive R^2 can be bounded above by a specific SDF. Section 3 presents two semi-parametric bounds when the SDF are bounded above by the maximum risk aversion or by the market Sharpe ratio. The results of applying these two bounds to return predictability are reported in Section 4. Finally, Section 5 summarizes and concludes.

1.2 Model

In this section, I show how to connect the predictive regression with asset pricing models and then derive an upper bound on the predictive R^2 with the variance of the stochastic discount factor (SDF).

1.2.1 Asset pricing model

The central idea of finance theory is that the price of any asset is uniquely determined by a Euler equation that satisfies

$$\mathrm{E}[m_{t+1}r_{j,t+1}|I_t] = 0, \ j = 0, 1, \cdots, N, \tag{1.1}$$

[3]See equation (1.2) in Section 2.

where m_{t+1} is the SDF, $r_{j,t+1}$ is the return of asset j in excess of the risk-free rate $R_{f,t}$.[4] Equation (1.1) says that the risk-adjusted return process defined by the product of the excess return $r_{j,t+1}$ and the SDF m_{t+1} is a martingale and is unpredictable using any information contained in I_t. This equation is so general that it can accommodate the case when the return itself is predictable, which does not necessarily conflict with the market efficiency hypothesis. The only case of $r_{j,t+1}$ being unpredictable is when m_{t+1} is constant over time.

According to Cochrane (2005), any asset pricing model is a particular specification of m_{t+1}. One default SDF, which satisfies (1.1) and prices the $N+1$ assets, is given by

$$m_{0,t+1} = R_{f,t}^{-1} + (1_N - R_{f,t}^{-1}\mu)'\Sigma^{-1}(R_{t+1} - \mu), \tag{1.2}$$

where R_{t+1} is the $N \times 1$ vector of gross returns on the N risky assets with mean μ and covariance Σ, and 1_N is an N-dimensional vector of ones. I assume that μ is not proportional to 1_N and the N risky assets are not redundant.

In what follows, when it is not necessary to be explicit about the difference between assets, I will suppress the subscripts and just write r_{t+1} rather than $r_{j,t+1}$.

1.2.2 Predictive regression

Predictive regression is widely used in the study of return predictability, and is expressed as

$$r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}, \tag{1.3}$$

[4] I use $R_{f,t}$ rather than $R_{f,t+1}$ since it is known at the beginning of the return period.

where z_t is a predictive variable known at the end of period t. The degree of predictability is measured by the predictive R^2,

$$R^2 = \frac{\mathrm{Var}(\alpha + \beta z_t)}{\mathrm{Var}(r_{t+1})}. \tag{1.4}$$

When $R^2 > 0$, r_{t+1} can be forecasted by z_t. Otherwise, it cannot be forecasted. Following this idea, numerous variables have been identified as predictors. Ludvigson and Ng (2007) and Goyal and Welch (2008) provide a comprehensive list of predictors.

1.2.3 Bound on R^2

Whether return predictability can be explained by asset pricing models is equivalent to ask whether the predictive R^2 in (1.4) can be derived from (1.1). For easy of exposition, I follow Balduzzi and Kallai (1997) and normalize the SDF

$$\tilde{m}_{t+1} = \frac{m_{t+1}}{\mathrm{E}(m_{t+1})}$$

such that $E(\tilde{m}_{t+1}) = 1$ and $\mathrm{E}(\tilde{m}_{t+1} r_{t+1}) = 0$. With a little abuse of notation, I still call this normalized SDF as the SDF in the sequel.

I assume that the predictor z_t in (1.3) has a mean zero and variance one throughout the paper. Following Kirby (1998) and Ferson and Siegel (2003), I multiply the pricing equation (1.1) by z_t in both sizes and apply the law of iterated expectations to obtain

$$\mathrm{E}(\tilde{m}_{t+1} r_{t+1} z_t) = 0, \tag{1.5}$$

which can be rewritten as

$$\mathrm{Cov}(r_{t+1}, z_t) = -\mathrm{Cov}(\tilde{m}_{t+1}, r_{t+1} z_t). \tag{1.6}$$

Since $\text{Cov}(r_{t+1}, z_t) = \text{E}(r_{t+1}z_t)$, equality (1.6) says that the expected excess return with z_t units of investment in the asset r_{t+1} is equal to the negative covariance between the SDF and the realized excess return of the investment. In other words, any dynamic trading strategy that exploits the predictability of r_{t+1} must be priced by the SDF.

Recall that $\text{Var}(z_t) = 1$ and $\beta = \text{Cov}(r_{t+1}, z_t)$. Combining (1.4) and (1.6) gives

$$R^2 = \frac{\text{Var}(\alpha + \beta z_t)}{\text{Var}(r_{t+1})} = \frac{\beta^2}{\text{Var}(r_{t+1})} = \frac{\text{Cov}^2(r_{t+1}, z_t)}{\text{Var}(r_{t+1})} = \frac{\text{Cov}^2(\tilde{m}_{t+1}, r_{t+1}z_t)}{\text{Var}(r_{t+1})}. \tag{1.7}$$

If an asset pricing model is true, i.e., the model can match the empirical evidence, the last equality of (1.7) should always hold. To test this hypothesis, Kirby (1998) uses the generalized method of moments (GMM) and finds that the R^2 calculated from the last equality of (1.7) is much smaller than that in (1.4) for established consumption- and factor-based asset pricing models at that moment. Therefore, he concludes that return predictability is inconsistent with what is expected. Kirby's method is parametric and depends on the specification of $\tilde{m}_{t+1}$. Since Kirby (1998), new asset pricing models, such as the habit formation model, the long-run risk model and the rare disaster model, have been developed. This implies that we need retest the conclusion of Kirby (1998) when a new model is proposed.

I solve Kirby's problem from another perspective by developing an upper bound on (1.7) which can serve as a benchmark for evaluating forecasts. Statistically, the larger the predictive R^2, the higher the degree of predictability. Both financial economists and investment practitioners have paid a lot attention in the past four decades in searching for variables that can produce a better R^2. This raises two issues. First, without theoretical guidance on the R^2 permitted by asset pricing models, an investor will never know whether the used predictor is the best one. Second, given hundreds of predictors that have been identified, how does an investor use them in investment decision making? Should an investor utilize all the possible predictors or just choose a subset of them?[5] An investor cannot run two million regressions and then decides which one is the best. However, if an investor knows the maximum predictability, he can stop searching when a predictor generates an R^2 that achieves or is close to the theoretical upper bound. Moreover, an investor can directly exclude those variables with R^2s much less than the bound.

Following Kan and Zhou (2006), I impose one structure on the SDF: $\tilde{m}_{t+1} = \tilde{m}(x_{t+1})$ is a function of a set of observable state variables x_{t+1}. This structure remains general enough to accommodate many asset pricing models. For example, factor-based models, such as the capital asset pricing model (CAPM) and the Fama-French three-factor model, specify $\tilde{m}_{t+1}$ as a linear function of factors. In consumption-based models, the state variables are the surplus consumption ratio and the consumption growth rate in the habit formation model (Campbell and Cochrane, 1999; Kan and Zhou, 2006), are the risk-free rate, the dividend-price ratio, and the consumption growth rate in the long-run risk model (Constantinides and Ghosh, 2011), and are the consumption growth rate and the dividend-price ratio in the rare disaster model (Wachter, 2012). In addition, Bansal and Viswanathan (1993) specify the SDF as a nonlinear function of the market portfolio, the Treasury bill yield, and the term spread. Dittmar (2002) specifies the SDF as a cubic function of aggregate wealth. Aït-Sahalia and Lo (2000) project the SDF onto stock returns to obtain an observable kernel, thereby avoiding the use of the consumption growth rate.

Now I am in a position to present the following proposition to explain that the predictive R^2 can be bounded above.

Proposition 1 *Suppose that the SDF* $\tilde{m}_{t+1} = \tilde{m}(x_{t+1})$ *is a function of* K-*dimensional state variable* x_{t+1} *and* $\mathrm{E}(\varepsilon_{t+1}|x_{t+1}) = 0$ *in the regression* $r_{t+1}z_t =$

[5]This echoes Cochrane (2011) who asks how multivariate information affects the understanding of price movements.

$a + b'x_{t+1} + \varepsilon_{t+1}$. *Then,*

$$R^2 \leq \phi^2_{x,rz} \mathrm{Var}(\tilde{m}_{t+1}), \tag{1.8}$$

where

$$\phi^2_{x,rz} = \frac{\rho^2_{x,rz} \mathrm{Var}(r_{t+1} z_t)}{\mathrm{Var}(r_{t+1})}, \tag{1.9}$$

and

$$\rho^2_{x,rz} = \frac{\mathrm{Cov}(x_{t+1}, r_{t+1} z_t)' \mathrm{Var}^{-1}(x_{t+1}) \mathrm{Cov}(x_{t+1}, r_{t+1} z_t)}{\mathrm{Var}(r_{t+1} z_t)}. \tag{1.10}$$

The formal proof is provided in the paper's Appendix. Here I give a simplified proof showing how the predictive R^2 can be bounded by the variance of the SDF. This is the key to restrict the regression analysis of return predictability by asset pricing models. Suppose x_{t+1} and $r_{t+1} z_t$ are jointly normally distributed conditional on time t. From (1.7), I have

$$\begin{aligned}
R^2 &= \frac{\mathrm{Cov}^2(\tilde{m}_{t+1}, r_{t+1} z_t)}{\mathrm{Var}(r_{t+1})} = \frac{\left[\mathrm{Cov}(x_{t+1}, r_{t+1} z_t)' \mathrm{Var}^{-1}(x_{t+1}) \mathrm{Cov}(\tilde{m}_{t+1}, x_{t+1})\right]^2}{\mathrm{Var}(r_{t+1})} && (1.11) \\
&\leq \left[\mathrm{Cov}(x_{t+1}, r_{t+1} z_t)' \mathrm{Var}^{-1}(x_{t+1}) \mathrm{Cov}(x_{t+1}, r_{t+1} z_t)\right] \\
&\quad \times \frac{\left(\mathrm{Cov}(\tilde{m}_{t+1}, x_{t+1})' \mathrm{Var}^{-1}(x_{t+1}) \mathrm{Cov}(\tilde{m}_{t+1}, x_{t+1})\right)}{\mathrm{Var}(r_{t+1})} && (1.12) \\
&= \frac{\rho^2_{x,rz} \mathrm{Var}(r_{t+1} z_t) \mathrm{Cov}(\tilde{m}_{t+1}, x_{t+1})' \mathrm{Var}^{-1}(x) \mathrm{Cov}(\tilde{m}_{t+1}, x_{t+1})}{\mathrm{Var}(r_{t+1})} && (1.13) \\
&\leq \frac{\rho^2_{x,rz} \mathrm{Var}(r_{t+1} z_t) \mathrm{Var}(\tilde{m}_{t+1})}{\mathrm{Var}(r_{t+1})} = \phi^2_{x,rz} \mathrm{Var}(\tilde{m}_{t+1}), && (1.14)
\end{aligned}$$

where (1.11) uses Stein's Lemma, which separates the underlying stochastic structure between r_{t+1} and x_{t+1} from the distortion of $\tilde{m}(\cdot)$ (Furman and Zitikis, 2008).

Inequalities (1.12) and (1.14) use the Cauchy-Schwarz inequality. This completes the proof of (1.8).

Equality (1.11) shows that the covariance $\mathrm{Cov}(\tilde{m}_{t+1}, r_{t+1}z_t) = \mathrm{E}(r_{t+1}z_t)$ is mainly dependent on two parts: one is covariance between the excess return with z_t units of investment in r_{t+1} and the state variable x_{t+1}, $\mathrm{Cov}(x_{t+1}, r_{t+1}z_t)$, and the other is the covariance between the SDF and the state variable, $\mathrm{Cov}(\tilde{m}_{t+1}, x_{t+1})$. In the asset pricing literature, expected returns are expressed by the covariance of the returns and the SDF. The failure of asset pricing models in explaining return puzzles or anomalies is usually attributed to the inability of preferences in capturing investor's behaviors. For this reason, many different preferences have been proposed over the past three decades. With a "moment-matching" approach (calibrating parameters with real data and investigating if the estimated parameters make sense or if what the model implies with given parameters is consistent with return moments), one specific utility is usually successful in explaining one or several puzzles, but not all of them. Proposition 1, however, shows that the failure of asset pricing models may be due to the insufficient state variables x_{t+1} rather than the utility functions $\tilde{m}(\cdot)$. The covariance between the return and the SDF may blur the main reason of the inability of asset pricing models.

Proposition 1 imposes a slightly stronger assumption

$$\mathrm{E}_t(u_{t+1}|x_{t+1}) = 0, \tag{1.15}$$

rather than the typical $\mathrm{E}_t(u_{t+1}) = 0$ and $\mathrm{Cov}_t(u_{t+1}, x_{t+1}) = 0$. One extreme case is $u_{t+1} = 0$ when $r_{t+1}z_t$ is the same as x_{t+1} and can be fully projected on x_{t+1}. Actually, the two assumptions are equivalent if the excess return $r_{t+1}z_t$ and the state variable x_{t+1} are jointly elliptically, conditionally distributed (Muirhead, 1982).

The bound in (1.8) is an improvement over the bound of Ross (2005) who finds

$$R^2 \leq \mathrm{Var}(\tilde{m}_{t+1}). \tag{1.16}$$

This improvement is due to the fact that I use the information of x_{t+1} in $\tilde{m}_{t+1}$. Comparing (1.8) and (1.16), Ross (2005) takes the extreme possibility that the state variable and the excess return are perfectly correlated. This is obviously not the case in the real equity market. Suppose that the correlation between the consumption growth rate and the market portfolio is 0.2 and that the SDF is driven by the consumption growth rate, bound (1.8) will be at least 25 times tighter than that derived by Ross (2005). Cochrane (2005) notes the fact that the low correlation between the consumption growth rate and stock returns exacerbates the risk premium puzzle, but does not develop this point with respect to return predictability. In summary, Ross' bound imposes almost no structure on the SDF other than the law of one price. The consequence is that it can deliver an R^2 bound that is applicable for all SDFs. However, the cost is that the bound is too loose to be meaningful in practice. Over my sample period, Ross' bound is as large as 4.78%, but the predictive R^2 in the existing literature is less than 1% in general. To the best of my knowledge, no single predictor can produce an R^2 of 4.78%.

The bound in (1.16) holds with respect to the default SDF, i.e.,

$$R^2 \leq \mathrm{Var}(\tilde{m}_{0,t+1}), \tag{1.17}$$

which can be tightened by Kan and Zhou (2007) who show that

$$\mathrm{Var}(\tilde{m}_{0,t+1}) \leq \rho^2_{x,\tilde{m}_0} \mathrm{Var}(\tilde{m}(x_{t+1})), \tag{1.18}$$

where $\rho_{x,\tilde{m}_0}$ is the multiple correlation between the state variable x_{t+1} and the default SDF. Combining these two inequalities, Zhou (2010) gives the following upper bound

$$R^2 \leq \rho^2_{x,\tilde{m}_0} \mathrm{Var}(\tilde{m}_{t+1}), \tag{1.19}$$

which is apparently tighter than Ross (2005) bound.

An interesting question at this point is whether the bound in (1.8) is tighter than (1.19). This is equivalent to exploring whether $\phi^2_{x,rz} < \rho^2_{x,m_0}$. While there is no analytical relation between them, empirical applications will show that $\phi^2_{x,rz}$ is always smaller than ρ^2_{x,m_0}.

It is important to highlight the implication of the proposed bound of the predictive R^2 on cross-sectional return predictability. In the literature, a large number of papers find that return predictability exists and varies across cross-sectional portfolios sorted by market capitalization (Ferson and Harvey, 1991; Kirby, 1998), book-to-market ratio (Ferson and Harvey, 1991), industry (Ferson and Harvey, 1991), and volatility (Han, Yang and Zhou, 2012). Proposition 1 says that the maximum predictability of any asset is directly determined by the parameter, $\phi^2_{x,r\tilde{z}}$, in the upper bound of R^2. An asset is allowed to be more predictable if it has a higher correlation with the state variables of the SDF, regardless of the specification of the SDF.

1.3 Upper Bound on $\mathrm{Var}(\tilde{m}_{t+1})$

Inequality (1.8) provides an upper bound on the predictive R^2. However, the SDF is model-specific and unobservable. The goal of this section is to develop an upper bound on $\mathrm{Var}(\tilde{m}_{t+1})$ that is observable and model-free.

There are two approaches for the SDF specification proffered in the literature, the *absolute* approach and the *relative* approach (Cochrane and Saá-Requejo, 2000). The absolute approach makes explicit assumptions about the representative investor's preference and endowment. Under these assumptions, the SDF is uniquely, endogenously determined by the form of preferences. Although this approach is precise, it is sensitive to model and parameter misspecification errors. The relative approach assumes the existence of a set of basis assets and the absence of arbitrage opportunities, restricting the set of the SDF to those that can correctly price the basis assets in the economy and assigning positive values to payoffs in every state. Without resorting to preferences or endowments, this approach is exogenously specified and robust to model specification. The drawback is that there are usually infinite SDFs that can price the basis assets. This implies that it is difficult to choose an correct asset pricing model when all the SDFs produce the same price.

To tackle this challenge, Cochrane and Saá-Requejo (2000) and Ross (2005) propose to integrate the absolute and the relative approaches by restricting the SDF to an economically meaningful set. In contrast to Hansen and Jagannathan (1991) who restrict the SDF with a lower bound, I assume an upper bound on the SDF volatility to exclude the opportunities that may generate arbitrages.

1.3.1 Bound $\mathrm{Var}(\tilde{m}_{t+1})$ with relative risk aversion

Ross (2005) shows that, in an incomplete market, if all investors are bounded above by a maximum risk aversion, the set of the SDFs can be restricted by the marginal investor's SDF.

Lemma 1 (Ross, 2005) *If a utility function, $U(w)$, is bounded above in the relative risk aversion by a utility function $V(w)$, i.e., the risk aversion of $U(w)$ is less than that of $V(w)$, then*

$$\mathrm{Var}(\tilde{m}_U) \leq \mathrm{Var}(\tilde{m}_V),$$

where $\tilde{m}_U$ and $\tilde{m}_V$ are the corresponding SDFs. Moreover, if $V(w)$ is a constant relative risk aversion utility function with risk aversion γ ($\gamma \neq 1$) and the optimal *wealth is lognormally distributed such as* $\log w \sim N(\mu_w, \sigma_w^2)$*, then*

$$\text{Var}(\tilde{m}_U) \leq \gamma^2 \sigma_w^2.$$

This lemma says that the variance of any SDF can be bounded above by a maximum risk aversion.

Applying Lemma 1, I present the first semi-parametric bound in this paper as follows.

Proposition 2 *Under conditions of Propositions 1 and Lemma 1, if investors are bounded above by the maximum risk aversion* γ*, the upper bound of the predictive* R^2 *is*

$$R^2 \leq \bar{R}_{RA}^2 = \phi_{x,rz}^2 \gamma^2 \sigma_w^2. \tag{1.20}$$

1.3.2 Bound $\text{Var}(\tilde{m}_{t+1})$ with market Sharpe ratio

Instead of maximum risk aversion, Ross (1976) advocates using the market Sharpe ratio to restrict the variability of the SDF. The intuition is that a high Sharpe ratio is not an arbitrage opportunity or a violation of the law of one price, but extremely high Sharpe ratios are unlikely to persist. In particular, Ross (1976) bounds the asset pricing theory residuals by assuming that no portfolio can have more than twice the market Sharpe ratio. With this idea, Cochrane and Saá-Requejo (2000) use the market Sharpe ratio to bound option prices when either market frictions or non-market risks violate simple arbitrage pricing. That is,

$$\text{Std}(\tilde{m}_{t+1}) \leq h \cdot \text{SR}(r_{S\&P500}), \tag{1.21}$$

where h is a parameter chosen by the marginal investor. Cochrane and Saá-Requejo (2000) choose $h = 2$ as the threshold for "good deals".

Proposition 3 *Under conditions of Propositions 1, if the volatility of the SDF is bounded by the market Sharpe ratio as in (1.21), the upper bound of the predictive* R^2 *is*

$$R^2 \leq \bar{R}^2_{SR} = \phi^2_{x,rz} \cdot h^2 \cdot \mathrm{SR}^2(r_{S\&P500}). \tag{1.22}$$

It is important to point out that the maximum risk aversion γ or h are the central parameters that a user must input to the calculation. When the upper bound of the SDF's volatility is violated, Shanken (1992) calls there have some "approximate arbitrage" opportunities. Ledoit (1995) calls a high Sharpe ratio a "δ arbitrage" that should be ruled out. Also, there are other ways to bound the volatility of $\tilde{m}_{t+1}$. For example, Bernardo and Ledoit (2000) bound the SDF as $a \leq \tilde{m}_{t+1} \leq b$, where a and b are two positive and finite real parameters. By applying the Grüss' inequality, one immediately has $\mathrm{Var}(\tilde{m}_{t+1}) \leq \frac{(b-a)^2}{4}$ for any distribution of $\tilde{m}_{t+1}$.

1.4 Empirical Results

This section explores empirically whether the predictive R^2s of predicting excess returns on the market portfolio and cross-sectional portfolios are smaller than the upper bounds derived from asset pricing models.

1.4.1 Data

The main data set used in this paper is from Goyal and Welch (2008) and the Ken French data library, spanning 1959:01-2010:12,[6] where the sources are described in

[6] I thank Amit Goyal and Ken French for making the data available.

detail. The excess return of the market portfolio is the gross return on the S&P 500 (including dividends) minus the gross return on a risk-free treasury bill. As discussed by Ferson and Korajczyk (1995), in the context of this paper, it is not appropriate to use continuously compounded returns, which are commonly used in the literature of return predictability. The basic pricing equation says that the expected returns are equal to the conditional covariances of returns with the marginal utility for wealth, which depends on the simple arithmetic return of the optimal portfolio. Moreover, continuously compounded portfolio returns are not the portfolio-weighted average of the compounded returns of the component securities. For these reasons, I use simple arithmetic returns.[7]

Ten popular economic variables in Goyal and Welch (2008) are used as predictors:

1. *Dividend-price ratio* (d/p): a 12-month moving sum of dividends paid on the S&P 500 index divided by the S&P 500 index;
2. *Earning-price ratio* (e/p): a 12-month moving sum of earnings on the S&P 500 index divided by the S&P 500 index;
3. *Dividend yield* (dy): a 12-month moving sum of dividends divided by the lagged S&P 500 index;
4. *Treasury bill rate* (tbl): the 3-month Treasury bill (secondary market) rate;
5. *Default yield spread* (dfy): the difference between BAA and AAA-rated corporate bond yields;
6. *Term spread* (tms): the difference between the long-term yield on government bonds the Treasury bill rate;
7. *Net equity expansion* (ntis): ratio of a twelve-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks;

[7] The predictive R^2s with compounded returns are generally larger than simple arithmetic returns. The results are available upon request.

8. *Inflation* (infl): the Consumer Price Index;

9. *Long-term return* (ltr): return on long-term government bonds;

10. *Equity risk premium volatility* (rvol): moving standard deviation of the monthly returns on the S&P 500 index (Mele, 2007):

$$\hat{\sigma}_t = \sqrt{\frac{\pi}{2}} \sum_{i=1}^{12} \frac{|r_{t+1-i}|}{\sqrt{12}}.$$

    I use this volatility measure rather than the realized volatility in Goyal and Welch (2008) to avoid two severe outliers in Octobers of 1987 and 2008.

To calculate the upper bounds of the predictive R^2, I need state variables. The most popular state variable in consumption-based asset pricing models is the consumption growth rate: the percentage change in the seasonally adjusted, aggregate, real per capita consumption expenditures on nondurable goods and services. The data are reported by the Bureau of Economic Analysis (BEA). In addition, I consider the linear factor models where the market index or the Fama-French three-factors are used as the state variables. I also use data on ten size portfolios, ten book-to-market portfolios, ten momentum portfolios, and ten industry portfolios, for cross-sectional predictability.

What is the reasonable maximum risk aversion has been and will continue to be a debate for a long time, although researchers admit that it should not be large. Mehra and Prescott (1985) argue that a reasonable upper bound of risk aversion is around 10. Ross (2005) uses the insurance premium to explain that a value of 5 is large enough. Barro and Ursúa (2012) think that "a γ [risk aversion] of 6 seems implausibly high." Empirically, Guiso, Sapienza and Zingales (2011) find that the average risk aversion increases from 2.85 before the 2008 crisis to 3.27 after the collapse of the financial market. Paravisini, Rappoport and Ravina (2012) estimate the risk aversion from investors' financial decisions and find that the average risk aversion is 2.85 with a median of 1.62. I follow Ross (2005) by setting the maximum risk aversion to be 5. Also, in the application, I assume that the optimal wealth for the marginal investor who has the maximum risk aversion is the market portfolio. During the sample period, the market portfolio has an annual risk premium 5.31% and a volatility 15.44%.

When the market Sharpe ratio is used to bound the predictive R^2, I follow Ross (1976) and Cochrane and Saá-Requejo (2000) by setting h equal to 2. I find that the upper bound with this value is close to that with the maximum risk aversion bound with a value of 5. This result indirectly supports Ross (2005) that the upper bound of risk aversion should not exceed 5.

1.4.2 Estimation and test

The parameters to calculate the predictive R^2 and its upper bounds involve only the mean and covariance of $y_{t+1} = (r_{t+1}, z_t, r_{t+1}z_t, x'_{t+1})'$, where x_{t+1} could be multi-dimensional. The moment conditions are

$$h(y_{t+1}, \theta) = \begin{pmatrix} y_{t+1} - \mu_y \\ y_{t+1}y'_{t+1} - (\Sigma_y + \mu_y\mu'_y) \end{pmatrix}, \tag{1.23}$$

where $\mu_y = E(y_{t+1})$ and $\Sigma_y = Cov(y_{t+1})$. The econometric specification in (1.23) is exactly identified, the GMM estimator of $\theta = (\mu'_y, \Sigma_y)$ is the value that sets $1/T \sum_{t=1}^{T} h(y_{t+1}, \theta)$ equal to zero.

The distribution of $\hat{\theta}$ takes the form

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, S), \tag{1.24}$$

where $S = \sum_{j=-\infty}^{\infty} E[h(y_{t+1}, \theta)h(y_{t+1-j}, \theta)']$.

We use a Wald test to evaluate whether $R^2 \leq \bar{R}^2_{RA}$ or $\bar{R}^2_{SR}$. This is equivalent to a one-sided test for $g(\theta_{RA}) = 0$ or $g(\theta_{Sh}) = 0$, where θ_{RA} and θ_{Sh} are the parameters used in $g(\theta_{RA}) = R^2 - \bar{R}^2_{RA}$ for the bound with risk aversion and $g(\theta_{Sh}) = R^2 - \bar{R}^2_{SR}$ for the bound with Sharpe ratio. Let Σ_{RA} and Σ_{Sh} be the corresponding covariances of θ_{RA} and θ_{Sh}. The Wald statistic is

$$W_{RA} = Tg(\hat{\theta}_{RA})\left[\frac{dg}{d\theta_{RA}}\hat{\Sigma}_{RA}\frac{dg}{d\theta_{RA}}\right]^{-1} g(\hat{\theta}_{RA}) \xrightarrow{d} \chi^2(1) \tag{1.25}$$

for the bound with risk aversion, and

$$W_{Sh} = Tg(\hat{\theta}_{Sh})\left[\frac{dg}{d\theta_{Sh}}\hat{\Sigma}_{Sh}\frac{dg}{d\theta_{Sh}}\right]^{-1} g(\hat{\theta}_{Sh}) \xrightarrow{d} \chi^2(1) \tag{1.26}$$

for the bound with Sharpe ratio.

The approach here is slightly different from the typical GMM estimation and testing by imposing the constraint $R^2 = \bar{R}^2_{RA}$ or $\bar{R}^2_{SR}$ in the econometric specification directly. With the property of GMM, the two approaches are asymptotically equivalent. The choice of this paper makes it easy to compare the difference between the predictive R^2 and the theoretical upper bounds apparently.

1.4.3 R^2 bounds on market portfolio predictability

Table 1.1 reports the predictive R^2 and its bounds for the regression model, $r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}$, where r_{t+1} is the excess return of the market portfolio and z_t is the predictor given in the table's first column. The value of R^2 and its bounds are all in percentage points. The state variable is the consumption growth rate and the default SDF m_0 is constructed by the market portfolio.[8] Panel A shows the results when the maximum risk aversion is 5. The predictive R^2s are given in the second column, which range from 0.04% for the net equity issues (ntis) to 1.23% for the equity risk premium volatility (rvol). Positive R^2s suggest that the excess return of the market portfolio

[8]Other portfolios, such as the Fama-French three factors or Fama-French 25 size book-to-market portfolios, can be easily used to construct m_0. This will change the multiple correlation ρ^2_{x,m_0}, but the change is very small. The results are available upon request.

is predictable and the degree of predictability varies across predictors. The upper bound of Ross (2005), $\bar{R}^2_{Ross}$, is 4.78% reported in Column 3 regardless of what the predictor is. Because this bound exceeds any R^2 in Column 2, time-varying expected return appears to be a perfect explanation of return predictability. To the best of my knowledge, however, there is no single predictor in the literature generating an R^2 as large as 4.78% with monthly data. The reason is that Ross' bound implicitly assumes a correlation of 1 between the excess return and the consumption growth rate. Hence, it is too loose to be meaningful.

Column 4 reports the correlation between the state variable and the default SDF and Column 5 reports the bound developed by Zhou (2010). Since the correlation is 0.17, the bound in Zhou (2010) is 0.13%, and thereby improves approximately 37 times relative to Ross (2005). Out of ten predictors, eight exhibit significantly higher predictive R^2s than this bound. The two exceptions are the earnings-price ratio (e/p) and the net equity issues (ntis).

Column 6 shows that the correlations between the state variable and the excess returns with trading strategy z_t. Surprisingly, all the correlations are pretty small and range from 0.02 to 0.06. Recall that the key parameter in the upper bounds in (1.20) and (1.22) is $\phi^2_{x,rz} = \rho^2_{x,rz}\text{Var}(r_{t+1}z_t)/\text{Var}(r_{t+1})$. $\text{Var}(r_{t+1}z_t)/\text{Var}(r_{t+1})$ is larger than one but less than 4 for any z_t of the ten predictors. This implies that small value of $\rho_{x,rz}$ makes the upper bounds of the predictive R^2 small. Actually, both bounds with the maximum risk aversion and the market Sharpe ratio are approximately zero. As a result, the proposed bounds are significantly less than the predictive R^2s. The low bound of R^2 is consistent with Hansen and Singleton (1983) who explore the joint dynamics of stock returns and consumption growth and find that the predictability of stock returns is proportional to that of consumption growth. The weak predictability of the consumption growth rate in turn implies that stock returns are almost unpredictable. This result is confirmed by Kirby (1998) with a formal GMM test.

Panel B considers the case when the maximum risk aversion is 10. In this case, the upper bounds R^2 can be obtained by multiplying the bounds in Panel A by 4. Due to the small value of $\rho_{x,rz}$, the increase in the risk aversion does not change the upper bounds significantly. This insensitivity of the R^2 bounds implies that changing the maximum risk aversion is not promising to reconcile the violations of the bounds. On the other hand, since the bound with the maximum risk aversion of 5 is close to the bound with the market Sharpe ratio (as shown in Panel A), I believe that 5 is a reasonable upper bound of risk aversion. In this sense, I will report results with the maximum risk aversion of 5 in the sequel.

One may be curious that the results in Table 1.1 are only valid to consumption-based asset pricing models since I only consider the consumption growth rate as the state variable of the SDF. In the literature, there are many factor-based asset pricing models. Table 1.2 reports the bounds with alternative state variables. In particular, Panel A assumes that the state variable is the market portfolio (the state variable of CAPM) and Panel B considers the Fama-French three factors. With these two cases, since the correlation ρ_{x,m_0} is approaching one, the bound of Zhou (2010) reduces to Ross (2005) and exceeds the predictive R^2s. However, the bounds proposed in this paper still work well. When the state variable is the market portfolio, eight predictors violate the bounds, either with the maximum risk aversion or the market Sharpe ratio. When the state variables are the Fama-French three factors, six predictors violate the bounds with the maximum risk aversion and seven violate the bounds with the market Sharpe ratio. When the momentum factor is added to the Fama-French three factors, the upper bounds of R^2 do not change significantly, and therefore, to conserve space, the results are not reported.

In summary, the predictive R^2s from the predictive regression are larger than the maximum predictability permitted by asset pricing models. Since $\text{Var}(r_{t+1}z_t)/\text{Var}(r_{t+1})$ has nothing with the asset pricing model, the failure to explain predictability is clearly due to the correlation between the excess return and the state variables of the SDF, the maximum risk aversion, or the volatility of the marginal investor' wealth. In the bound with the market Sharpe ratio, the parameter of the maximum risk aversion is replaced by the parameter h which is the threshold of excluding arbitrage opportunities. Among them, I have already considered the case of a risk aversion of 10. The marginal investor's wealth is assumed to be the market portfolio, which may be more volatile than the real wealth with other non-financial assets. Therefore, the only reason is that the correlation between the excess return and the state variables in the SDF is too low (as shown in Column 6 in Tables 1.1 and 1.2). This explanation is obvious when the state variables are the Fama-French three factors, which have a much higher correlation with the excess returns and so generate higher bounds on the predictive R^2s. The findings of this section suggest that the state variables are more important than investor's preferences in explaining return predictability. This explanation is consistent with Cochrane and Hansen (1992), Campbell and Cochrane (1999), and Albuquerque, Eichenbaum, and Rebelo (2012) who attribute the failure of consumption-based asset pricing models to the low correlation between asset returns and the state variables of the SDF.

1.4.4 R^2 bounds with recently developed models

This subsection discusses whether the habit formation model, the long-run risk model, or the rare disaster model can explain the predictability of the market portfolio when the state variables are from one of these three asset pricing models.

Habit formation

The state variables in the habit formation model are the consumption growth rate and the surplus consumption ratio s_t that is unobservable since the habit level is latent. I follow Campbell and Cochrane (1999) by extracting s_t from the model and calculate the multiple correlation between the state variables $x_t = (\triangle c_t, \triangle s_t)$ and the excess return with z_t units of investment in the market portfolio. The results are reported in Panel A of Table 1.3. With an additional state variable $\triangle s_t$, the correlation between the excess return and the state variables approximately doubles relative to the traditional Consumption-based models. However, it is still very small. Nine out of ten correlations (since different predictor implies different correlation) are less than 0.1. As a result, both bounds with the maximum risk aversion and the market Sharpe ratio are still close to zero, significantly less than the predictive R^2s.

Long-run risk

The long-run risk model focuses on the low-frequency properties of the time series of dividends and aggregate consumption, and can explain simultaneously the equity risk premium puzzle, the risk-free rate puzzle, and the high level of market volatility. The key assumptions in the long-run risk model are that the consumption growth rate and the dividend growth rate follow the following joint dynamics:

$$\begin{aligned}
\triangle c_{t+1} &= \mu_c + \mu_{c,t} + \sigma_t \epsilon_{c,t+1}, \\
\mu_{c,t+1} &= \rho_\mu \mu_{c,t} + \psi_c \sigma_t \epsilon_{\mu,t+1}, \\
\sigma^2_{t+1} &= (1-\nu)\bar{\sigma}^2 + \nu \sigma_t^2 + \sigma_w \epsilon_{\sigma,t+1}, \\
\triangle d_{t+1} &= \mu_d + \phi \mu_{c,t} + \phi \sigma_t \epsilon_{d,t+1},
\end{aligned}$$

where c_{t+1} is the log aggregate consumption and d_{t+1} is the log dividends. The shocks $\epsilon_{c,t+1}$, $\epsilon_{\mu,t+1}$, $\epsilon_{\sigma,t+1}$, and $\epsilon_{d,t+1}$ are assumed to be i.i.d. normally distributed.[9]

With log-affine approximation, the SDF is

$$\log m_{t+1} = A_0 + A_1\mu_{c,t} + A_2\sigma_t^2 + A_3\triangle c_{t+1} + A_4\mu_{c,t+1} + A_5\sigma_{t+1}^2, \tag{1.27}$$

where $A_0, \cdots, A_5$ are parameters to be estimated. There are two latent state variables in the SDF, the conditional mean of the consumption growth rate y_t and the conditional variance of its innovation σ_t^2, which are difficult to be measured in the data. Motivated by Dai and Singleton (2000), Constantinides and Ghosh (2011) bridge this gap and find that these two latent variables can be projected on the log risk-free rate $r_{f,t}$ and the log dividend-price ratio dp_t:

$$\begin{aligned}
\mu_{c,t} &= \alpha_0 + \alpha_1 r_{f,t} + \alpha_2 dp_t, \\
\sigma_t^2 &= \beta_0 + \beta_1 r_{f,t} + \beta_2 dp_t.
\end{aligned}$$

In this way, the log SDF is an affine function of the log risk-free rate, the log dividend-price ratio, and the consumption growth rate:

$$\log m_{t+1} = B_0 + B_1 r_{f,t} + B_2 dp_t + B_3 r_{f,t+1} + B_4 dp_{t+1} + B_5\triangle c_{t+1}.$$

Panel B of Table 1.3 shows the R^2 bounds when the state variables in the SDF are

$$x_{t+1} = (\triangle c_{t+1}, r_{f,t+1}, dp_{t+1})'. \tag{1.28}$$

The correlations between x_{t+1} and the excess returns conditional information z_t are around 0.1, implying that the R^2 bounds will be significantly larger than that in the

[9] I use $\mu_{c,t}$ rather than x_t to denote the persistent component of consumption since x_t has been used as the state variables of the SDF.

habit formation model. However, both bounds are still less than the predictive R^2s. These results are consistent with Constantinides and Ghosh (2011) and Bansal, Kiku, and Yaron (2012) who find that the permitted degree of predictability is extremely low in the long-run risk framework.

Rare disaster

The rare disaster model revived by Barro (2006) is intended to solve the equity risk premium puzzle and does not accommodate time-varying expected returns. Gabaix (2012) allows for time-varying probability and size of disasters, thereby generating volatility of price-dividend ratios and implying return predictability to some extent. Gourio (2008) exclusively studies whether the predictability generated by the rare disaster model can match the magnitude of predictability observed in market data. In so doing, he introduces an exogenous, persistent, time-varying disaster probability in the rare disaster framework. With numerical simulation, to best match the predictive power of the dividend-price ratio, the model needs to have an average equity premium as high as 13.71%, which is obviously not reasonable. As a result, Gourio concludes "with Epstein-Zin utility, the model can fit the facts qualitatively, and to some extent quantitatively, if we allow for a highly variable probability of disaster, leverage and an IES above unity" to explain return predictability.

The basic assumption for the rare disaster model is that the consumption growth rate follows the stochastic process:

$$\triangle c_{t+1} = \begin{cases} \mu_c + \sigma\epsilon_{t+1}, & \text{with probability } 1 - p_t; \\ \mu_c + \sigma\epsilon_{t+1} + \log(1-b), & \text{with probability } p_t. \end{cases} \tag{1.29}$$

where ϵ_{t+1} is i.i.d. $N(0,1)$ and $0 < b < 1$ is the size of the disaster. The crucial question is to find a variable to proxy the unobservable probability of disasters. Wachter (2012) considers the rare disaster model in a continuous-time setting and find that the dividend-price ratio is a strictly increasing function of the disaster probability, which implies that one can invert this function to find the disaster probability given the observations of the dividend-price ratio. Hence, in addition to the consumption growth rate, the dividend-price ratio can be used as an observable state variable for the rare disaster model. That is,

$$x_{t+1} = (\triangle c_{t+1}, dp_{t+1})'.$$

The predictive R^2 bounds are exhibited in Panel C of Table 1.3. Again, ten predictive R^2s exceed the bounds significantly. These results are approximately the same as that in the habit formation model. Therefore, consistent with Gourio's (2008) numerical simulation, it is difficult for the rare disaster model to match the observed return predictability.

1.4.5 R^2 bounds with market frictions

One interesting question is what happens when the market is not frictionless. The proposed bounds in this paper assume that investors can trade freely without transaction costs and constraints. It may be the case that the profits documented in the literature are not attainable for investors because of transaction costs and constraints. The limit of arbitrage forces investors to deviate from the trading strategy that seeks to exploit predictability in the market.

Market frictions refer to trading costs that can be the transaction cost in He and Modest (1995), the marginal value of liquidity services of tradeable assets in Holmström and Tirole (2001), the transaction cost in Acharya and Pedersen (2005), the funding liquidity in Brunnermeier and Pedersen (2009), or the execution cost in Hasbrouck (2009).[10] Nagel (2012) reviews these models and finds that the SDF

[10] Amihud, Mendelson, and Pedersen (2005) give an excellent literature review on the relationship between transaction costs of different dimensions and asset prices.

in frictionless market can be augmented with a factor Λ_t that captures the state of transaction costs

$$\tilde{m}_{t+1}^F = \tilde{m}_{t+1} \frac{\Lambda_t}{\Lambda_{t+1}}. \tag{1.30}$$

Let $\triangle\lambda_{t+1} = \log(\Lambda_{t+1}/\Lambda_t)$. Then I can rewrite $\tilde{m}_{t+1}^F$ as

$$\tilde{m}_{t+1}^F = \tilde{m}^F(x_{t+1}, \triangle\lambda_{t+1}). \tag{1.31}$$

In this way, a higher $\triangle\lambda_{t+1}$ means a higher transaction cost, and an asset paying well in the state of higher $\triangle\lambda_{t+1}$ earns a lower expected return. The bounds in this paper can be adjusted easily by including $\triangle\lambda_{t+1}$ into the state variables.

I use the liquidity factor constructed by Pástor and Stambaugh (2003) as the proxy of transaction cost. Table 1.4 reports the R^2 bounds on the market portfolio forecasts with the ten macroeconomic variables. Panel A considers the case when the state variable is the consumption growth rate. In this case, the bound with either the maximum risk aversion or the market Sharpe ratio is marginally improved relative to that without considering transaction (Panel A of Table 1.1). All the ten R^2s are significantly larger than the two bounds. Where the Fama-French three factors are used as the state variables in Panel B, the results are almost the same as Panel B of Table 1.2. Six R^2s exceed the two proposed bounds significantly. The results in Table 1.4 are in contrast to de Roon and Szymanowska (2012) who point out that the finding in Kirby (1998) can be reconciled with transaction costs. The reason is that they consider fixed transaction cost while I focus on time-varying cost.

1.4.6 R^2 bounds on cross-sectional portfolio predictability

One interesting question is whether the proposed bounds work well for cross-sectional portfolio forecasts. Theoretically, Propositions 2 and 3 show that individual portfolios should have different predictability since they have different correlations with the state variables. For this reason, I report results on ten size portfolios, ten value portfolios (formed based on the book-to-market ratio), ten momentum portfolios, and ten industry portfolios. I consider two cases for the state variables: the consumption growth rate and the Fama-French three factors. The maximum risk aversion is assumed to be 5. To save space, I report the results when the predictor is the dividend-price ratio or the term spread. The results for the other predictors exhibit similar characteristics and are available upon request.

Portfolio forecasts with dividend-price ratio

Size portfolios The predictability of size portfolios (i.e., portfolios formed based on market capitalization) has been extensively investigated (Ferson and Harvey, 1991; Ferson and Korajczyk, 1995; Kirby, 1998). The basic pattern is that portfolios with small size are more predictable than portfolios with large size. Table 2.6 reports the predictive R^2s when the dividend-price ratio is used as the predictor, and the upper bounds proposed in this paper. Panel A considers the case when the state variable is the consumption growth rate. Surprisingly, the predictability of size portfolios in Column 2 does not show the monotonic pattern reported by Kirby (1998).[11] The minimum predictability is the smallest size portfolio with an R^2 of 0.09%. The maximum predictability is the 4th smallest size portfolio with an R^2 of 0.48%. The predictive R^2 for the largest size portfolio is 0.25%. The bound developed by Ross (2005) is 4.78%, larger than any predictive R^2, suggesting that the predictability of size portfolios can be explained. However, the bound in Zhou (2010) is 0.13%, smaller than all R^2s except for the smallest size portfolio. With respect to the proposed bounds

[11]Kirby (1998) forecasts the size portfolios by using five predictors simultaneously (the excess return on the equally weighted NYSE index, a dummy variable for the month of January, the 1-month 90-day Treasury bill rate less than the 30-day Treasury bill rate, the yield on Moody's Baa rated bonds less the yield on Moody's Aaa rated bonds, and the dividend yield on the S&P 500 stock index less the 30-day Treasury bill rate).

in this paper, both bounds with the maximum risk aversion and the market Sharpe ratio are close to zero and significantly smaller than the corresponding R^2s.

Panel B of Table 2.6 considers the case when the Fama-French three factors are used as the state variables. Since the Fama-French three factor model includes the market portfolio as a factor, which has high correlations with component portfolios, the bound in Zhou (2010) loses the power to diagnose the predictability. However, my bounds remain valid. Except for the first two smallest size portfolios, the other eight portfolios display R^2s larger than the bounds, either the bound with the maximum risk aversion or the bound with the Sharpe ratio. Cross-sectionally, the proposed bounds are monotonically decreasing in firm size. The inability of the dividend-price ratio to generate monotonic predictive R^2s may be due to the fact that the dividend-price ratio uses the sum of dividends paid on the S&P 500 index. Big firms usually pay more dividends than small firms. As a result, the dividend-price ratio is more informative for large size portfolios, exhibiting higher predictive power.

Value portfolios The predictability of the value premium reported in the literature is mixed. Lettau and Ludvigson (2001) show some positive evidence, but Lewellen and Nagel (2006) find that the time-variation in the expected value premium is marginal and hence unpredictable. Table 1.6 reports the results when the dividend-price ratio is used to forecast the ten value portfolios formed based on the book-to-market ratio. The predictive R^2s are 0.11% for the 1st decile portfolio (growth portfolio) and 0.32% for the 10th decile portfolio (value portfolio). This suggests that when the difference between the value and the growth portfolio is used as a proxy of the value premium, the value premium should be significantly predicted by the dividend-price ratio. Panel A shows that the proposed bounds, as well as those in Zhou (2010), are less than the predictive R^2s when the state variable is the consumption growth rate. When the Fama-French three factors are used, the two proposed bounds are still less than the predictive R^2s. While the predictive R^2s are more than 0.25% except for the growth portfolio. The difference between the R^2s for the value portfolios versus the size portfolios is that the proposed bounds do not show a monotonic pattern with respect to the book-to-market ratio.

Momentum portfolios When the dividend-price ratio is used to forecast the ten decile momentum portfolios, the predictive R^2s vary significantly, ranging from 0.10% for the 5th portfolio to 0.83% for the 9th portfolio, as shown in Table 1.7. The R^2s for the lowest- and the highest-momentum portfolios are 0.19% and 0.24%, respectively. Again, all R^2s exceed the proposed bounds when the state variable is the consumption growth rate. When the Fama-French three factors are used, the predictive R^2s, except for the 5th portfolio, exceed the two bounds.

Industry portfolios Ferson and Harvey (1991) and Ferson and Korajczky (1995) show significant predictability for industry portfolios. In Table 2.8, when the dividend-price ratio is used as the predictor, eight out of ten industries (with two exceptions, manufacturing and energy) show strong performance. The most predictable industry is nondurable goods with an R^2 of 0.64%. Panel A shows that all the predictive R^2s except Enrgy exceed the proposed bounds when the state variable is the consumption growth rate. Panel B identifies that five industries that have larger R^2s than the bounds when the state variables are the Fama-French three factors. This result indicates that asset pricing models can generate more predictability for some industry portfolios than others.

Portfolio forecasts with term spread

This section discusses the results when the term spread is used to forecast cross-sectional portfolios. While the term spread exhibits stronger predictive ability, the overall pattern is similar to the case when the predictor is the dividend-price ratio.

When the consumption growth rate is used as the state variable, asset pricing models do not show any hope of explaining return predictability. Instead, when the Fama-French three factors are used, they generate larger bounds. Here I only report the results for the case of the Fama-French three factors (see Tables 1.9 and 1.10), which are summarized as follows. First, the predictive ability of the term spread is too strong to be explained by asset pricing models. That is, all the predictive R^2s, with three exceptions, exceed the proposed bounds. Second, the predictability varies significantly across different portfolios. Third, the failure of current asset pricing models lies in the poor ability of the state variable in capturing the cross-sectional characteristics of individual portfolios.

The results are robust to the habit formation model, the long-run risk model, and the rare disaster model, and also robust to the case with transaction costs. Overall, the cross-sectional results echo the market portfolio forecast that time-varying expected return can only explain a small fraction of predictability.

1.5 Conclusion

This paper asks whether the overall pattern of return predictability is consistent with asset pricing models. To answer this question, I develop two upper bounds on the predictive R^2. When one of ten established macroeconomic variables in Goyal and Welch (2008) is used to forecast the excess returns of the market portfolio and cross-sectional portfolios, the predictive R^2s almost always exceed the upper bounds, implying that return predictability cannot be fully explained by extant asset pricing models. The reason is the low correlation between the forecasted excess return and the state variables used in the SDF.

There are also many other reasons to explain why the predictive R^2s violate the upper bounds. There may be structural breaks in the specific models over the long-term period investigated in this study. For example, Goyal and Welch (2008), Rapach, Strauss, and Zhou (2010), Henkel, and Martin and Nardari (2011) find strong evidence of fairly frequent breaks in the predictive regression. Most macro fundamental variables exhibit significant power of return predictability during economic recessions, but perform badly during economic expansions. It may be necessary to incorporate regime changes into the upper bounds. Also, an alternative explanation is behavioral bias that leads investors to under- or over-react to private or public news, generating return predictability.

This paper focuses on the stock market. It will be of interest to investigate whether any asset pricing model can explain return predictability on the bond market, housing market, commodity market, currency market, and international markets.

Appendix

Proof of Proposition 1. Since $\mathrm{E}_t(\varepsilon_{t+1}|x_{t+1}) = 0$ in the regression $r_{t+1}z_t = a + bx_{t+1} + \varepsilon_{t+1}$, I have

$$\mathrm{Cov}(\varepsilon_{t+1}, \tilde{m}(x_{t+1}) = \mathrm{E}[\mathrm{E}(\varepsilon_{t+1}|x_{t+1})\tilde{m}(x_{t+1})] = 0.$$

Then

$$\mathrm{Cov}(r_{t+1}z_t, \tilde{m}(x_{t+1})) = \mathrm{Cov}[b'x_{t+1}, \tilde{m}(x_{t+1})] = b'\Sigma_{x\tilde{m}}.$$

The Cauchy-Schwarz inequality generates

$$\mathrm{Cov}[r_{t+1}z_t, \tilde{m}(x_{t+1})]^2 = (b'\Sigma_{xx}^{1/2}\Sigma_{xx}^{-1/2}\Sigma_{x\tilde{m}})^2 \leq (b'\Sigma_{xx}b)(\Sigma'_{x\tilde{m}}\Sigma_{xx}^{-1}\Sigma_{x\tilde{m}}).$$

From the regression $r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}$, the R^2 is

$$\begin{aligned}
R^2 &= \frac{\beta \mathrm{Var}(z_t)\beta}{\mathrm{Var}(r_{t+1})} \leq \frac{\mathrm{Cov}^2(r_{t+1}, z_t)}{\mathrm{Var}(r_{t+1})} \\
&= \frac{\mathrm{Cov}^2(\tilde{m}_{t+1}, r_{t+1}z_t)}{\mathrm{Var}(r_{t+1})} \\
&\leq \frac{b'\Sigma_{xx}b}{\mathrm{Var}(r_{t+1}z_t)} \frac{\mathrm{Var}(r_{t+1}z_t)(\Sigma'_{x\tilde{m}}\Sigma_{xx}^{-1}\Sigma_{x\tilde{m}})}{\mathrm{Var}(r_{t+1})} \\
&\leq \rho_{x,rz}^2 \frac{\mathrm{Var}(r_{t+1}z_t)}{\mathrm{Var}(r_{t+1})} \mathrm{Var}(\tilde{m}_{t+1}) \\
&= \phi_{x,rz}^2 \mathrm{Var}(\tilde{m}_{t+1}).
\end{aligned}$$

This completes the proof.

1.6 References

Acharya, V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 374–410.

Aït-Sahalia, Y., Lo. A.W., 2000. Nonparametric risk management and implied risk aversion. Journal of Econometrics 94, 9–51.

Albuquerque, R., Eichenbaum, M., Rebelo, S., 2012. Valuation risk and asset pricing. Working paper, Boston University.

Amihud, Y., Mendelson, H., Pedersen, L.H., 2005. Liquidity and asset prices. Foundations and Trends in Finance 1, 269–364.

Ang, A., Bekaert, G., 2007. Return predictability: is it there? Review of Financial Studies 17, 651–707.

Baker, M., Wurgler, J., 2000. The equity share in new issues and aggregate stock returns. Journal of Finance 55, 2219–2257.

Balduzzi, P., Kallal, H., 1997. Risk premia and variance bounds. Journal of Finance 52, 1913–1997.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

Bansal, R., Kiku, D., Yaron, A., 2012. An empirical evaluation of the long-run risks model for asset prices. Critical Finance Review 1, 183–221.

Bansal, R., Viswanathan, S., 1993. No arbitrage and arbitrage pricing: A new approach. Journal of Finance 48, 1231–1278.

Barro, R.J., 2006. Rare disasters and asset markets in the twentieth centrury. Quarterly Journal of Economics 121, 823–866.

Barro, R.J., Ursúa, J.F., 2012. Rare macroeconomic disasters. Annual Review of Economics, forthcoming.

Bernardo, A.E., Ledoit, O., 2000. Gain, loss, and asset pricing. Journal of Political Economy 108, 144–172.

Breen, W., Glosten, L., Jagannathan, R., 1989. Economic significance of predictable variations in stock index returns. Journal of Finance 64, 1177–1189.

Brunnermeier, M.K., Pedersen, L.H., 2009. Market liquidity and funding liquidity. Review of Financial Studies 22, 2201–2238.

Campbell, J.Y., 1987. Stock returns and term structure. Journal of Financial Economics 18, 373–399.

Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.

Campbell, J., Yogo, M., 2006. Efficient tests of stock return predictability. Journal of Financial Economics 81, 27–60.

Campbell, J.Y., Shiller, R., 1988. Stock prices, earnings, and expected dividends. Journal of Finance 43, 661–676.

Campbell, J., Vuolteenaho, T., 2004. Inflation illusion and stock prices. American Economic Review 94, 19–23.

Cochrane, J.H., 2005. Asset Pricing. Princeton, NJ.

Cochrane, J.H., 2011. Presidential address: discount rates. Journal of Finance 66, 1047–1108.

Cochrane, J.H., Hansen, L.P., 1992. Asset pricing explorations for macroeconomics. NBER Macroeconomics Annual 1992 7, 115-182. MIT Press.

Cochrane, J.H., Saá-Requejo, J., 2000. Beyond arbitrage: good-deal asset price bounds in incomplete markets. Journal of Political Economy 108, 79–119.

Constantinides, G.M., Ghosh, A., 2011. Asset pricing tests with long-run risks in consumption growth. Revew of Asset Pricing Studies 1, 96–136.

Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. Review of Financial Studies 22, 2801–2833.

Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.

de Roon, F., Szymanowska, M., 2012. Asset pricing restrictions on predictability: frictions matter. Management Science 58, 1916–1932.

Dittmar, R.F., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. Journal of Finance 57, 369–403.

Driesprong, G., Jacobsen, B., Maat, B., 2008. Striking oil: another puzzle? Journal of Financial Economics 89, 307–327.

Fama, E., French, K., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.

Fama, E., French, K., 1993. Common risk factors in the returns on stock and bonds. Journal of Financial Economics 33, 3–56.

Fama, E., Schwert, G.W., 1977. Asset returns and inflation. Journal of Financial Economics 5, 115–146.

Ferson, W., Harvey, C., 1991. The variation in economic risk premiums. Journal of Political Economy 99, 385–415.

Ferson, W., Korajczyk, R.A., 1995. Do arbitrage pricing models explain the predictability of stock returns? Journal of Business 68, 309–349.

Ferson, W., Siegel, A., 2003. Stochastic discount factor bounds with conditioning information. Review of Financial Studies 16, 567–595.

French, K.R, Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 3, 3–29.

Furman, E., Zitikis, R., 2008. Weighted risk capital allocations. Insurance: Mathematics and Economics 43, 263–269.

Gabaix, X., 2012. Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. Quarterly Journal of Economics 127, 645–700.

Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium performance. Review of Financial Studies 21, 1455–1508.

Gourio, F., 2008. Time-series predictability in the disaster model. Finance Reserach Letters 5, 191–203.

Gourio, F., 2012. Disaster risk and business cycles. American Economic Review, forthcoming.

Guiso, L., Sapienza P., Zingales, L., 2011. Time varying risk aversion, Working paper.

Guo, H., 2006. On the out-of-sample predictability of stock market returns. Journal of Business 79, 645–670.

Han, Y., Yang, K., Zhou, G., 2012. A new anomaly: the cross-sectional profitability of technial analysis. Journal of Financial and Quantitative Analysis, forthcoming.

Hansen, L., Singleton, K., 1983. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. Journal of Political Economy 91, 249–264.

Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of Political Economy 99, 225–262.

Hasbrouck, J., 2009. Trading costs and returns for US equities: Estimating effective costs from daily data. Journal of Finance 64, 1445–1477.

He, H., Modest, D.M., 1995. Market frictions and consumption-based asset pricing. Journal of Political Economy 103, 94–117.

Henkel, S.J., Martin, J.S., Nadari, F., 2011. Time-varying short-horizon predictability. Journal of Financial Economics 99, 560–580.

Holmström, B., Tirole, J., 2001. LAPM: A liquidity-based asset pricing model. Journal of Finance 56, 1837–1867.

Hong, H., Yogo, M., 2012. What does futures maket interest tell us about the macroeconomy and asset prices? Journal of Financial Economics 105, 473–490.

Kan, R., Zhou, G., 2006. A new variance bound on the stochastic discount factor. Journal of Business 79, 941–961.

Kirby, C., 1998. The restrictions on predictability implied by rational asset pricing models. Review of Financial Studies 11, 343–382.

Kothari, S.P., Shanken, J., 1997. Book-to-market, dividend yield, and expected market returns: a time series analysis. Journal of Financial Economics 44, 169–203.

Ledoit, O., 1995. Essays on Risk and Return in the Stock Market. Ph.D. dissertation, MIT.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. Journal of Finance 56, 815–849.

Lewellen J., Nagel, S., 2006. The conditional CAPM does not explain asset-pricing anomalies. Journal of Financial Economics 82, 289–314.

Ludvigson, S.C., Ng, S., 2007. The empirical risk-return relation: A factor analysis approach. Journal of Financial Economics 83, 171–222.

Mehra, R., Prescott, E., 1985. The equity premium puzzle. Journal of Monetary Economics 15, 145–161.

Mele, A., 2007. Asymmetric stock market volatility and the cyclical behavior of expected returns. Journal of Financial Economics 86, 446–478.

Muirhead, R.J., 1982. Aspects of multivariate statistical theory. New York: Wiley.

Nagel, S., Singleton, K.J., 2011. Estimation and evaluation of conditional asset pricing models. Journal of Finance 66, 873–909.

Paravisini, D., Rappoport, V., Ravina, E., 2012. Risk aversion and wealth: evidence from person-to-person lending portfolios. Working Paper, Columbia GSB.

Pástor, Ľ., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Rangvid, J., 2006. Output and expected returns. Journal of Financial Economics 81, 595–624.

Rapach, D.E., Strauss, J.K., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Review of Financial Studies 23, 821–862.

Ross, S.A., 1976. The arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–360.

Ross, S.A., 2005. Neoclassical Finance. Princeton University Press.

Shanken, J., 1992. The current state of the arbitrage pricing theory. Journal of Finance 47, 1569–1574.

Wachter, J.A., 2012. Can time-varying risk of rare disasters explain aggregate stock market volatility? Journal of Finance, forthcoming.

Zhou, G., 2010. How much stock return predictability can we expect from an asset pricing model? Economics Letters 108, 184–186.

Table 1.1: R^2 bounds on market portfolio forecast

The table reports the bounds of the predictive R^2 from the regression $r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}$, where r_{t+1} is the excess return of the market portfolio and z_t is one of the ten predictors given in the first column. The state variable x in the SDF is the consumption growth rate. The marginal investor's risk aversion is 5 in Panel A and 10 in Panel B. $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

z	R^2(%)	$\bar{R}^2_{Ross}$(%)	ρ_{x,m_0}	$\bar{R}^2_{Zhou}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
Panel A: maximum risk aversion $\gamma = 5$							
d/p	0.27	4.78	0.17	0.13***	0.02	0.00***	0.00***
e/p	0.11	4.78	0.17	0.13	0.03	0.01***	0.01***
dy	0.33	4.78	0.17	0.13***	0.02	0.00***	0.00***
tbl	0.21	4.78	0.17	0.13***	0.06	0.02***	0.02***
dfy	0.21	4.78	0.17	0.13***	0.02	0.00***	0.00***
tms	0.56	4.78	0.17	0.13***	0.04	0.01***	0.01***
ntis	0.04	4.78	0.17	0.13	0.02	0.00***	0.00***
infl	0.42	4.78	0.17	0.13***	0.02	0.00***	0.00***
ltr	1.04	4.78	0.17	0.13***	0.02	0.00***	0.00***
rvol	1.23	4.78	0.17	0.13***	0.03	0.00***	0.00***
Panel B: maximum risk aversion $\gamma = 10$							
d/p	0.27	19.11	0.17	0.54	0.02	0.01***	0.00***
e/p	0.11	19.11	0.17	0.54	0.03	0.03***	0.01***
dy	0.33	19.11	0.17	0.54	0.02	0.01***	0.00***
tbl	0.21	19.11	0.17	0.54	0.06	0.07***	0.02***
dfy	0.21	19.11	0.17	0.54	0.02	0.01***	0.00***
tms	0.56	19.11	0.17	0.54**	0.04	0.03***	0.01***
ntis	0.04	19.11	0.17	0.54	0.02	0.01**	0.00***
infl	0.42	19.11	0.17	0.54	0.02	0.01***	0.00***
ltr	1.04	19.11	0.17	0.54***	0.02	0.01***	0.00***
rvol	1.23	19.11	0.17	0.54***	0.03	0.01***	0.00***

Table 1.2: R^2 bounds on market portfolio forecast

The table reports the bounds of the predictive R^2 from the regression $r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}$, where r_{t+1} is the excess return of the market portfolio and z_t is one of the ten predictors given in the first column. The state variables x_t in the SDF are the market portfolio in Panel A or the Fama-French three factors in Panel B. $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

z	$R^2(\%)$	$\bar{R}^2_{Ross}(\%)$	ρ_{x,m_0}	$\bar{R}^2_{Zhou}(\%)$	$\rho_{x,rz}$	$\bar{R}^2_{RA}(\%)$	$\bar{R}^2_{SR}(\%)$
Panel A: x_t is the return of the market portfolio							
d/p	0.27	4.78	1.00	4.76	0.10	0.06***	0.05***
e/p	0.11	4.78	1.00	4.76	0.05	0.02***	0.02***
dy	0.33	4.78	1.00	4.76	0.06	0.02***	0.02***
tbl	0.21	4.78	1.00	4.76	0.07	0.03***	0.02***
dfy	0.21	4.78	1.00	4.76	0.25	0.47	0.40
tms	0.56	4.78	1.00	4.76	0.03	0.00***	0.00***
ntis	0.04	4.78	1.00	4.76	0.19	0.24	0.21
infl	0.42	4.78	1.00	4.76	0.10	0.07***	0.06***
ltr	1.04	4.78	1.00	4.76	0.00	0.00***	0.00***
rvol	1.23	4.78	1.00	4.76	0.25	0.32***	0.27***
Panel B: x_t are the Fama-French three-factors							
d/p	0.27	4.78	1.00	4.76	0.11	0.07***	0.06***
e/p	0.11	4.78	1.00	4.76	0.12	0.10	0.09**
dy	0.33	4.78	1.00	4.76	0.08	0.04***	0.03***
tbl	0.21	4.78	1.00	4.76	0.25	0.35	0.30
dfy	0.21	4.78	1.00	4.76	0.32	0.75	0.64
tms	0.56	4.78	1.00	4.76	0.18	0.16***	0.13***
ntis	0.04	4.78	1.00	4.76	0.23	0.33	0.28
infl	0.42	4.78	1.00	4.76	0.16	0.16***	0.14***
ltr	1.04	4.78	1.00	4.76	0.09	0.06***	0.04***
rvol	1.23	4.78	1.00	4.76	0.28	0.40***	0.34***

Table 1.3: R^2 bounds on market portfolio forecast with recently developed models

The table reports the bounds on the R^2 from the regression $r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}$, where r_{t+1} is the excess return of the market portfolio and z_t is one of the ten predictors in Column 1. $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

z	$R^2(\%)$	$\bar{R}^2_{Ross}(\%)$	ρ_{x,m_0}	$\bar{R}^2_{Zhou}(\%)$	$\rho_{x,rz}$	$\bar{R}^2_{RA}(\%)$	$\bar{R}^2_{SR}(\%)$
Panel A: Habit formation model							
d/p	0.27	4.78	0.18	0.15***	0.07	0.03***	0.02***
e/p	0.11	4.78	0.18	0.15	0.04	0.01***	0.01***
dy	0.33	4.78	0.18	0.15***	0.06	0.02***	0.02***
tbl	0.21	4.78	0.18	0.15***	0.06	0.02***	0.02***
dfy	0.21	4.78	0.18	0.15***	0.11	0.10***	0.08***
tms	0.56	4.78	0.18	0.15***	0.04	0.01***	0.01***
ntis	0.04	4.78	0.18	0.15	0.02	0.00***	0.00***
infl	0.42	4.78	0.18	0.15***	0.03	0.00***	0.00***
ltr	1.04	4.78	0.18	0.15***	0.04	0.01***	0.01***
rvol	1.23	4.78	0.18	0.15***	0.03	0.00***	0.00***
Panel B: Long-run risk model							
d/p	0.27	4.78	0.17	0.14***	0.11	0.07***	0.06***
e/p	0.11	4.78	0.17	0.14	0.07	0.04***	0.03***
dy	0.33	4.78	0.17	0.14***	0.12	0.09***	0.07***
tbl	0.21	4.78	0.17	0.14***	0.08	0.03***	0.03***
dfy	0.21	4.78	0.17	0.14***	0.09	0.06***	0.05***
tms	0.56	4.78	0.17	0.14***	0.09	0.04***	0.03***
ntis	0.04	4.78	0.17	0.14	0.11	0.08	0.07
infl	0.42	4.78	0.17	0.14***	0.13	0.10***	0.08***
ltr	1.04	4.78	0.17	0.14***	0.10	0.06***	0.05***
rvol	1.23	4.78	0.17	0.14***	0.12	0.07***	0.06***
Panel C: Rare disaster model							
d/p	0.27	4.78	0.17	0.14***	0.08	0.04***	0.03***
e/p	0.11	4.78	0.17	0.14	0.06	0.02***	0.02***
dy	0.33	4.78	0.17	0.14***	0.08	0.04***	0.03***
tbl	0.21	4.78	0.17	0.14***	0.05	0.02***	0.01***
dfy	0.21	4.78	0.17	0.14***	0.08	0.05***	0.04***
tms	0.56	4.78	0.17	0.14***	0.07	0.03***	0.02***
ntis	0.04	4.78	0.17	0.14	0.06	0.02***	0.02***
infl	0.42	4.78	0.17	0.14***	0.03	0.01***	0.00***
ltr	1.04	4.78	0.17	0.14***	0.09	0.05***	0.04***
rvol	1.23	4.78	0.17	0.14***	0.04	0.01***	0.01***

Table 1.4: R^2 bounds on market portfolio forecast with transaction cost

The table reports the bounds on the R^2 from the regression $r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}$, where r_{t+1} is the excess return of the market portfolio and z_t is one of the 10 predictors in Column 1. The state variables x_t in the SDF are the consumption growth rate in Panel A and the Fama-French three factors in Panel B. The transaction cost is measured by the liquidity factor of Pástor and Stambaugh (2003). $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

z	$R^2(\%)$	$\bar{R}^2_{Ross}(\%)$	ρ_{x,m_0}	$\bar{R}^2_{Zhou}(\%)$	$\rho_{x,rz}$	$\bar{R}^2_{RA}(\%)$	$\bar{R}^2_{SR}(\%)$
Panel A: x_t is consumption growth rate							
d/p	0.29	4.88	0.34	0.55	0.05	0.01***	0.01***
e/p	0.10	4.88	0.34	0.55	0.03	0.01***	0.01***
dy	0.31	4.88	0.34	0.55	0.06	0.02***	0.02***
tbl	0.24	4.88	0.34	0.55	0.04	0.01***	0.01***
dfy	0.20	4.88	0.34	0.55	0.01	0.00***	0.00***
tms	0.53	4.88	0.34	0.55	0.04	0.01***	0.01***
ntis	0.04	4.88	0.34	0.55	0.06	0.02**	0.02**
infl	0.55	4.88	0.34	0.54	0.09	0.05***	0.04***
ltr	1.10	4.88	0.34	0.55***	0.09	0.05***	0.04***
rvol	1.19	4.88	0.34	0.55***	0.03	0.00***	0.00***
Panel B: x_t are the Fama-French three-factors							
d/p	0.29	4.88	1.00	4.86	0.14	0.12***	0.11***
e/p	0.10	4.88	1.00	4.86	0.13	0.12	0.10
dy	0.31	4.88	1.00	4.86	0.12	0.08***	0.07***
tbl	0.24	4.88	1.00	4.86	0.26	0.39	0.33
dfy	0.20	4.88	1.00	4.86	0.34	0.84	0.72
tms	0.53	4.88	1.00	4.86	0.18	0.16***	0.14***
ntis	0.04	4.88	1.00	4.86	0.22	0.32	0.28
infl	0.55	4.88	1.00	4.86	0.17	0.19***	0.16***
ltr	1.10	4.88	1.00	4.86	0.12	0.09***	0.08***
rvol	1.19	4.88	1.00	4.86	0.29	0.42***	0.36***

Table 1.5: R^2 bounds on size portfolio forecasts

The table reports the bounds of the predictive R^2 from the regression $r_{j,t+1} = \alpha_j + \beta_j z_t + \varepsilon_{j,t+1}$, where $r_{j,t+1}$ is the excess return on one of the ten size portfolios and z_t is the dividend-price ratio. The state variables x_t in the SDF are the consumption growth rate (Panel A) or the Fama-French three factors (Panel B). $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

r	R^2(%)	$\bar{R}^2_{Ross}$(%)	ρ_{x,m_0}	$\bar{R}^2_{Zhou}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
Panel A: x_t is the consumption growth rate							
Small	0.09	4.78	0.17	0.13	0.08	0.04***	0.03***
2	0.20	4.78	0.17	0.13***	0.05	0.02***	0.02***
3	0.32	4.78	0.17	0.13***	0.06	0.02***	0.02***
4	0.48	4.78	0.17	0.13***	0.04	0.01***	0.01***
5	0.43	4.78	0.17	0.13***	0.03	0.01***	0.01***
6	0.43	4.78	0.17	0.13***	0.03	0.01***	0.00***
7	0.29	4.78	0.17	0.13***	0.03	0.01***	0.00***
8	0.28	4.78	0.17	0.13***	0.00	0.00***	0.00***
9	0.24	4.78	0.17	0.13***	0.02	0.00***	0.00***
Large	0.25	4.78	0.17	0.13***	0.01	0.00***	0.00***
Panel B: x_t are the Fama-French three-factors							
Small	0.09	4.78	1.00	4.76	0.23	0.31	0.26
2	0.20	4.78	1.00	4.76	0.23	0.34	0.29
3	0.32	4.78	1.00	4.76	0.20	0.23***	0.19***
4	0.48	4.78	1.00	4.76	0.19	0.18***	0.16***
5	0.43	4.78	1.00	4.76	0.17	0.18***	0.15***
6	0.43	4.78	1.00	4.76	0.14	0.12***	0.10***
7	0.29	4.78	1.00	4.76	0.15	0.14***	0.11***
8	0.28	4.78	1.00	4.76	0.14	0.12***	0.10***
9	0.24	4.78	1.00	4.76	0.12	0.09***	0.08***
Large	0.25	4.78	1.00	4.76	0.09	0.06***	0.05***

Table 1.6: R^2 bounds on value portfolio forecasts

The table reports the bounds of the predictive R^2 from the regression $r_{j,t+1} = \alpha_j + \beta_j z_t + \varepsilon_{j,t+1}$, where $r_{j,t+1}$ is the excess return on one of the ten value portfolios and z_t is the dividend-price ratio. The state variables x_t in the SDF are the consumption growth rate (Panel A) or the Fama-French three factors (Panel B). $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

r	$R^2(\%)$	$\bar{R}^2_{Ross}(\%)$	ρ_{x,m_0}	$\bar{R}^2_{Zhou}(\%)$	$\rho_{x,rz}$	$\bar{R}^2_{RA}(\%)$	$\bar{R}^2_{SR}(\%)$
Panel A: x_t is the consumption growth rate							
Low	0.11	4.78	0.17	0.13	0.04	0.01^{***}	0.01^{***}
2	0.33	4.78	0.17	0.13^{***}	0.03	0.01^{***}	0.00^{***}
3	0.28	4.78	0.17	0.13^{***}	0.03	0.01^{***}	0.00^{***}
4	0.25	4.78	0.17	0.13^{***}	0.03	0.01^{***}	0.01^{***}
5	0.19	4.78	0.17	0.13^{***}	0.03	0.01^{***}	0.01^{***}
6	0.39	4.78	0.17	0.13^{***}	0.02	0.00^{***}	0.00^{***}
7	0.23	4.78	0.17	0.13^{***}	0.04	0.01^{***}	0.01^{***}
8	0.28	4.78	0.17	0.13^{***}	0.02	0.00^{***}	0.00^{***}
9	0.41	4.78	0.17	0.13^{***}	0.04	0.01^{***}	0.01^{***}
High	0.32	4.78	0.17	0.13^{***}	0.06	0.02^{***}	0.02^{***}
Panel B: x_t are the Fama-French three-factors							
Low	0.11	4.78	1.00	4.76	0.11	0.09^{**}	0.08^{***}
2	0.33	4.78	1.00	4.76	0.13	0.09^{***}	0.08^{***}
3	0.28	4.78	1.00	4.76	0.15	0.14^{***}	0.12^{***}
4	0.25	4.78	1.00	4.76	0.16	0.15^{***}	0.13^{***}
5	0.19	4.78	1.00	4.76	0.15	0.14^{***}	0.12^{***}
6	0.39	4.78	1.00	4.76	0.12	0.08^{***}	0.07^{***}
7	0.23	4.78	1.00	4.76	0.17	0.16^{***}	0.14^{***}
8	0.28	4.78	1.00	4.76	0.18	0.18^{***}	0.15^{***}
9	0.41	4.78	1.00	4.76	0.15	0.12^{***}	0.10^{***}
High	0.32	4.78	1.00	4.76	0.18	0.16^{***}	0.14^{***}

Table 1.7: R^2 bounds on momentum portfolio forecasts

The table reports the bounds of the predictive R^2 from the regression $r_{j,t+1} = \alpha_j + \beta_j z_t + \varepsilon_{j,t+1}$, where $r_{j,t+1}$ is the excess return on one of the ten momentum portfolios and z_t is the dividend-price ratio. The state variables x_t in the SDF are the consumption growth rate (Panel A) or the Fama-French three factors (Panel B). $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

r	R^2(%)	$\bar{R}^2_{Ross}$(%)	ρ_{x,m_0}	$\bar{R}^2_{Zhou}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
Panel A: x_t is the consumption growth rate							
Low	0.19	4.78	0.17	0.13***	0.01	0.00***	0.00***
2	0.23	4.78	0.17	0.13***	0.04	0.01***	0.01***
3	0.48	4.78	0.17	0.13***	0.04	0.01***	0.01***
4	0.24	4.78	0.17	0.13***	0.06	0.02***	0.02***
5	0.10	4.78	0.17	0.13	0.08	0.03***	0.01***
6	0.56	4.78	0.17	0.13***	0.07	0.03***	0.02***
7	0.31	4.78	0.17	0.13***	0.07	0.03***	0.03***
8	0.31	4.78	0.17	0.13***	0.03	0.01***	0.01***
9	0.83	4.78	0.17	0.13***	0.02	0.00***	0.00***
High	0.24	4.78	0.17	0.13***	0.01	0.00***	0.00***
Panel B: x_t are the Fama-French three-factors							
Low	0.19	4.78	1.00	4.76	0.12	0.08***	0.07***
2	0.23	4.78	1.00	4.76	0.10	0.06***	0.05***
3	0.48	4.78	1.00	4.76	0.12	0.13***	0.11***
4	0.24	4.78	1.00	4.76	0.18	0.18***	0.15***
5	0.10	4.78	1.00	4.76	0.17	0.17	0.15
6	0.56	4.78	1.00	4.76	0.19	0.22***	0.19***
7	0.31	4.78	1.00	4.76	0.21	0.27**	0.23***
8	0.31	4.78	1.00	4.76	0.19	0.24***	0.20***
9	0.83	4.78	1.00	4.76	0.16	0.17***	0.14***
High	0.24	4.78	1.00	4.76	0.15	0.16***	0.13***

Table 1.8: R^2 bounds on industry portfolio forecasts

The table reports the bounds of the predictive R^2 from the regression $r_{j,t+1} = \alpha_j + \beta_j z_t + \varepsilon_{j,t+1}$, where $r_{j,t+1}$ is the excess return on one of the ten industry portfolios and z_t is the dividend-price ratio. The state variables x_t in the SDF are the consumption growth rate (Panel A) or the Fama-French three factors (Panel B). $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

r	R^2(%)	$\bar{R}^2_{Ross}$(%)	ρ_{x,m_0}	$\bar{R}^2_{Zhou}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
Panel A: x_t is the consumption growth rate							
NoDur	0.64	4.78	0.17	0.13***	0.09	0.05***	0.04***
Durbl	0.19	4.78	0.17	0.13***	0.06	0.02***	0.02***
Manuf	0.03	4.78	0.17	0.13	0.03	0.01***	0.01***
Enrgy	0.01	4.78	0.17	0.13	0.00	0.00	0.00
HiTec	0.10	4.78	0.17	0.13	0.01	0.00***	0.00***
Telcm	0.52	4.78	0.17	0.13***	0.02	0.00***	0.00***
Shops	0.47	4.78	0.17	0.13***	0.04	0.01***	0.01***
Hlth	0.25	4.78	0.17	0.13***	0.08	0.04***	0.03***
Utils	0.16	4.78	0.17	0.13***	0.00	0.00***	0.00***
Other	0.44	4.78	0.17	0.13***	0.05	0.01***	0.01***
Panel B: x_t are the Fama-French three-factors							
NoDur	0.64	4.78	1.00	4.76	0.27	0.42***	0.35***
Durbl	0.19	4.78	1.00	4.76	0.09	0.04***	0.03***
Manuf	0.03	4.78	1.00	4.76	0.16	0.16	0.14
Enrgy	0.01	4.78	1.00	4.76	0.20	0.29	0.24
HiTec	0.10	4.78	1.00	4.76	0.21	0.32	0.27
Telcm	0.52	4.78	1.00	4.76	0.16	0.16***	0.13***
Shops	0.47	4.78	1.00	4.76	0.18	0.18***	0.15***
Hlth	0.25	4.78	1.00	4.76	0.23	0.32	0.27
Utils	0.16	4.78	1.00	4.76	0.25	0.40	0.34
Other	0.44	4.78	1.00	4.76	0.22	0.28***	0.24***

Table 1.9: R^2 bounds on size and value portfolio forecasts

The table reports the bounds on R^2 from the regression $r_{j,t+1} = \alpha_j + \beta_j z_t + \varepsilon_{j,t+1}$, where z_t is the term spread and $r_{j,t+1}$ is the excess return on one of the 10 size portfolios (Panel A) or the 10 value portfolios (Panel B). The state variables x_t in the SDF are the Fama-French three factors. $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

r	$R^2(\%)$	$\bar{R}^2_{Ross}(\%)$	ρ_{x,m_0}	$\bar{R}^2_{Zhou}(\%)$	$\rho_{x,rz}$	$\bar{R}^2_{RA}(\%)$	$\bar{R}^2_{SR}(\%)$
			Panel A: Size portfolios				
Small	0.83	4.78	1.00	4.76	0.23	0.26***	0.22***
2	0.61	4.78	1.00	4.76	0.20	0.19***	0.16***
3	0.54	4.78	1.00	4.76	0.17	0.15***	0.13***
4	0.63	4.78	1.00	4.76	0.17	0.16***	0.13***
5	0.58	4.78	1.00	4.76	0.17	0.14***	0.12***
6	0.81	4.78	1.00	4.76	0.16	0.13***	0.11***
7	0.69	4.78	1.00	4.76	0.16	0.14***	0.12***
8	0.52	4.78	1.00	4.76	0.16	0.14***	0.12***
9	0.58	4.78	1.00	4.76	0.18	0.18***	0.13***
Large	0.54	4.78	1.00	4.76	0.18	0.17***	0.14***
			Panel B: Value portfolios				
Low	0.52	4.78	1.00	4.76	0.18	0.16***	0.14***
2	0.51	4.78	1.00	4.76	0.13	0.08***	0.07***
3	0.73	4.78	1.00	4.76	0.12	0.08***	0.06***
4	0.68	4.78	1.00	4.76	0.14	0.10***	0.09***
5	0.65	4.78	1.00	4.76	0.12	0.08***	0.07***
6	0.43	4.78	1.00	4.76	0.14	0.10***	0.09***
7	0.28	4.78	1.00	4.76	0.11	0.06***	0.05***
8	0.14	4.78	1.00	4.76	0.16	0.14***	0.11***
9	0.29	4.78	1.00	4.76	0.16	0.14***	0.12***
High	0.49	4.78	1.00	4.76	0.20	0.24***	0.21***

Table 1.10: R^2 bounds on momentum and industry portfolio forecasts

The table reports the bounds on R^2 from the regression $r_{j,t+1} = \alpha_j + \beta_j z_t + \varepsilon_{j,t+1}$, where z_t is the term spread and $r_{j,t+1}$ is the excess return on one of the 10 momentum portfolios (Panel A) or the 10 industry portfolios (Panel B). The state variables x_t in the SDF are the Fama-French three factors. $\bar{R}^2_{Ross}$ and $\bar{R}^2_{Zhou}$ denote the bounds proposed by Ross (2005) and Zhou (2010). $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ denote the bounds developed in this paper with the maximum risk aversion and the market Sharpe ratio, respectively. ρ_{x,m_0} and $\rho_{x,rz}$ denote the multiple correlations, where the default SDF m_0 is constructed by the market portfolio. Statistical significance is assessed by the Wald statistic for testing that the predictive R^2 is less than the theoretical upper bound. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

r	$R^2(\%)$	$\bar{R}^2_{Ross}(\%)$	ρ_{x,m_0}	$\bar{R}^2_{Zhou}(\%)$	$\rho_{x,rz}$	$\bar{R}^2_{RA}(\%)$	$\bar{R}^2_{SR}(\%)$
			Panel A: Momentum portfolios				
Low	0.54	4.78	1.00	4.76	0.22	0.25^{***}	0.21^{***}
2	0.43	4.78	1.00	4.76	0.18	0.16^{***}	0.14^{***}
3	0.51	4.78	1.00	4.76	0.15	0.12^{***}	0.10^{***}
4	0.60	2.78	1.00	4.76	0.15	0.11^{***}	0.10^{***}
5	0.50	4.78	1.00	4.76	0.17	0.14^{***}	0.12^{***}
6	0.80	4.78	1.00	4.76	0.15	0.13^{***}	0.11^{***}
7	1.11	4.78	1.00	4.76	0.14	0.11^{***}	0.09^{***}
8	0.56	4.78	1.00	4.76	0.16	0.14^{***}	0.12^{***}
9	0.37	4.78	1.00	4.76	0.15	0.13^{***}	0.11^{***}
High	0.33	4.78	1.00	4.76	0.19	0.21^{***}	0.18^{***}
			Panel B: Industry portfolios				
NoDur	0.40	4.78	1.00	4.76	0.16	0.14^{***}	0.12^{***}
Durbl	1.41	4.78	1.00	4.76	0.21	0.24^{***}	0.20^{***}
Manuf	0.80	4.78	1.00	4.76	0.15	0.12^{***}	0.10^{***}
Enrgy	0.16	4.78	1.00	4.76	0.12	0.08^{***}	0.07^{***}
HiTec	0.61	4.78	1.00	4.76	0.21	0.21^{***}	0.18^{***}
Telcm	0.21	4.78	1.00	4.76	0.21	0.21	0.18^{*}
Shops	0.43	4.78	1.00	4.76	0.10	0.05^{***}	0.04^{***}
Hlth	0.00	4.78	1.00	4.76	0.12	0.06	0.05
Utils	0.19	4.78	1.00	4.76	0.03	0.00^{***}	0.00^{***}
Other	0.40	4.78	1.00	4.76	0.16	0.14^{***}	0.12^{***}

Chapter 2

Economic and Market Conditions: Two State Variables that Predict the Stock Market[1]

Stock market predictability is of great interest to both researchers and practitioners. In this paper, motivated by both economic theory and investment practice, we identify two new predictors that capture the state of the economy and the market condition, and outperform the well-known variables in predicting the market risk premium. Moreover, the new predictors forecast the stock market not only during the down turns of the economy, but also during the up turns when extant predictors fail. Cross-sectionally, we find that the same idea also provides new and effective predictors for component portfolios sorted by size, book-to-market ratio, industry, and long- and short-term reversals.

2.1 Introduction

Forecasting the stock market is of great interest to both academic researchers and investment practitioners. There is a huge literature on predictability, and numerous economic variables have been identified as predictors of the market risk premium. For example, Rozeff (1984), Fama and French (1988), and Campbell and Shiller (1988a, 1988b) present evidence that various valuation ratios, such as the dividend yield, have forecasting power. On the other hand, Keim and Stambaugh (1986), Campbell (1987), Breen, Glosten, and Jaganathan (1989), and Fama and French (1989) find that

[1] This is a joint work with Guofu Zhou.

the nominal interest rates and the interest rate spreads can predict the market risk premium, while Nelson (1976) and Fama and Schwert (1977) find that the inflation rate can predict the market too. Recent studies continue to confirm the predictability using valuation ratios (Cochrane, 2008; Pastor and Stambaugh, 2009), interest rates (Ang and Bekaert, 2007), and inflation (Campbell and Vuolteenaho, 2004). Other studies identify useful economic variables as new predictors, including the corporate issuing activity (Baker and Wurgler, 2000; Boudoukh, Michaely, Richardson, and Roberts, 2007), the consumption-wealth ratio (Lettau and Ludvigson, 2001), and stock volatility (Guo, 2006). While these studies provide mostly in-sample predictability, Rapach, Strauss, and Zhou (2010), Henkel, Martin, and Nardari (2011), Ferreira and Santa-Clara (2011), Dangl and Halling (2012), and Neely, Rapach, Tu, and Zhou (2012), among others, show that the market risk premium can be reliably predicted out-of-sample as well. However, despite the diversity of existing economic predictors, none of them seems to capture the overall outlook of the economy and the mean-reversion behavior of the stock market.

In this paper, we identify two new predictors of the stock market. The first predictor is an economic indicator of the future state of the overall economy. The classic Merton (1971) model provides the theoretical basis for the state of the economy or the changing investment opportunity set as the source for the time-varying market risk premium. Surprisingly, in the vast empirical literature on predictability, no studies have ever used a simple aggregate measure of the economy to predict the market, though various individual economic variables have been used. Since each individual economic variable summarizes only one aspect of the economy, it does not capture the overall state of the economy. To utilize fully the insight from Merton (1971), an overall measure of the state of the economy is needed. Fortunately, the leading economic indicator (we call it ECON hereafter) of The Conference Board is designed exactly for this purpose, and is widely available today from Bloomberg and other news sources. In the finance literature, ECON has been used as a measure of economic state in earlier studies by Perez-quiros and Timmermann (2000), Lamont, Polk, and Saá-Requejo (2001), Ozoguz (2009), and Lee (2012), but not for forecasting the market risk premium. More importantly, instead of revised data, we use vintage data (those were actually released at the time) that are more appropriate for out-of-sample forecasting.[2]

The second predictor is a technical indicator (TECH hereafter) that measures the mean-reversion behavior of the stock market.[3] Practitioners have long held the view that the stock market fluctuates around its long-term mean. For example, John Bogle (2012), the legendary investor and founder of the Vanguard Group that manages billions of retirement funds for teachers and college professors, says that the number one rule of investing (out of his ten rules) is "Remember reversion to the mean." What is hot today may not be hot tomorrow. The stock market reverts to its long-term mean over the long run. To capture this mean-reversion idea, we simply define the TECH indicator at any given time as the past year cumulative return minus its long term mean, and standardize it by its annualized volatility. The intuition is that, when we look at the market this month, if the cumulative return since one year ago has already been 26%, the stock market will be more likely to go down than to go up next month since the long-term mean is less than 13%.[4]

This paper also makes an econometric contribution to the classical predictive regression model by advocating the use of up- and down-market regressions. While

[2] We are very grateful to Ataman Ozyildirim of The Conference Board for providing us with the vintage data.

[3] This indicator belongs to the domain of Technical Analysis that uses past price or trading volume data to predict future price movements. Brock, Lakonishok, and LeBaron (1992), Neely, Weller, and Dittmar (1997), Sullivan, Timmermann, and White (1999), Lo, Mamaysky, and Wang (2000), Kavajecz and Odders-White (2004), Menkhoff and Taylor (2007), Han, Yang, and Zhou (2012), among others, find significant economic values of technical trading strategies.

[4] This is also in line with various newspapers and investment publications that post 52-week (roughly a year) high, low and return on individual stocks and the market. George and Hwang (2004) and Li and Yu (2012) study the anchor effects of the 52-week high. But we focus here on the market mean-reversion effects which is not necessarily psychologically based.

academic studies on the asymmetry in stock returns in up- and down-markets are fairly recent (see, e.g., Ang and Chen, 2002 and Cooper, Gutierrez, and Hameed, 2004), practitioners have long characterized the stock price movements as up- and down-markets. The most popular concept of an up-market is defined as the periods when the stock market price is above its 200-day moving average price. Otherwise, a down-market occurs. In the finance literature, Cooper, Gutierrez, and Hameed (2004) appear the first to use the up- and down-market regressions in their study of the momentum strategy and find that the profitability depends on the state of the market.[5] However, their definition of the up- and down-markets has nothing to do with the moving averages.

In this paper, we follow practitioners' definition of the up- and down-markets and incorporate them into the standard predictive regression model. The extended model shares the simple feature of the predictive regression model, but allowing for asymmetric reactions of the stock market to its predictors in the up- and down-markets. It nests the usual predictive regression model as a special case if the up- and down-market reactions are the same. Statistically, one may use a different lag, say 100 days, to re-define the moving average, or use an optimal lag that can yield the greatest out-of-sample predictability than reported below. However, to mitigate concerns of data mining and data snooping (see., e.g., Lo and MacKinlay, 1990), we simply use the 200-day moving average that had been used by practitioners for decades before our out-of-sample periods. Indeed, according to Siegel (1994), the analysis of the moving averages goes back at least to the 1930s, and Gordon (1968) finds that, over the period of 1897 to 1967, up to seven times returns can be earned by buying stocks above their moving averages than sticking to the buy-and-hold strategy. Siegel (1994) continues to find the value of the moving average investment strategy

[5] To capture the leptokurtosis of the momentum strategy, Daniel, Jagannathan, and Kim (2012) identify the state of the stock market as "calm" or "turbulent" with a two-state hidden Markov Chain model, and find that severe losses mainly occur in the "turbulent" state.

till the 1990s. In practice, the 200-day moving average has been widely plotted for years in investment letters, trading softwares, and newspapers (such as Investor Business Daily). Economically, since the 200-day moving average is widely followed, its effect might be easy to understand. If enough investors believe it, they may herd on this information, thereby generating impact on the market price (see, e.g., Froot, Schaferstein, and Stein, 1992, and Bikhchandani, Hirshleifer and Welch, 1992), and making it necessary to study the predictability across the up- and down-markets.

Econometrically, our extended predictive regression model with the up- and down-markets may capture some common regime effects of sophisticated econometric models such as those of Hamilton (1989), Perez-Quiros and Timmermann (2000), Lettau and Van Nieuwerburgh (2008), and Tu (2010). We do not use these models for two reasons. First, it is well known that complex models can be counter-productive in out-of-sample forecasting due to estimation errors, which is why the simple predictive regression model is the primary model used in the predictability literature. Second, our definition is what many investors are actually using to assess the market state. It is economically interesting to see how it works in practice.

Empirically, an analysis over the up- and down-markets does reveal some fundamentally different behaviors of the stock market across the market states. For example, a large daily drop of 5% or more in the stock market occurs five times more often in down-markets than in up-markets, and a daily drop of 10% or more happens only in the down-markets in terms of the S&P500 index over January 1959 to December 2011.[6] In addition, the root mean-squared pricing error of the well-known Fama-French (1993) three-factor model increases 70% in the down-markets.

More interestingly, the market's responses to our measure of mean-reversion have different signs across the market state. The regression slope of the market risk premium on TECH is negative with a value of -0.21 in the up-markets, but it becomes

[6]With respect to the Dow Jones Industrial Average, it is about 4 times over the much longer period of May 1896 to December 2011.

positive, 0.41, in the down-markets, implying that a 1% increase in TECH in the down-markets is likely to increase next month's return by 0.41%. To understand the sign in the up-markets, suppose that the market is under-valued (as measured by a negative TECH value). Since investors have abundant capital due to past rising prices and have less constraints in borrowing, they buy aggressively when TECH is negative, which drives the price up and lifts the future return back to its long-term mean or above it. For the sign in the down-markets, there are two intuitive reasons. First, as many investors follow the down-market indicator, they may not buy aggressively in a down-market or even start selling to reduce stock exposure. Second, those investors who use leverage are likely forced to sell as margins relative to asset values are increased. Both explanations contribute to the empirical fact that selling generates more selling in the down-markets, resulting in a positive regression slope of the market risk premium on TECH.

How well do the new predictors, ECON and TECH, predict the market risk premium? The most stringent criterion is the out-of-sample R-square, R^2_{OS}, defined by Campbell and Thompson (2008). When pooling information across 14 commonly used macroeconomic variables, the combination method of Rapach, Strauss, and Zhou (2010) provides an R^2_{OS} of only 1% for monthly market risk premium forecasting. With more sophisticated strategies, the R^2_{OS}s are improved to about 1.3% and 1.8%, respectively, by Ferreira and Santa-Clara (2011) and Neely, Rapach, Tu, and Zhou (2012). In contrast, our simple and intuitive predictors yield an R^2_{OS} of 3.02%, the best to date. Moreover, while existing predictors predict the stock market primarily during recessions and do not have significant predictive power during economic expansions, ECON and TECH predict the market risk premium in both expansions and recessions with R^2_{OS}s of 1.96% and 6.70%, respectively.

An interesting aspect on ECON and TECH is that their predictive power is complementary to each other. When measured by both in- and out-of-sample R-squares, the total R-square is close to the sum of both individual R^2s. This implies that the economic condition captured by ECON is fairly unrelated to the mean-reversion measure. Another interesting aspect is that ECON and TECH can also predict risk premiums significantly on cross-sectional portfolios sorted by size, book-to-market ratio, industry, long- and short-term reversals. Moreover, their predictive ability is robust when TECH is re-defined by the mean reversion measure of individual portfolios.

Why can ECON and TECH forecast the stock market? While the predictability of ECON is obvious since it captures the future state of the economy, the predictability of TECH deserves more discussion. One explanation is the existence of mean-reverting variation in expected returns, which implies that the best prediction of next month's return is the long-term mean plus a correction term that depends on the deviation of current return from the long-term mean, and therefore, generating return predictability (Fama and French, 1988). Empirically, Conrad and Kaul (1988) find that expected stock return is time-varying and reverts back to its mean over time. The source of mean-reversion can be due to investors' asymmetric response to uncertainty on the state of the economy. Theoretically, Veronesi (1999) shows that, when the market shifts between two unobservable states, investors overreact to bad news in good times and underreact to good news in bad times. In this perspective, TECH can be regarded as a proxy of the news. A negative TECH is a bad news that drives investors to overreact in an up-market, suggesting that the return is expected to revert in the future, so the prediction of TECH is negative in an up-market. On the other hand, a positive TECH is a good news that drives investors to underreact in a down-market, suggesting that the current return will continue. However, since the market is more often in an up trend than a down trend, stock returns will be reverting overall. This mean-reversion behavior has been well recognized. For example, Cecchetti, Lam, and Mark (1990), Bessembinder, Coughenour, Seguin, and Smoller

(1995), Pastor and Stambaugh (2012), among others, explicitly use mean-reverting models of stock returns to address various important questions in finance.

The rest of the paper is organized as follows. Section 1 introduces the econometric methodology. Section 2 shows how the two new predictors predict the market risk premium significantly. Section 3 investigates cross-sectional portfolio predictability, which is followed by Section 4 with a brief conclusion.

2.2 Econometric Methodology

The standard predictive regression model for the market risk premium forecast is

$$r_{t+1} = \alpha + \beta x_t + \varepsilon_{t+1}, \tag{2.1}$$

where r_{t+1} is the excess return of the S&P 500 index, x_t is a predictor, and ε_{t+1} is the error term. The out-of-sample forecast of next period's market risk premium based on (2.1) is naturally computed as

$$\hat{r}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t x_t, \tag{2.2}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the ordinary least squares (OLS) estimates of α and β, respectively, based on data from the start of the available sample through t. The in-sample forecast is computed the same as above except that the $\hat{\alpha}_t$ and $\hat{\beta}_t$ are replaced by those estimated by using the entire sample. In the finance literature, almost all of the predictability studies prior to Goyal and Welch (2008), such as Rozeff (1984), Fama and French (1988), and Campbell and Shiller (1988a, 1988b), are based on in-sample results.

With respect to out-of-sample forecasting, however, Goyal and Welch (2008) find that the market risk premium cannot be reliably predicted in the past 30 years with the use of numerous established macroeconomic variables, which raises the question of whether the predictability really exists. One reason for their finding is that the data-generating process for stock returns may be subject to parameter instability and regime shifts, but regression (2.1) does not consider this possibility. Clearly, predictors may have in general different predictive abilities in different time periods or market states.

In this paper, with the idea similar to Cooper, Gutierrez, and Hammed (2004), we consider an extension of the standard predictive regression model (1) by allowing for two states: the up- and down-markets. We follow the wide investment practice and define an up-market as time periods when the market is above its 200-day moving average price, which is defined as

$$I_{up,t} = \begin{cases} 1, & \text{if } P_t \geq \frac{1}{200}\sum_{i=1}^{200} P_{t+1-i}; \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

where P_t is the daily price level of the market index.

To assess the importance of up- and down-markets, we present two characteristics of stock returns in different market states. The first characteristic is about large daily drops of the stock market. Table 2.1 reports the numbers of daily drops for the Dow Jones Industrial Average (DJIA) and the S&P 500 index when the drop is larger than 3%, 5%, and 10%, respectively, where the drop is measured by the daily arithmetic return without dividends. The market state, up or down, on day $t+1$ is determined by the 200-day moving average indicator on date t. The data on DJIA is from May 26, 1896 to December 30, 2011, downloaded from Robert Shiller's homepage. During this period, there are 84 days with return drops larger than 5%, among which, 66 occur in the down-markets and the remaining 18 happen in the up-markets. Moreover, all six daily drops larger than 10% occur only in the down-markets. During the sample period (January 2, 1959 to December 30, 2011) studied in this paper for return predictability, DJIA has 3 daily drops larger than 5% in the up-markets, and 16 such drops in the down-markets. Accordingly, the S&P 500 index has 3 drops in the up-markets, and 19 in the down-markets. These results show a clear difference on large return drops over the up- and down-markets.

The second characteristic is the pricing error of the Fama-French three-factor model on their 25 portfolios sorted by firm size and book-to-market ratio over the up- and down-markets, which is reported in Table 2.2. The portfolio betas are estimated with all data from January 1959 to December 2011, but the pricing errors, alphas, are evaluated in each of the up- and down-markets, respectively. One may interpret that a portfolio is underpriced if its pricing error is positive, and is overpriced if its pricing error is negative. With this interpretation, 12 out of 25 portfolios are overpriced in the up-markets (Panel A), and 9 are overpriced in the down-markets (Panel B). In particular, except for the smallest size and lowest book-to-market ratio portfolio, all the remaining smallest size portfolios are underpriced in the up-markets and overpriced in the down-markets. This suggests that small size portfolios are more likely to be affected by the market state.

To measure the aggregate market pricing error, we investigate the root mean-squared pricing error (RMSE),

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{25} \alpha_i^2}{25}}, \tag{2.4}$$

where α_i is the pricing error for portfolio i. Table 2.2 shows that the root mean-squared error increases from 0.13% in the up-markets to 0.22% in the down-markets, implying that the aggregate pricing error increases 70% around in the down-markets. Both Tables 2.1 and 2.2 suggest that it is important to divide the market into up- and down-markets.

In this paper, to allow for different predictive power of predictors in up- and down-markets, we modify the usual predictive regression as

$$r_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot x_t + \varepsilon_{up,t+1}, & \text{up-market;} \\ \alpha_{down} + \beta_{down} \cdot x_t + \varepsilon_{down,t+1}, & \text{down-market.} \end{cases} \tag{2.5}$$

Apparently, (2.5) nests regression (2.1) as a special case when the up- and down-market reactions are the same.

We follow Campbell and Thompson (2008) and use R^2_{OS} to measure the out-of-sample performance, which is defined as

$$R^2_{OS} = 1 - \frac{\sum_{t=1}^{T}(\hat{r}_t - r_t)^2}{\sum_{t=1}^{T}(\bar{r}_t - r_t)^2}, \tag{2.6}$$

where T is the out-of-sample number, $\hat{r}_t$ is the excess return forecast estimated from regression (2.5), and $\bar{r}_t$ is the historical average return, both of which are estimated using data up to month $t-1$. If the predictor x_t is viable, R^2_{OS} will be positive, which implies a lower mean-squared forecast error (hereafter MSFE) relative to the forecast based on the historical average return. All the forecasts are estimated with the expanding windows approach. We use the first 26 years of data for in-sample training and the remaining 27 years of data for the out-of-sample evaluation. That is, our out-of-sample period starts in January 1985 and ends in December 2011.[7]

If R^2_{OS} is positive, the forecast outperforms the historical average return in terms of the MSFE. The null hypothesis of interest is therefore $R^2_{OS} \leq 0$ against the alternative hypothesis that $R^2_{OS} > 0$. We test this hypothesis by using the Clark and West (2007) MSFE-adjusted statistic. Define

$$f_{t+1} = (r_{t+1} - \bar{r}_{t+1})^2 - [(r_{t+1} - \hat{r}_{t+1})^2 - (\bar{r}_{t+1} - \hat{r}_{t+1})^2]. \tag{2.7}$$

[7] The data on ECON starts from January 1959 but the first vintage is only available in December 1968. So, we have effectively used 16 years of vintage data for the in-sample training.

Then, the Clark and West (2007) MSFE-adjusted statistic is the t-statistic from the regression of f_{t+1} on a constant.

2.3 Empirical Results

In this section, we describe first the data sets and provide the in- and out-of-sample results for predicting the market risk premium. Then, we examine the results over business cycles. Finally, we assess the performance of our predictors in forecasting economic activity.

2.3.1 Data

The data sets span from January 1959 through December 2011. We follow the literature and focus on monthly market risk premium predictability. One reason is that our ECON indicator is released with a monthly frequency, and another reason is that Ang and Bekaert (2007) and Boudoukh, Richardson, and Whitelaw (2008) suggest short horizon predictability to avoid the overestimation issue in long-horizon regressions.

The market risk premium is the log return on the S&P 500 index (including dividends) minus the log return on a risk-free bill. For comparison, we use ten of the most popular predictors in Goyal and Welch (2008).

1. *Dividend-price ratio* (d/p): log of a twelve-month moving sum of dividends paid on the S&P 500 index minus the log of stock prices (S&P 500 index).
2. *Earnings-price ratio* (e/p): log of a twelve-month moving sum of earnings on the S&P 500 index minus the log of stock prices.
3. *Book-to-market ratio* (b/m): book-to-market value ratio for the Dow Jones Industrial Average.
4. *Treasury bill rate* (tbl): three-month Treasury bill rate (secondary market).

5. *Default yield spread* (dfy): difference between Moody's BAA- and AAA-rated corporate bond yields.
6. *Term spread* (tms): difference between the long-term yield on government bonds and the Treasury bill rate.
7. *Net equity expansion* (ntis): ratio of a twelve-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks.
8. *Inflation* (infl): consumer price index (all urban consumers). As in Goyal and Welch (2008), we use inflation with one month lag to account for the delay in releases.
9. *Long-term return* (ltr): return on long-term government bonds.
10. *Stock variance* (svar): monthly sum of squared daily returns on the S&P 500 index.

The ECON indicator, published by The Conference Board on a monthly basis, is constructed to predict economic turning points (peaks and troughs) over the business cycle. As a composite index, ECON consists of ten individual economic leading indicators: 1) average weekly hours (manufacturing), 2) average weekly initial claims for unemployment insurance, 3) manufacturers' new orders (consumer goods and materials), 4) vendor performance (slower deliveries diffusion index), 5) manufacturers' new orders (nondefense capital goods), 6) building permits (new private housing units), 7) stock prices (S&P 500 Index), 8) Money supply (M2), 9) interest rate spread (10-year Treasury bonds less Federal Funds rate), and 10) The Conference Board index of consumer expectations. All of these indicators have an established tendency to decline before recessions and rise before recoveries. Following the standard macroeconomic forecasting, we use the simple linear detrending approach to remove the trend of ECON before using it in the predictive regression. As other macroeconomic indices, the data released today on the past may include possible revisions and adjustments. We follow the common practice by using this data for in-sample forecasting. However, we use the vintage data, which are the actual released data and have no updates, for out-of-sample forecasting to obtain the only practically feasible forecasts over time.

The TECH indicator, as a proxy of mean-reversion, at any time t is defined by

$$\mathrm{TECH}_t = \frac{r_{t-12\to t} - \mu}{\sigma_{t-12\to t}}, \tag{2.8}$$

where $r_{t-12\to t}$ is the cumulative return of the S&P 500 index over the year from month $t-12$ to month t, μ is its long-term mean,[8] and $\sigma_{t-12\to t}$ is the annualized moving standard deviation estimator (Officer, 1973; Mele, 2007), which appears more appropriate to use for investors when they lookback at the volatility over the past year.[9]

2.3.2 Forecasting results

Table 2.3 reports the estimation results of forecasting the market risk premium with the modified predictive regression (2.5). Column 1 presents the predictors. Columns 2 and 4 are regression intercepts in the up- and down-markets. The corresponding regression slopes are shown in Columns 3 and 5. The t-statistics are included in the parentheses. The last two columns are the in-sample and out-of-sample R-squares. For comparison, we first report the results for the ten well-known macroeconomic variables introduced in Section 2, among which, five predictors exhibit significant predictive ability in the up-markets (the default yield spread (dfy), the term spread (tms), the net equity expansion (ntis), the inflation rate (infl), and the stock variance (svar)), and two show significance in the down-markets (the long-term return (ltr)

[8] The long-term mean is calculated with full sample up to date.

[9] We have examined some alternative volatility measures such as the realized volatility in Goyal and Welch (2008) and find that they do not alter the qualitative conclusions.

and the stock variance (svar)). The in-sample R^2s are all positive, ranging from 0.59% for the earnings-to-price ratio (e/p) to 3.26% for the stock variance (svar). In an unreported table where regression (2.1) is used, all the ten predictors deliver smaller R^2s than those in Column 6. The overall performance in Table 2.3 suggests that the predictive ability of macroeconomic variables varies over the market states. However, when we turn to the out-of-sample forecast, only the earnings-price ratio, the inflation and the stock variance show better performance in predicting the market risk premium than the historical average return. In particular, the out-of-sample R^2_{OS} of the stock variance is 1.52%, significant at the 10% level. When the R^2_{OS}s are calculated separately over economic expansions and recessions, their values are 0.21% and 6.10%, respectively. This is consistent with the literature that the predictive power of existing economic variables is mainly from economic recessions.

When ECON is used as the sole predictor, both the up- and down-market regression slopes are negative, implying that an increase in ECON predicts a decrease in the future market risk premium. This may not be surprising since ECON is designed to predict the future economic activity. This finding is consistent with Cooper and Priestley (2009) who find that the output gap (de-trended industrial production) is negatively related to the market risk premium. Interestingly, unlike many other economic variables, the regression slopes of ECON are almost the same in the up- and down-markets, with values of -0.50 and -0.57, respectively. The in-sample R^2 is 1.29%, about the average of all the economic variables. However, its out-of-sample R^2_{OS} is 0.87%, which is significant at the 10% level and is higher than all existing variables except the stock variance (svar). As a leading indicator, it is not a surprise that ECON outperforms pure macro variables such as the long-term (ltr) and the inflation rate (infl). However, the stock variance (svar), as a market-based variable, seems to possess sufficient future information to do better than ECON.

In contrast to ECON, TECH exhibits different predictive patterns across the market states. Its regression slope is negative in the up-markets and positive in the down-markets, and both are significant at the 5% level. This asymmetric predictive ability implies that investors are likely to overreact to bad news in an up-market when there is a negative shock, and underreact to good news in a down-market. The in-sample and out-of-sample R-squares based TECH alone are 3.69% and 2.34%, respectively. In comparison with all the ten well-known predictors, TECH outperforms them both in-sample and out-of-sample. In fact, the only good existing predictor is the stock variance (svar), which has an R^2_{OS} of 1.52% and while others have negative or virtually zero out-of-sample R-squares. The stock variance (svar) is more like a market condition measure and has a correlation of -0.22 with TECH.

To examine the predictive ability of using both ECON and TECH, we run the following regression,

$$r_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot \mathrm{TECH}_t + \beta \cdot \mathrm{ECON}_t + \varepsilon_{up,t+1}, & \text{up-market;} \\ \alpha_{down} + \beta_{down} \cdot \mathrm{TECH}_t + \beta \cdot \mathrm{ECON}_t + \varepsilon_{down,t+1}, & \text{down-market,} \end{cases} \tag{2.9}$$

Note that, to reduce the number of parameters that minimizes estimation error, we do not distinguish the predictive ability of ECON in the up- and down-markets since its slopes are close to each other in the two states, as shown earlier in the case when ECON is used as the sole predictor. The results are reported at the bottom of Table 2.3. The in-sample R^2 is 3.98%, which is the best to date and is about doubling the size of most existing predictors in the literature. Consistent with this evidence, the out-of-sample R^2_{OS} is 3.02%,[10] far greater than 1.3% and 1.8%, the maximum R^2_{OS}s obtained recently by Ferreira and Santa-Clara (2011), and Neely, Rapach, Tu and Zhou (2012) with sophisticated approaches and predictors. Interestingly, the

[10]The in-sample and out-of-sample R-squares are 5.03% and 2.90%, respectively, when ECON is allowed to have different slopes in the up- and down-markets.

out-of-sample R_{OS}^2 of using both ECON and TECH is close to the sum of both, suggesting that the economic condition captured by ECON is fairly unrelated to the mean-reversion behavior of the stock market.

Goyal and Welch (2003, 2008) provide an interesting graphical approach to evaluate the out-of-sample predictive power. This device depicts recursively the residuals to show whether the predictive regression forecast has a lower MSFE than the historical average return for any period by simply comparing the height of the curve at the beginning and end points of the segment corresponding to the period of interest. If the curve is higher at the end of the segment relative to the beginning, the predictive forecast has a lower MSFE during the period. A predictive forecast that always outperforms the historical average will have a slope that is positive everywhere.

Figure 2.1 plots the cumulative sum-squared error from the historical average return forecast minus the cumulative sum-squared error from our competing forecasts. The positive slopes reveal that either ECON, TECH, or both outperforms the historical average forecast consistently over time. This is in contrast to Goyal and Welch (2008) who find that the out-of-sample predictive ability of a number of other economic variables deteriorates markedly after the oil shock of the mid-1970s. Our out-of-sample period, 1985:01–2011:12, also allows us to analyze how our predictors behave over the recent market period characterized by the collapse of the technology bubble and the 2008–2009 mortgage crisis. As shown in Figure 2.1, even during these turbulent periods, either ECON or TECH or they jointly continue to predict the market risk premium. Finally, the cumulative sum-squared errors from using both ECON and TECH are the sum of the errors from using each individually, echoing the earlier results about their R^2s that the sum of using each equals closely to using both simultaneously.

2.3.3 Performance over business cycle

This section investigates whether ECON and TECH can forecast the market risk premium during expansions, as well as recessions. This is of interest since recent studies, such as Rapach, Strauss, and Zhou (2010), Henkel, Martin, and Nardari (2011), and Dangl and Halling (2012), find that the forecasting power of traditional macroeconomic variables is significant during recessions but insignificant during expansions.

We calculate the R^2_{OS}s during expansions and recessions over the NBER business cycle dates, separately. Table 2.4 reports the results. When ECON is the unique predictor in regression (2.1), the R^2_{OS}s are 0.10% and 1.66% during expansions and recessions, respectively, in which 1.66% is significant at the 5% level.[11] In contrast, when the only predictor is TECH, the R^2_{OS} is 1.57% in expansions, significant at the 5% level, and 5.02% in recessions, insignificant.[12] These results indicate that ECON performs better in recessions while TECH works better in expansions. When they are put together, the R^2_{OS}s are 1.96% and 6.70% during expansions and recessions, and are significant at the 5% level and the 10% level, respectively. Again, the R^2_{OS}s from the joint prediction are slightly larger than the sum of individual R^2_{OS}s, implying that ECON and TECH are complementary with each other in the market risk premium forecasting.

Why is the performance of ECON and TECH in predicting the market risk premium weaker during economic expansions than during recessions? One explanation is the countercyclical pattern of the market risk premium. In expansions when consumption, output, and investment are strong, investors are less risk-averse and require a lower premium for risk taking. In addition, they are less constrained and have ample

[11] When the predictive abilities of ECON are distinguished by running (2.5) in up- and down-markets, the out-of-sample R-squares are -0.02% and 3.98% over expansions and recessions. The overall R^2_{OS} is 0.87%, reported in Table 2.3.

[12] The insignificance is likely due to the small sample size, i.e., there are only 34 months during our out-of-sample period identified as economic recessions. However, this small sample size does not alter the significance when ECON and TECH are jointly used.

capital to eliminate any arbitrage opportunity. As a result, any information/news will be incorporated quickly into the market, and the predictability is weaker. On the contrary, in recessions when consumption, output and investment are weak, investors are more risk-averse and require a higher risk premium (Campbell and Cochrane, 1999; Cochrane, 2011). In this case, investors suffer from both leverage and capital constraints (He and Krishnamurthy, 2012). Hence, the lack of sufficient arbitrage capital may limit the speed of news diffusion, resulting in stronger prediction in recessions.

The above intuitive explanation can yield some analytical insights via a simple model. Suppose that the market risk premium is governed by the following process

$$r_{t+1} = \zeta_{s_t}\mu_t + \sigma_{s_t}\varepsilon_{t+1}, \tag{2.10}$$

$$\mu_{t+1} = (1-\rho)\mu_0 + \rho\mu_t + u_{t+1}, \tag{2.11}$$

where $s_t = G, B$ is the business cycle barometer of expansion or recession, and is independent of ε_{t+1} and u_{t+1}. When ζ_{s_t} and σ_{s_t} are constant, the assumed return process reduces to Pastor and Stambaugh (2009) where the risk premium is time-varying and follows an $AR(1)$ process. When μ_t is constant, it reduces to the simplest regime shifting process (Ang and Timmermann, 2012). Our assumption says that the expected return is time-varying and may shift from a high-growth state to a low-growth state at random times (Veronesi, 1999; Ozoguz, 2009).

The Sharpe ratio conditional on business cycle is

$$Sh_{s_t} = \mathrm{E}_{s_t}\left[\frac{\zeta_{s_t}\mu_t}{\mathrm{std}_t(r_{t+1})}\right] = \frac{\zeta_{s_t}\mu_0}{\sigma_{s_t}}. \tag{2.12}$$

According to Lettau and Ludvigson (2010) and Lustig and Verdelhan (2012), the Sharpe ratio is higher in economic recessions. This implies that $\frac{\zeta_B}{\sigma_B} > \frac{\zeta_G}{\sigma_G}$ (the unconditional expected return $\mu_0 > 0$ is self-evident). Moreover, the predictive R-square over business cycles is

$$R_{s_t}^2 = \frac{\mathrm{Var}[E_t(r_{t+1})|s_t]}{\mathrm{Var}(r_{t+1}|s_t)} = \frac{\zeta_{s_t}^2 \mathrm{Var}(\mu_t)}{\zeta_{s_t}^2 \mathrm{Var}(\mu_t) + \sigma_{s_t}^2}. \tag{2.13}$$

Obviously, if $\frac{\zeta_B}{\sigma_B} > \frac{\zeta_G}{\sigma_G}$, then the R-square during recessions must be larger than that during expansions, that is,

$$R_B^2 > R_G^2. \tag{2.14}$$

This helps to understand why stock returns are more predictable in recessions than expansions.

2.3.4 Economic activity forecast

One interesting question is whether ECON and TECH have the power of forecasting economic activity. Ang, Piazzesi, and Wei (2006) and Hong, Torous, and Valkanov (2007) find that financial ratios that forecast the market risk premium can also forecast economic activity.

Following Diebold and Rudebusch (1991), we use the index of industrial production (IP) as the proxy of economic activity. We use the revised IP data as our target to predict. The reason is that revised data is believed to be closer to the truth. Also, as pointed out by McGuckin and Ozyildirim (2004), the use of revised data in lagged values of the dependent variable gives the autoregressive element an advantage vis-a-vis the contribution of our predictors. This may make it more difficult to improve the forecast relative to the lags of IP.

We run the following regression

$$\mathrm{IP}_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot x_t + \varepsilon_{up,t+1}, & \text{up-market;} \\ \alpha_{down} + \beta_{down} \cdot x_t + \varepsilon_{down,t+1}, & \text{down-market,} \end{cases} \tag{2.15}$$

where IP_{t+1} is the growth rate in industrial production and x_t are ECON, TECH, or both. In the calculation of the out-of-sample R^2_{OS}, we adopt the historical sample mean as the benchmark so that the interpretation is consistent with our market risk premium prediction. Table 2.5 reports the regression intercepts, slopes in the up- and down-markets, and the in- and out-of-sample R-squares. When ECON is used as the predictor, it significantly predicts the industrial production growth in the down-markets. The in- and out-of-sample R-squares are 11.18% and 7.47%, respectively.

When TECH is used to forecast IP, its regression slope is significant in the down-markets. Overall, the in-sample R^2 is 18.03% and the out-of-sample R^2_{OS} is 10.33%. Surprisingly, TECH has even better forecasting performance than ECON. One possible reason is that TECH is more timely to forecast IP while ECON is released with one-month lag. When ECON and TECH are jointly used, the R-square is 18.99% for in-sample forecast and 10.61% for out-of-sample forecast.

In brief, we can conclude with a high degree of confidence that both ECON and TECH can predict the future economic activity.

2.4 Portfolio Risk Premium Forecast

In this section, we show that ECON and TECH can also predict risk premiums on cross-sectional portfolios sorted by size, book-to-market ratio, industry, long- and short-term reversals.

2.4.1 Size portfolios

Return predictability on size portfolios has been extensively investigated in the literature (Ferson and Harvey, 1991; Ferson and Korajczyk, 1995; Kirby, 1998) and the basic characteristic is that smaller size portfolios are more predictable than larger size portfolios.

Table 2.6 presents the out-of-sample R^2_{OS}s on the ten decile portfolio risk premium forecasts. We report the results for ECON, TECH, and both, respectively. The market state, up or down, is determined by the 200-day moving average indicator of the S&P 500 index, as defined in (2.3). TECH is constructed by either the returns on the S&P 500 index or the individual portfolio returns. Panel A considers the case when TECH is constructed on the market portfolio. When ECON serves as the predictor, the R^2_{OS}s are generally decreasing from 2.00% for the smallest size portfolio to 0.42% for the largest size portfolio, with one exception of the second smallest size portfolio whose R^2_{OS} is 2.10%. Among the ten portfolios, only the performance for the largest size portfolio is not significant. This finding is consistent with our results in Table 2.2 that small firms are more positively affected by improving economic fundamentals but more vulnerable during economic downturns (Perez-Quiros and Timmermann, 2000).

When TECH is the predictor, it shows significant forecasting ability too. Without a strictly monotonic predictability trend, the R^2_{OS}s increase from small size portfolios to large size portfolios, and are significant at least at the 10% level. The reason for this increasing predictability is the increasing correlation between the size portfolios and the market portfolio, which ranges from 0.55 for the smallest size portfolio to 0.99 for the largest size portfolio. The returns we use in this paper are value-weighted, which implies that a larger size portfolio should have a higher correlation with the market portfolio than a smaller size portfolio. As shown in the previous section, TECH should exhibit stronger predictive ability for portfolios with a larger weight in the market portfolio, since it can significantly forecast the market risk premium. The smallest size portfolio does not have a high correlation with the market portfolio but is predicted by TECH significantly. This may be due to the fact that the small size portfolio is highly mispriced (as shown in Table 2.2). When ECON and TECH are jointly used, all R^2_{OS}s are significant at the 5% level, with values ranging from 2.25%

to 3.00%. As in the market risk premium forecasting, the R^2_{OS} from joint use of ECON and TECH is almost equal to the sum of R^2_{OS}s from individual ECON and TECH forecasts. This implies that ECON and TECH are complementary in predicting size portfolio risk premiums.

Panel B of Table 6 shows the performance when individual portfolio returns are used to construct TECH. That is, the TECH indicator for the first decile portfolio risk premium forecast is constructed by using the first portfolio returns, and so on. In this case, the general pattern is that TECH can still forecast portfolio risk premium but with weaker performance. Six portfolios show positive R^2_{OS}s. The weaker performance is due to the fact that individual size portfolios have different trends. For example, when the 100-day moving average is used as the market state indicator, most R^2_{OS}s are improved significantly. To keep consistent, however, we insist on the 200-day moving average throughout the paper. This phenomenon applies to other component portfolio forecasts too. When ECON and TECH are combined, the R^2_{OS}s range from 1.33% to 2.96% and all are significant at the 5% level.

2.4.2 Book-to-market portfolios

Whether value premium is predictable has received considerable attention in the past two decades. Janannathan and Wang (1996), Pontiff and Schall (1999), and Chen, Petkova, and Zhang (2008) document positive evidence, while Lewellen and Nagel (2006) find that the covariance between the value-minus-growth risk and the aggregate risk premium is small and therefore the value premium is unpredictable. We revisit this problem by considering ten book-to-market portfolio risk premium forecast.

Panel A of Table 2.7 presents the results when TECH is constructed with market portfolio returns. In the second column, ECON delivers increasing R^2_{OS}s from the lowest book-to-market (growth) portfolio to the highest book-to-market (value) portfolio. When TECH is used, the R^2_{OS}s do not show a monotonic pattern but nine are significant at the 5% level and the remaining one is insignificant but positive. The R^2_{OS}s for joint prediction range from 1.97% for the first decile portfolio to 3.59% for the fourth portfolio.

Panel B reports R^2_{OS}s when individual portfolio returns are used to construct TECH. The performance is again weaker than simply using the market portfolio returns. However, all R^2_{OS}s are positive during our sample period and five are significant at least at the 10% level. When TECH is augmented by ECON, all the forecasts are significant at the 10% level.

2.4.3 Industry portfolios

Studies on industry portfolio risk premium are relatively limited. Ferson and Harvey (1991) and Ferson and Korajczyk (1995) consider this problem on a small set of economic variables that serve as predictors. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) provide supporting evidence that some industry portfolios are predictable while others are not.

Table 2.8 provides the results on which industries can be forecasted by ECON and TECH. It is clear that ECON can produce significant R^2_{OS}s on four industry portfolio forecasts, which are consumer durables (Durbl), business equipment (HiTec), Wholesale, retail, and some services (Shops), and Other. In contrast, when TECH is constructed by the market portfolio returns (Panel A), it generates positive R^2_{OS}s for all ten industries and seven of them are significant at the 5% level. When both ECON and TECH are used, nine industries, except Telcm, are significantly predicted at the 10% level. Panel B reports the results when TECH is constructed by the individual portfolio returns. Again, the predictive ability of TECH is weaker in this case. The reason is that we use the 200-day moving average indicator to define the up- or down-market states, which is unlikely to always work due to the fact that different industries may have different up- or down-market cycles and these cycles may not coincide exactly with the market cycles.

2.4.4 Long- and short-term reversal portfolios

For robustness, we investigate further on ten long-term reversal and ten short-term reversal portfolios whose data are readily available from Ken French Library. The long-term reversal portfolios at month t are constructed based on prior returns from month $t-60$ to month $t-13$ while the short-term reversal portfolios are based on the previous month's return.

The results are reported in Tables 2.9 and 2.10. The overall patterns can be summarized as follows. First, ECON shows decreasing predictive ability for the long-term reversal portfolios and stable ability for the short-term reversal portfolios. All the R^2_{OS}s are positive. Second, TECH generates increasing R^2_{OS}s for both long- and short-term reversal portfolios. Third, when our two predictors are simultaneously used, the R^2_{OS}s are significantly improved and are approximately equal to the sum of R^2_{OS}s predicted with individual predictors. With several exceptions, all R^2_{OS}s are significant at the 5% level.

2.5 Conclusion

In this paper, we provide two new predictors, ECON and TECH, that measure the state of the economy and the market condition. We find that the two state variables, complementary to each other, can predict the market risk premium significantly in both up- and down-markets or in both business expansions and recessions. ECON and TECH are simple to compute, easy to interpret and perform far better than numerous predictors found in the large finance literature. Moreover, the same idea can also be applied to forecast portfolio returns sorted on firm size, book-to-market ratio, industry, long- and short-term reversals.

Our study focuses on the stock market. It will be of interest to investigate the predictive ability of similar ECON and TECH predictors in the bond market, commodity market, currency market and international markets. Since the pricing errors of factor models and predictability vary substantially over the up- and down-markets, our study also calls for developing theoretical models to understand them, and for exploring implications on corporate decision making.

2.6 References

Ang, A., and G. Bekaert. 2007. Return predictability: is it there? *Review of Financial Studies* 20: 651–707.

Ang, A., and J. Chen. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63: 443–494.

Ang, A., M. Piazzesi, and M. Wei. 2006. What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131: 359–403.

Ang, A., and A. Timmermann. 2012. Regimes changes and financial markets. *Annual Review of Financial Economics* 4: 313–337.

Baker, M., and J. Wurgler. 2000. The equity share in new issues and aggregate stock returns. *Journal of Finance* 55: 2219–2257.

Bessembinder, H., J.F. Coughenour, P.J. Seguin, and M.M. Smoller. 1995. Mean reversion in equilibrium asset prices: evidence from the futures term structure. *Journal of Finance* 50: 361–375.

Bikhchandani, S., D. Hirshleifer, and I. Welch, 1992. A theory of fads, fashion, custom, and cultural change as informatioal cascades. *Journal of Political Economy* 100: 992–1026.

Bogle, J.C. 2012. *The Clash of the Cultures: Investment vs. Speculation*. John Wiley & Sons.

Boudoukh, J., R. Michaely, M.P. Richardson, and M.R. Roberts. 2007. On the importance of measuring payout yield: implications for empirical asset pricing. *Journal of Finance* 62: 877–915.

Boudoukh, J., M.P. Richardson, and R.F. Whitelaw. 2008. The myth of long-horizon predictability. *Review of Financial Studies* 21: 1577–1605.

Breen, W., L.R. Glosten, and R. Jagannathan. 1989. Economic significance of predictable variation in stock index returns. *Journal of Finance* 44: 1177–1189.

Brock, W., J. Lakonishok, and B. LeBaron. 1992. Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance* 47: 1731–1764.

Campbell, J.Y. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18: 373–399.

Campbell, J.Y., and J.H. Cochrane. 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107: 205–251.

Campbell, J.Y., and R.J. Shiller. 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1: 195–228.

Campbell, J.Y., and R.J. Shiller. 1988b. Stock prices, earnings, and expected dividends. *Journal of Finance* 43: 661–676.

Campbell, J.Y., and S.B. Thompson. 2008. Predicting excess stock returns out of sample: can anything beat the historical average? *Review of Financial Studies* 21: 1509–1531.

Campbell, J., and T. Vuolteenaho. 2004. Inflation illusion and stock prices. *American Economic Review* 94: 19–23.

Cecchetti, S.G., P.S. Lam, and N.C. Mark. 1990. Mean reversion in equilibrium asset prices. *American Economic Review* 80: 398–418.

Chen, L., R. Petkova, and L. Zhang. 2008. The expected value premium. *Journal of Financial Economics* 87: 269–280.

Clark, T.E., and K.D. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138: 291–311.

Cochrane, J.H. 2008. The dog that did not bark: a defense of return predictability. *Review of Financial Studies* 21: 1533–1575.

Cochrane, J.H. 2011. Discount rates. *Journal of Finance* 66: 1047–1108.

Cohen, L., and A. Frazzini. 2008. Economic links and predictable returns. *Journal of Finance* 63: 1977–2011.

Conrad, J., and G. Kaul. 1988. Time-variation in expected returns. *Journal of Business* 61: 409–425.

Cooper, I., and R. Priestly. 2009. Time-varying risk premiums and the output gap. *Review of Financial Studies* 22: 2801–2833.

Cooper, M.J., R.C. Gutierrez Jr., and A. Hameed. 2004. Market states and momentum. *Journal of Finance* 59: 1345–1365.

Dangl, T., and M. Halling. 2012. Predictive regressions with time-varying coefficients. *Journal of Financial Economics* 106: 157–181.

Daniel, K., R. Jagannathan, and S. Kim. 2012. Tail risk in momentum strategy returns. Working paper, Columbia Unversity and Northwestern University.

Diebold, F.X., and G.D. Rudebusch. 1991. Forecasting output with the composite leading index: A real-time analysis. *Journal of the American Statistical Association* 86: 603–610.

Fama, E., and G.W. Schwert. 1977. Asset returns and inflation. *Journal of Financial Economics* 5: 115–146.

Fama, E., and K. French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22, 3–25.

Fama, E., and K. French. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25: 23–49.

Fama, E., and K. French. 1993. Common risk factors in the returns on stock and bonds. Journal of Financial Economics 33: 3–56.

Ferreira, M.I., and P. Santa-Clara. 2011. Forecasting stock market returns: the sum of the parts is more than the whole. *Journal of Financial Economics* 100: 514–537.

Ferson, W., and C. Harvey. 1991. The variation in economic risk premiums. *Journal of Political Economy* 99: 385–415.

Ferson, W., and R.A. Korajczyk. 1995. Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business* 68: 309–349.

Froot, K., Schaferstein, D.S., and J.J. Stein, 1992. Herd on the street: Informational inefficiencies in a market with short-term speculation. *Journal of Finance* 47: 1461–1484.

George, T., and C. Hwang. 2004. The 52-week high and momentum investing. *Journal of Finance* 59: 2145–2176.

Gordon, W. 1968. *The Stock Market Indicators.* Palisades, N.J.: Investors Press.

Goyal, A., and I. Welch. 2003. Predicting the equity premium with dividend ratios. *Management Science* 49: 639–654.

Goyal, A., and I. Welch. 2008. A comprehensive look at the empirical performance of equity premium performance. *Review of Financial Studies* 21: 1455–1508.

Guo, H. 2006. On the out-of-sample predictability of stock market returns. *Journal of Business* 79: 645–670.

Hamilton, J.D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57: 357–384.

Han, Y., K. Yang, and G. Zhou. 2012. A new anomaly: the cross-sectional profitability of technical analysis. *Journal of Financial and Quantitative Analysis*, forthcoming.

He, Z., and A. Krishnamurthy. 2012. Intermediary asset pricing. *American Economic Review*, forthcoming.

Henkel, S.J., J.S., Martin, and F. Nadari. 2011. Time-varying short-horizon predictability. *Journal of Financial Economics* 99: 560–580.

Hong, H., W. Torous, and R. Valkanov. 2007. Do industries lead stock markets? *Journal of Financial Economics* 83: 367–396.

Keim, D.B., and R.F. Stambaugh. 1986. Predicting returns in the stock and bond markets. *Journal of Financial Economics* 17: 357–390.

Jagannathan, R., and Z. Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51: 3–53.

Kavajecz, K., and E. Odders-White. 2004. Technical analysis and liquidity provision. *Review of Financial Studies* 17: 1043–1071.

Kirby, C. 1998. The restrictions on predictability implied by rational asset pricing models. *Review of Financial Studies* 11: 343–382.

Lamont, O., C. Polk, and J. Saá-Requejo. 2001. Financial constraints and stock returns. *Review of Financial Studies* 14: 529–554.

Li, J., and J. Yu. 2012. Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics* 104: 401–419.

Lee, S.S. 2012. Jumps and information flow in financial markets. *Review of Financial Studies* 25: 439–479.

Lettau, M., and S. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56: 815–849.

Lettau, M., and S.C. Ludvigson. 2010. Measuring and modeling variation in the risk-return trade-off. In Ait-Sahalia, Y., Hasen, L.P. (ed.), *Handbook of Financial Econometrics*. Elsevier, pp. 617–690.

Lewellen, J., and S. Nagel. 2006. The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics* 82: 289–314.

Lo, A.W., and A.C. MacKinlay, 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3: 431–467.

Lo, A.W., H. Mamaysky, and J. Wang. 2000. Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementaiton. *Journal of Finance* 55: 1705–1765.

Lustig, H.N., and A. Verdelhan. 2012. Business cyclical variation in the risk-return trade-off. *Journal of Monetary Economics*, forthcoming.

McGuckin, R.H., and A. Ozyildirim. 2004. Real-time tests of the leading economic index: do changes in the index composition matter? *Journal of Business Cycle Measurement and Analysis* 1: 171–191.

Mele, A. 2007. Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of Financial Economics* 86: 446–478.

Menkhoff, L., and M.P. Taylor. 2007. The obstinate passion of foreign exchange professionsls: Technical analysis. *Journal of Economic Literature* 45: 936–972.

Menzly, L., and O. Ozbas. 2010. Market segmentation and the cross-predictability of returns. *Journal of Finance* 65: 1555–1580.

Merton, R.C. 1971. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* 3: 373–413.

Neely, C.J., D.E. Rapach, J. Tu, and G. Zhou. 2012. Forecasting the equity risk premium: the role of technical indicators. Federal Reserve Bank of St. Louis working paper 2010-008C.

Neely, C.J., P. Weller, and R. Dittmar. 1997. Is technical anlaysis in the foreign exchange market profitable? A genetic programming approach. *Journal of Financial and Quantitative Analysis* 32: 405–426.

Nelson, C.R. 1976. Inflation and the rates of returns on common stock. *Journal of Finance* 31: 471–483.

Officer, R.R. 1973. The variability of the market factor of the New York Stock Exchange. *Journal of Business* 46: 434–453.

Ozoguz, A. 2009. Good times or bad times? Investors' uncertainty and stock returns. *Review of Financial Studies* 22: 4377–4422.

Pastor, L., and R.F. Stambaugh. 2009. Predictive systems: living with imperfect predictors. *Journal of Finance* 64: 1583–1628.

Pastor, L., and R.F. Stambaugh. 2012. Are stocks really less volatile in the long run? *Journal of Finance* 67: 431–478.

Perez-Quiros, G., and A. Timmermann. 2000. Firm size and cyclical variations in stock returns. *Journal of Finance* 55: 1229–1262.

Pontiff, J., and L. Schall. 1999. Book to market as a predictor of market returns. *Journal of Financial Economics* 49: 141–160.

Rapach, D.E., J.K. Strauss, and G. Zhou. 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. *Review of Financial Studies* 23: 821–862.

Rozeff, M.S. 1984. Dividend yields are equity risk premiums. *Journal of Portfolio Management* 11: 68–75.

Siegel, J.J. 1994. *Stocks for the Long Run: A Guide to Selecting Markets for Long-Term Growth*. Burr Ridge, Ill.: Irwin Professional Publishing.

Sullivan, R., A. Timmermann, and H. White. 1999. Data snooping, technical trading rule performance, and the bootstrap. *Journal of Finance* 54: 1647–1691.

Tu, J. 2010. Is regime switching in stock returns important in portfolio decisions? *Management Science* 56: 1198–1215.

Veronesi, P., 1999. Stock market overreaction to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies* 12: 975–1007.

Table 2.1: Large market daily drops in up- and down-markets

	$r < -3\%$			$r < -5\%$			$r < -10\%$		
	up	down	overall	up	down	overall	up	down	overall
Panel A: DJIA									
5/26/1896–12/30/2011	122	254	376	18	66	84	0	6	6
5/26/1896–12/31/1958	103	190	293	15	50	65	0	5	5
1/2/1959–12/30/2011	19	64	83	3	16	19	0	1	1
Panel B: S&P 500									
1/2/1959–12/30/2011	20	71	91	3	19	22	0	1	1

This table reports the numbers of daily big drops in up- and down-markets. The up-market state on day $t+1$ is defined by the market index above its 200-day moving average on day t. The market return, r, is the daily simple arithmetic return (excluding dividends).

Table 2.2: Pricing errors (alphas) in up- and down-markets

	Low	2	3	4	High	
Panel A: Pricing errors (alphas) in up-markets						
Small	−0.31	0.08	0.05	0.24	0.23	
2	−0.23	−0.12	0.06	0.04	−0.06	
3	−0.11	0.04	−0.06	0.05	0.02	
4	0.07	−0.18	−0.07	0.06	−0.09	
Big	0.20	0.03	−0.10	−0.12	−0.08	
Root mean-squared error (RMSE)						0.13
Panel B: Pricing errors (alphas) in down-markets						
Small	−0.80	−0.18	−0.10	−0.01	−0.11	
2	−0.03	0.15	0.21	0.19	0.10	
3	0.06	0.16	0.15	0.08	0.20	
4	0.24	0.16	0.12	0.11	−0.14	
Big	0.09	0.05	0.16	−0.16	−0.38	
Root mean-squared error (RMSE)						0.22

This table reports the pricing errors (alphas) of the Fama-French three-factor model in up- and down-markets on the the Fama-French 25 portfolios formed on size and book-to-market ratio:

$$R_t - R_t^f = \alpha + \beta_1 \cdot (R_{mt} - R_t^f) + \beta_2 \cdot \mathrm{SMB}_t + \beta_3 \cdot \mathrm{HML}_t + \varepsilon_t.$$

We estimate the betas with the entire sample (January 1959 to December 2011) and then calculate the pricing errors in up- and down-markets, respectively. The root mean-squared pricing error (RMSE) is defined as RMSE $= \sqrt{\sum_{i=1}^{25} \alpha_i^2 / 25}$.

Table 2.3: Forecasts of market risk premium

x_t	α_{up}	β_{up}	α_{down}	β_{down}	R^2	R^2_{OS}
d/p	0.01	0.00	0.05	0.01	2.47	-0.38
	(0.69)	(0.38)	(1.58)	(1.63)		
e/p	0.00	-0.00	0.02	0.01	1.10	0.09
	(0.23)	(-0.13)	(0.65)	(0.69)		
b/m	0.01	-0.00	-0.01	0.01	2.47	-0.51
	(1.36)	(-0.06)	(-0.93)	(0.87)		
tbl	0.01	-0.11	-0.00	0.01	1.23	-0.16
	(3.20)	(-1.54)	(-0.25)	(0.06)		
dfy	-0.00	1.02	-0.00	-0.08	1.16	-0.18
	(-0.94)	(2.29)	(-0.07)	(-0.08)		
tms	0.00	0.25	-0.00	0.08	0.59	-1.09
	(0.45)	(1.95)	(-0.54)	(0.33)		
ntis	0.01	-0.20	-0.00	0.16	2.35	-0.28
	(3.74)	(-2.11)	(-0.69)	(0.75)		
infl	0.01	-1.45	-0.00	0.15	1.73	0.17^{*}
	(4.46)	(-2.31)	(-0.37)	(0.16)		
ltr	0.01	0.10	-0.00	0.19	0.79	-0.54
	(2.72)	(1.31)	(-0.76)	(1.93)		
svar	0.00	2.57	0.00	-1.19	3.26	1.52^{*}
	(1.08)	(1.70)	(0.82)	(-2.56)		
ECON	0.01	-0.50	-0.00	-0.57	1.29	0.87^{*}
	(3.25)	(-1.83)	(-0.40)	(-1.01)		
TECH	0.01	-0.21	0.01	0.41	3.69	2.34^{***}
	(3.98)	(-2.02)	(1.75)	(1.99)		
ECON+ TECH					3.98	3.02^{***}

This table reports the results of forecasting the market risk premium:

$$r_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot x_t + \varepsilon_{up,t+1}, & \text{up market;} \\ \alpha_{down} + \beta_{down} \cdot x_t + \varepsilon_{down,t+1}, & \text{down market,} \end{cases}$$

where r_{t+1} is the log return (including dividends) on the S&P 500 index minus the log return on a risk-free bill, x_t is one of predictors given in the first column. ECON is the leading economic indicator. TECH is the past year cumulative return minus its long term mean and standardized by its annualized moving standard deviation. R^2 is the in-sample R-square and R^2_{OS} is the Campbell and Thompson (2008) out-of-sample R-square over 1985:01–2011:12. The values in parentheses are the t-statistics. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. *** and * indicate significance at the 1% and 10% level, respectively.

Table 2.4: Forecasts of market risk premium over business cycles

	ECON	TECH	ECON+TECH
Expansion	0.11	1.57**	1.96**
Recession	1.66**	5.02	6.70*
Overall	0.45	2.34**	3.02***

This table reports the Campbell and Thompson (2008) out-of-sample R^2_{OS}s in predicting market risk premium with ECON and TECH over 1985:01–2011:12. The R^2_{OS}s are calculated based on the NBER-dated periods of expansions and recessions, respectively. The results in the column of ECON are based on the model

$$r_{t+1} = \alpha + \beta \cdot \mathrm{ECON}_t + \varepsilon_{t+1},$$

in the column of TECH are based on

$$r_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot \mathrm{TECH}_t + \varepsilon_{up,t+1}, & \text{up-market;} \\ \alpha_{down} + \beta_{down} \cdot \mathrm{TECH}_t + \varepsilon_{down,t+1}, & \text{down-market,} \end{cases}$$

and in the column of ECON+TECH are based on

$$r_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot \mathrm{TECH}_t + \beta \cdot \mathrm{ECON}_t + \varepsilon_{up,t+1}, & \text{up-market;} \\ \alpha_{down} + \beta_{down} \cdot \mathrm{TECH}_t + \beta \cdot \mathrm{ECON}_t + \varepsilon_{down,t+1}, & \text{down-market,} \end{cases}$$

where r_{t+1} is the log return (including dividends) on the S&P 500 index minus the log return on a risk-free bill. ECON is the leading economic indicator. TECH is the past year cumulative return minus its long term mean and standardized by its annualized moving standard deviation. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 2.5: Forecasts of industrial production growth

x_t	α_{up}	β_{up}	α_{up}	β_{down}	R^2	R^2_{OS}
ECON	0.38	0.94	-0.08	3.57	11.18	7.47***
	(9.45)	(1.63)	(-1.22)	(4.40)		
TECH	0.36	3.19	0.20	10.27	18.03	10.33***
	(7.73)	(1.33)	(2.25)	(3.35)		
ECON + TECH					18.99	10.61***

This table reports estimates from OLS regressions of current industrial production growth rates on ECON, TECH, or both, i.e.,

$$\mathrm{IP}_{t+1} = \begin{cases} \alpha_{up} + \beta_{up} \cdot x_t + \varepsilon_{up,t+1}, & \text{up-market;} \\ \alpha_{down} + \beta_{down} \cdot x_t + \varepsilon_{down,t+1}, & \text{down-market.} \end{cases}$$

ECON is the leading economic indicator. TECH is the past year cumulative return minus its long term mean and standardized by its annualized moving standard deviation. R^2 is the in-sample R-square and R^2_{OS} is the Campbell and Thompson (2008) out-of-sample R-square over 1985:01–2011:12. The t-statistics are reported in parentheses. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. *** indicates significance at the
1% level.

Table 2.6: Forecasts of size portfolios

	ECON	TECH	ECON+TECH
	Panel A: Market TECH		
Small	2.00***	1.21**	2.28***
2	2.10***	0.20*	2.48***
3	1.95***	0.80**	3.00***
4	1.76***	0.94*	2.93***
5	1.78***	0.84*	2.84**
6	1.67***	0.91*	2.82***
7	1.12**	0.83*	2.25**
8	1.05**	1.28**	2.64***
9	0.85**	1.40**	2.54**
Large	0.42	2.35**	2.96***
	Panel B: Individual TECH		
Small	2.00***	1.06**	2.79***
2	2.10***	-0.04	2.12***
3	1.95***	0.14	2.08***
4	1.76***	-0.01	1.84**
5	1.78***	-0.01	1.74**
6	1.67***	-0.09	1.58**
7	1.12**	0.21	1.33**
8	1.05**	0.38	1.56**
9	0.85**	0.73*	1.73**
Large	0.42	2.43**	2.96***

This table reports the Campbell and Thompson (2008) out-of-sample R^2_{OS}s of predicting size portfolio risk premiums over 1985:01–2011:12. ECON is the leading economic indicator. TECH is constructed by the S&P 500 index in Panel A and by individual portfolio in Panel B. The results in columns ECON, TECH, and ECON+TECH are based on the models in Table 2.4. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 2.7: Forecasts of book-to-market portfolios

	ECON	TECH	ECON+TECH
	Panel A: Market TECH		
Low	0.42	1.48**	1.97**
2	0.67*	1.86**	2.84***
3	0.58	2.00***	2.88***
4	0.47	2.70**	3.59***
5	0.65*	1.71**	2.68***
6	0.57*	1.68**	2.50**
7	0.62*	2.25**	3.09***
8	0.64**	2.63**	3.56**
9	0.73**	2.45**	3.54***
High	1.07***	1.50	2.81**
	Panel B: Individual TECH		
Small	0.42	1.82**	2.22**
2	0.67*	1.40**	2.47***
3	0.58	1.79**	2.79***
4	0.47	0.64	1.29**
5	0.65*	1.23*	2.21**
6	0.57*	0.10	0.98*
7	0.62*	1.26**	2.08**
8	0.64**	0.15	1.01*
9	0.73**	0.12	1.04*
Large	1.07***	0.47	1.66**

This table reports the Campbell and Thompson (2008) out-of-sample R^2_{OS}s of predicting book-to-market portfolio risk premiums over 1985:01–2011:12. ECON is the leading economic indicator. TECH is constructed by the S&P 500 index in Panel A and by individual portfolio in Panel B. The results in columns ECON, TECH, and ECON+TECH are based on the models in Table 2.4. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 2.8: Forecasts of industry portfolios

	ECON	TECH	ECON+TECH
	Panel A: Market TECH		
NoDur	-0.11	2.94***	3.33***
Durbl	1.31***	0.59	2.22**
Manuf	0.25	1.87**	2.32**
Enrgy	-0.16	2.24***	2.08***
HiTec	0.45*	0.55	0.98*
Telcm	-0.02	0.25	0.22
Shops	1.66**	1.13**	3.41***
Hlth	0.20	1.28**	1.59**
Utils	-0.39	2.50***	2.11***
Other	1.05**	2.05**	3.41***
	Panel B: Individual TECH		
NoDur	-0.11	-0.35	-0.42
Durbl	1.31***	-0.14	1.26*
Manuf	0.25	1.51**	1.81**
Enrgy	-0.16	1.06**	0.83*
HiTec	0.45*	-0.27	0.19
Telcm	-0.02	0.43	0.08
Shops	1.66**	-0.13	1.77**
Hlth	0.20	0.61	0.77
Utils	-0.39	0.00	-0.77
Other	1.05**	0.79	2.21**

This table reports the Campbell and Thompson (2008) out-of-sample R^2_{OS}s of predicting industry portfolio risk premiums over 1985:01–2011:12. ECON is the leading economic indicator. TECH is constructed by the S&P 500 index in Panel A and by individual portfolio in Panel B. The results in columns ECON, TECH, and ECON+TECH are based on the models in Table 2.4. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 2.9: Forecasts of long-term reversal portfolios

	ECON	TECH	ECON+TECH
		Panel A: Market TECH	
Low	1.36***	0.98*	2.63***
2	1.72***	0.97*	3.12***
3	0.80*	1.18**	2.31**
4	1.43***	1.27*	3.12***
5	0.72*	1.53*	2.58**
6	0.60*	2.40***	3.30***
7	0.38	1.83**	2.43**
8	0.49*	2.47**	3.28***
9	0.25	3.22**	3.75***
High	0.38	2.40**	2.85***
		Panel B: Individual TECH	
Low	1.36***	-0.36	1.04*
2	1.72***	-0.07	2.45**
3	0.80*	-0.34	1.08*
4	1.43***	0.13	2.07**
5	0.72*	0.92*	1.97**
6	0.60*	1.33**	2.08**
7	0.38	0.45	0.93*
8	0.49*	0.80*	1.57**
9	0.25	0.80*	1.09**
High	0.38	1.75**	2.08**

This table reports the Campbell and Thompson (2008) out-of-sample R^2_{OS}s of predicting long-term reversal portfolio risk premiums over 1985:01–2011:12. ECON is the leading economic indicator. TECH is constructed by the S&P 500 index in Panel A and by individual portfolio in Panel B. The results in columns ECON, TECH, and ECON+TECH are based on the models in Table 2.4. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 2.10: Forecasts of short-term reversal portfolios

	ECON	TECH	ECON+TECH
	Panel A: Market TECH		
Low	0.54**	0.57	1.28**
2	0.55*	0.56	1.32*
3	0.56*	1.83**	2.55**
4	0.47	2.16**	2.82**
5	0.15	1.89**	2.25**
6	0.40	1.80**	2.45**
7	0.63*	1.39**	2.22**
8	0.47	2.51**	3.29***
9	0.71*	2.01**	2.86***
High	0.52*	2.53***	3.15***
	Panel B: Individual TECH		
Low	0.54**	-0.22	0.31
2	0.55*	-0.37	0.19
3	0.56*	1.12*	1.82**
4	0.47	0.38	1.11*
5	0.15	0.85*	1.07*
6	0.40	1.22**	2.25***
7	0.63*	0.78*	1.53**
8	0.47	2.48**	3.12***
9	0.71*	2.42**	3.11***
High	0.52*	2.20**	2.72***

This table reports the Campbell and Thompson (2008) out-of-sample R^2_{OS}s of predicting short-term reversal portfolio risk premiums over 1985:01–2011:12. ECON is the leading economic indicator. TECH is constructed by the S&P 500 index in Panel A and by individual portfolio in Panel B. The results in columns ECON, TECH, and ECON+TECH are based on the models in Table 2.4. Statistical significance for R^2_{OS} is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

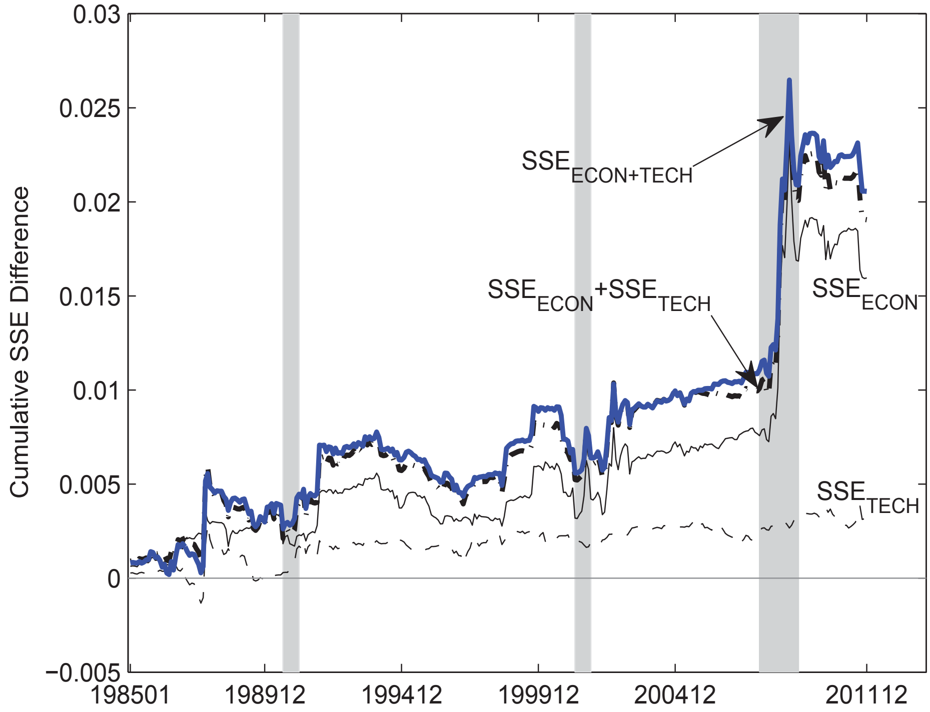

Figure 2.1: Cumulative relative out-of-sample sum-squared errors (SSE)

This figure depicts differences in the cumulative squared forecast errors for the historical average forecast relative to the predictive regression forecast based on ECON and TECH. Vertical bars depict the NBER-dated recessions. $\mathrm{SSE}_{\mathrm{TECH}}$ indicates that TECH is the predictor in the predictive regression, $\mathrm{SSE}_{\mathrm{ECON}}$ indicates that ECON is the predictor, and $\mathrm{SSE}_{\mathrm{ECON+TECH}}$ indicates that both ECON and TECH are used as the predictors. $\mathrm{SSE}_{\mathrm{ECON}} + \mathrm{SSE}_{\mathrm{ECON}}$ is the sum of $\mathrm{SSE}_{\mathrm{ECON}}$ and $\mathrm{SSE}_{\mathrm{TECH}}$.

Chapter 3

Side-by-Side Management

We present a common agency model to study side-by-side (SBS) management when a manager simultaneously manages two funds and separately contracts with the two fund principals. The contracting is decentralized and includes two types of externalities: the manager's efforts are substitutable and the performance in one fund generates a spillover effect on the other fund, which implies that the two principals can choose competition or free-riding at the equilibrium. Under public contracting, competition is more likely to dominate free-riding. Under private contracting, however, free-riding becomes more important. In either case, SBS could generate better performance than standalone management.

3.1 Introduction

Side-by-side management (SBS) refers to the growing financial practice that the same fund manager/fund family simultaneously manages two or more funds, and separately contracts with the two or more fund principals via different compensation schemes.[1] Critics contend that this practice creates conflicts of interest that SBS managers may favor one fund over the other because of limited and substitutable efforts. Even worse, an SBS manager may directly shift performance from the lower

[1] In the hedge fund industry, each manager operates on average 1.84 funds according to the AltVest database (Kolokolova, 2011). In contrast, in the equity mutual fund industry according to the CRSP Mutual Fund Database, the average number of funds under SBS management is 2.2 (Yadav, 2011). Also, the proportion of SBS managers amounts to 30%, managing 36% of the total assets (Agarwal and Ma, 2012).

incentivized fund to the higher incentivized fund.[2] However, industry proponents argue that investors may benefit from their affiliation funds due to spillover effect if the funds are delegated to a skilled manager or if one of the funds under SBS management shows star performance. In addition, with reputation effect and SEC regulation, SBS managers many have a strong incentive to avoid favoritism.[3]

Since SBS management directly influences fund performance and hence shareholders' interest, it is logical to investigate the effects that SBS managers will have on the SBS funds. What are the costs and benefits of SBS management and why do some fund investors/principals choose to delegate different fund management to a common fund manager instead of the traditional exclusive managers (standalone management)? With SBS management, what kind of contract should the principals offer, public contracting or private contracting?

This paper develops a common agency model to answer these questions by investigating the benefits and the costs of SBS management and to find the condition under which SBS management is beneficial relative to standalone management. In my model, two risk-neutral fund principals have two channels to delegate the fund management to risk-neutral managers who are skilled and have some special ability to generate abnormal returns. The principals can use the traditional standalone management and sign exclusive contracts with two different managers. This is the standard one principal-one agent delegation. Alternatively, they can choose SBS management, i.e., they delegate the two funds and sign different contracts with a common manager separately. In so doing, they may enjoy a spillover performance among SBS funds, which is well recognized in practice. For example, effort may generate a positive externality, i.e., a manager may efficiently exert efforts and transfer information across funds by holding common stocks among the different funds. This is achieved

[2]Cici, Gibson and Moussawi (2010) list six examples where favoritism can take different forms.

[3]As of 2006, all managers are required to disclose all of the funds and accounts they manage in the Statements of Additional Information.

by holding his best stocks in all of his funds he has better stock-picking or timing skill. With common holdings, he can spend his limited efforts on concentrated stocks and hence obtains more precise information. Another type of spillover is performance externality. Yadav (2011) finds that a fund receives 7.8% higher new money inflows per year if one of the other funds under SBS management is a star fund.

However, the concern of SBS management is that both principals face conflicts of interest that prevent them from dealing with the fund management at arms-length. With different incentives, the manager may favor one principal that offers him more incentives over the other one.[4] That is, SBS management may raise favoritism, unequal trading costs, different trading priorities and disproportionate allocations of securities, even among funds with nearly identical objectives and investment philosophies. To capture this concern by SEC and investors, I assume in this paper that the manager's efforts exerted on the two funds are substitutable (Homstrom and Milgrom, 1991; Peng and Röell, 2008; Liang and Nan, 2011), which generates an indirectly negative externality between the two contracts, implying that an increase in one fund's incentive will lead to an increase in the marginal cost of the other fund.

Under SBS management, the two principals can choose either public contracting or private contracting. Under public contracting, the two fund principals make the efforts public and contract on efforts exerted on the two funds. Under private contracting, each principal cannot observe the effort exerted on the other fund and hence can only control the effort on her own fund. Given the other fund's contract, the fund principal can compete by providing more incentives to the manager, or reduce the incentives by free-riding but enjoying the spillover effect. My model predicts that competition is more likely to be dominant in the public contracting while free-riding occurs more frequently in the private contracting. This in turn implies that public

[4] "Making special efforts to [...] keep them engaged in their work and the company, as well as appropriately rewarded," Hewitt reports, "getting a Bonus Instead of a Raise", The Wall Street Journal, December 29, 2004.

contracting is more likely to occur when the SBS funds are from the same fund family, whereas the private contracting is more often to occur when the SBS funds are from different fund families.[5]

It should be noted that, under standalone management, the principal' expected payoff corresponding to a given equilibrium effort is well defined, but under SBS management, this value needs to be defined more precisely either with full or incomplete information. Under SBS management, the binding participation constraint of the least-efficient manager determines only the sum of performance generated to both principals. There is an indeterminancy in the exact sharing of this surplus, and therefore only the sum of the principals' payoff is determined. If a specific sharing rule is not chosen, it will introduce multiple perfect Bayesian equilibria of the game in which SBS management is chosen, as long as the corresponding sharing rules satisfy both principals' reservation constraints. Because the principals' total payoff and the levels of effort are the same under any of these equilibria, there is no loss of generality in restricting ourselves to the symmetric, differential equilibria.

Another striking characteristic of SBS is that there are multiple symmetric, differential equilibra even in the case of complete information when the contracting is private. The reason underlying is that principals in SBS can provide nonlinear compensations that are unchosen in equilibrium, which generates intense competition between the principals, serving as implicit threats to prevent the rival principal from deviating from the equilibrium allocation. That is, these out-of-equilibrium offers are irrelevant to the offering principal's payoff but have impact on the rival principal's strategy.

The model proposed in this paper can reconcile and rationalize the contradicting evidence in Cici, Gibson, and Moussawi (2010) and Nohel, Wang, and Zheng (2010).

[5] This is consistent with Chen, Hong, Jiang and Kubik (2012) who document that mutual fund families outsource the management of their funds to unaffiliated advisory firms and find that funds managed externally significantly under-perform those run internally. They attribute this to the inability for fund families to coordinate with the external firm who manage other funds simultaneously.

Cici, Gibson, and Moussawi (2010) track 71 advisory *firms* engaged in both hedge and mutual fund management and find SBS mutual funds *underperform* unaffiliated mutual funds. In a contemporaneous paper, however, Nohel, Wang, and Zheng (2010) study 344 portfolio *managers* who manage a total of 693 mutual funds and 538 hedge funds simultaneously, and find that SBS mutual funds *outperform* unaffiliated funds and SBS hedge funds perform at best on par as good as unaffiliated funds. Agarwal and Ma (2011) examine the determinants and consequences of SBS in the mutual fund industry and find that well-performing and more experienced managers are more likely to switch to SBS by either taking over other funds within fund companies (i.e., acquired funds) that are poorly performing or launching new funds, and that funds managed prior to multitasking experience significant performance deterioration improve within two years after the switch. Yadav (2011) examines the portfolio management strategies of SBS mutual fund managers who are classfied to be low-match managers or high-match managers according to whether they hold low or high fraction of common holdings across the different funds they manage. He shows that high-match managers perform significantly better than low-match managers. The star performance of a fund results in high level of new money flows not only to the fund itself but also to the other funds managed by the manager.[6]

This paper differs from the traditional centralized contracting where the manager has no other option than accepts or refuses the contract offered by the principal. The participation constraint is most often modeled by an exogenously reservation utility for the manager. Under decentralized contracting, the manager may either accept all contracts at once or accept only a subset of the offers he receives. Also, a major difficulty in the common agency literature is to understand the new frictions

[6]Our paper is also related to the literature on fund family where performance spillover and favoritism are well acknowledged (Massa, 2003; Nanda, Wang and Zheng, 2004; Gaspar, Massa and Matos, 2006; Elton, Grumber and Green 2007; Dangl, Wu and J. Zechner (2008); Ruenzi and Kempf, 2008; Agarwal, Boyson and Naik, 2009; Gavazza, 2011; Brown and Wu, 2011; and Sialm and Tham, 2011).

due to the principals' non-cooperative behaviors. This requires a careful study of the contracting possibilities available to the principals, which may either alleviate or exacerbate distortions depending on the contexts. Readers may refer to Bernheim and Whinston (1986a), Dixit (1997), Martimort (1996), Martimort (2007), Martimort and Stole (2009) for common agency theory.

The rest of the paper proceeds as follows. Section 2 introduces the basic model. Section 3 discusses standalone management, which will serves as a benchmark. Section 4 studies SBS management under complete information, which is followed by Section 5 with incomplete information. Section 6 extends the model to the case where the SBS manager has ownerships in both funds. Section 7 summarizes the findings and concludes.

3.2 Model

This section describes the sequence of events that arises SBS and introduces a common agency model to analyze the effects of SBS on the manager's effort exertion and fund performance. For simplicity, I assume that all players are risk-neutral throughout the paper.

Suppose in the market there are two funds, A and B. To obtain abnormal returns, the fund principals delegate the funds to managers who have some scarce skills. There are two mechanisms for delegation, standalone management and SBS management. With standalone management, the two fund principals delegate their funds to two exclusive fund managers who will respectively generate risk-adjusted returns as

$$r_A = e_A + \varepsilon_A, \tag{3.1}$$

$$r_B = \sigma e_B + \varepsilon_B, \tag{3.2}$$

where e_A and e_B are efforts exerted by fund managers, and ε_A and ε_B are two i.i.d residuals with mean zero and variances σ^2.[7] Since everyone is risk-neutral, the mean-zero residual terms and their distributions have no effect on the results, but are introduced to demonstrate their generality (Laffont and Tirol, 1986). Also, this paper assumes that standalone managers have the same cost function as $C(e) = \frac{e^2}{2}$.

With SBS management, the two fund principals delegate the two funds to one common manager and contract with the manager separately. The benefit is the spillover effect between the two funds, i.e, the effort exerted on one fund has a positive effect on the other fund. The cost is the substitutable effect of efforts, i.e., the manager opts to allocate his limited efforts to align with each specific principal. To capture these two characteristics, I assume that the fund returns follow

$$r_A = e_A + \theta e_B + \varepsilon_A, \tag{3.3}$$

$$r_B = \sigma(e_B + \theta e_A) + \varepsilon_B, \tag{3.4}$$

and that the cost function is

$$C(e_A, e_B, \delta) = \frac{1}{2}(e_A^2 + e_B^2) + \delta e_A e_B,$$

where $\theta \in (0, 1)$ measures the spillover effect and $\delta \in [0, 1]$ captures the substitute effect. When a fund principal raises the manager's incentive to attract more efforts, the manager will reduce his effort to the other fund principal since his skills are substitutable. Specifically, when $\delta = 0$, the efforts are independent of each other and SBS management has an obvious benefit due to the spillover effect. However, when δ approaches ρ, the efforts for the two funds are perfectly substitutable and therefore, the manager will exert all his efforts on the fund with more incentives. The

[7] Generally, we may assume the fund excess return as $R = \sigma e + \beta(R_m - R_f) + \varepsilon$. This implies that σe is the fund alpha or the risk-adjusted return.

substitutable assumption ensures that the manager's efforts are limited ($\frac{\partial C(e_A,e_B,\delta)}{\partial e_A \partial e_B} \geq 0$). This is in the spirit of Homstrom and Milgrom (1991), Peng and Röell (2008) and Liang and Nan (2011) who consider the one principle one agent with multitask problems, while I explore the multi-principal one agent contract.

The key principle in SBS management is the desirability of keeping a balance between incentives across funds to avoid a form of fund arbitrage by the manager that results in one fund being neglected, i.e, when higher effort on one fund raises the marginal cost of effort on the other fund. In the absence of conflict, SBS is beneficial for both fund principals who prefer to hiring a common manager to manage the funds rather than hiring one manager per fund. However, introducing conflicts change the results because the manager may extract more rents from one fund by exerting more effort.

An alternative explanation for the effort substitutability is that the manager's ability consists of two components, market-timing ability (e_A) and stock-picking ability (e_B). Each type of ability can generate a positive abnormal return to the two funds, but the contributions are different. With the assumption of $\theta \in (0, 1)$, fund A is perhaps more dependent on the market-timing ability while fund B may be more dependent on the stock-picking ability. The difference may be due to the fact that each fund uses different investment strategies and is subject to different restrictions. For example, fund B is a mutual fund and hence is subject to short-sale constraint. As a result, fund B's performance is more likely to be attributed by the manager's stock-picking ability. Instead, fund A is a hedge fund and its performance is more likely to be attributed by the manager's market-timing ability.

Both the spillover effect θ and the substitute parameter δ can be used to measure the manager's type. The larger θ or the smaller δ, the higher the manager's ability. For simplicity, I fix θ and only use δ to define the manager's type throughout the paper. The information asymmetry naturally means that the manager's ability δ is only known to the manager but unknown to both fund principals. However, I assume that it follows a special distribution that is common knowledge among the two fund principals and the SBS manager.

The SBS manager has a quasi-linear utility function:

$$U(\delta) = w_A + w_B - C(e_A, e_B, \delta), \tag{3.5}$$

where w_A and w_B are compensation schedules that will be specified in detail according to the contract context. Specifically, I will consider in this paper two compensation schemes with SBS management, public contracting or private contracting. Under public contracting, both fund principals can contract with the manager on both e_A and e_B, i.e., $w_A = w_A(e_A, e_B)$ and $w_B = w_B(e_A, e_B)$. Under private contracting, however, principal A can only contract with the manager on effort e_A and e_B is unobservable to him, i.e., $w_A = w_A(e_A)$, and principal B can only contract on effort e_B, i.e., $w_B = w_A(e_B)$.

Although the two principals can change the compensation w to change the manager's incentives, SBS management introduces some new features that are not present under centralized contracting. For example, the manager's effort contracted upon by principal A enters directly into principal B's objective function, and vice versa. This arises a direct contractual externality (Martimort and Stole, 2003). Another type of externality occurs because the manager's cost function for fund A depends also on contractual effort e_B. With these two externalities, there may have multiple equilibria in the case of private contracting even under complete information. The multiplicity is driven by the nonlinear compensations available to the manager that will not be chosen in equilibrium. Offering a nonlinear compensation schedule that is unchosen by the manager can prevent the other fund principal from deviating from the equilibrium (Martimort and Stole, 2003). With SBS as a special characterization of common agency, the two compensation schemes are not determined at the equilibrium although the sum is determined that jointly satisfies the manager's participation constraint. In this sense, I only consider symmetric, differential equilibriums that equally distribute the joint payoff at the equilibrium between the two fund principals. In this sense, SBS is beneficial if the joint payoff is larger than the sum of individual payoffs from standalone management. Lastly, I assume that when the two principals decide SBS management, the manager cannot reject one but accept the other one, thereby he can only reject or accept both.

To sum up, the sequence of events is illustrated in Figure 1 and proceeds as follows:

1. Nature draws δ. This parameter is known only by the SBS manager in the case of asymmetric information or by all players under complete information. The manager reports $\tilde{\delta}$ to both fund principals.[8]

2. Principals choose the contracting scheme, standalone or SBS. If they choose SBS management, they need a new step to choose either public contracting or private contracting.[9] Then they propose separate contracts to the manager.

3. The common manager accepts or rejects both contracts.

4. If the SBS manager rejects, he gets his reservation utility zero. If he accepts, he chooses efforts to exert on both funds. Then, the final payoffs realize.

I am ready now to define the equilibrium of SBS as a triplet that includs the manger's effort level and the two incentive schemes offered by the two principals, (e^*, w_A^*, w_B^*), such that

[8] The manager could reports different values to the two principals. With truthful equilibrium, the manager will only report the same value.

[9] I will not consider the situation where one fund principal uses public contracting while the other principal uses private contracting, which could be the third type of SBS management. Maier and Ottaviani (2009) show that the welfare with this semi-public contracting is the same as public contracting.

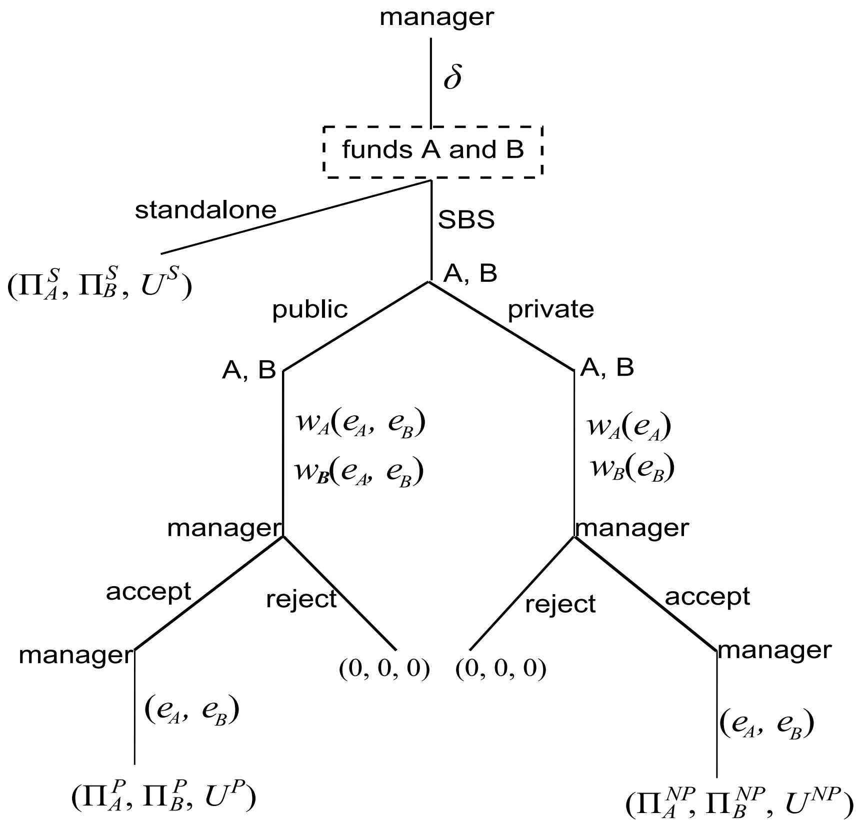

Figure 3.1: Sequence of events.

1. the manager chooses effort levels (e^*) to maximize his expected utility, taking the incentive schemes offered by the two principals as given, and

2. each principal offers an incentive scheme (w_A^* or w_B^*) that gives her the highest expected payoff, taking as given the incentive scheme provided by the other principal and the manager's optimal efforts.

3.3 Standalone Management

For comparison, in this section I consider standalone management where one manager is only allowed to manager one fund (Bhattacharya and Pfleiderer, 1985; Admati and Pfleiderer, 1997). Since the manager's effort is contractible, the problem for principal A is to choose e_A and w_A to maximize the expected payoff subject to the manager's participation constraint.

$$\max_{e_A, w_A} \quad E_{\varepsilon_A}[e_A + \varepsilon_A] - w_A \tag{3.6}$$

$$\text{s.t.} \quad w_A - C(e_A) \geq 0, \tag{3.7}$$

where we assume the manager's reservation utility is zero which holds throughout this paper.

It is clear that at the equilibrium, the participant constraint is binding and the optimal effort is

$$e_A^S = 1. \tag{3.8}$$

In this case, the principal provides the manager with a contract as

$$w_A = \frac{1}{2},$$

which equals the cost of exerting e_A^S. The expected payoff for principal A is

$$\Pi^S(A) = \frac{1}{2}. \tag{3.9}$$

Similarly, the optimal effort for principal B with standalone management is

$$e_B^S = \sigma,$$

and the expected payoff is

$$\Pi^S(B) = \frac{\sigma^2}{2}.$$

The sum of expected payoffs for the two fund principals is

$$\Pi^S(A, B) = \frac{1 + \sigma^2}{2}.$$

3.4 SBS with Complete Information

To characterize the complexity of SBS management, this section considers the simplest case where both principals can observe the manager's type δ. The reasons for focusing on complete information are at least twofold. First, SBS management is different from the traditional standalone management as "significant complexities arise even when there is no private information" in a common agency game (Bernheim and Whinston, 1986b). Second, ruling out information asymmetry permits one to isolate the effect on the outcomes of the competition between the two principals from the effect of incomplete information.

The contract of management can be formulated as a two-objective optimization problem subject to a participant constraint:

$$\max_{w_A} \quad E_{\varepsilon_A}[r_A] - w_A \tag{3.10}$$

$$\max_{w_B} \quad E_{\varepsilon_B}[r_B] - w_B \tag{3.11}$$

$$\text{s.t.} \quad w_A + w_B - C(e_A, e_B, \delta) \geq 0. \tag{3.12}$$

Comparing this problem with standalone management, the difference is that the contract is decentralized and there are two competing objectives in the contract formulation. This is also the main difference between one principal- and multiple principal-agency contracts. Subject to a joint participant constraint, the two fund principals choose incentives to maximize their own expected payoffs. The trilateral relationships make the problem much more complex. On the one hand, the two principals should provide incentives to make the manager work efficiently. On the other hand, when one principal makes an offer to the manager, she has to consider the potential conflict of interest from the other principal who may attenuate or intensify the manager's incentives. Because the efforts are substitutable, a small change in the compensation scheme offered by one fund principal affects the other fund performance. This interaction between the two funds' incentive schemes complicates the analysis of SBS management.

3.4.1 Public contracting

I now consider the public contracting where the effort exerted on one fund is observable to the other fund principal. This indicates that the two principals agree to rely on common efforts (e_A, e_B) but simultaneously and independently provide contracts to the manager. The conflict of interest is that each principal affects the manager's efforts and thus tries to mislead the other principal via the SBS manager. For technical purpose, I restrict attention to contracts which are differentiable almost everywhere.

Proposition 4 *Under complete information, the symmetric, differential equilibrium efforts are*

$$e_A^{P,fb} = \frac{1 - \sigma\delta + \theta(\sigma - \delta)}{1 - \delta^2}, \tag{3.13}$$

$$e_B^{P,fb} = \frac{\sigma - \delta + \theta(1 - \delta\sigma)}{1 - \delta^2} \tag{3.14}$$

When $\delta \leq \rho\theta$, the spillover effect dominates the substitute effect and SBS management is over-incentivized relative to standalone management, i.e., $e^{P,fb} \geq e^S$.[10] *Otherwise,*

[10]The superscript "P" and "fb" denote the *public* contracting and *first-best*, respectively. Accordingly, in the sequel, I will use "NP" and "sb" to denote *private* and *second-best* for the case of incomplete information. It should be mentioned that the first-best equilibrium refers to the case when the manager's type is known to both principals. The second-best equilibrium refers to the case when the manager's type is unknown to both fund principals.

the substitute effect becomes more important and hence SBS management is under-incentivized.

The proof of existence of equilibrium is an adaptation to Bernheim and Whinston (1986b) and I omit it here. However, it is important to point out the difference between SBS management and standalone management. First, due to complete information, the individual rationality (IR) constraint (3.12) is binding at any equilibrium. Otherwise, either principal A or B can obtain a higher payoff by decreasing w_A or w_B. Compared with the IR constraint (3.7) in standalone management, the SBS manager's compensations are from two funds.

Second, given $w_A(e)$ and $w_B(e)$, the manager chooses e to maximize his utility as

$$\max_{e} \quad w_A(e) + w_B(e) - C(e),$$

which yields the first-order conditions as

$$\frac{\partial w_A}{\partial e_A} + \frac{\partial w_B}{\partial e_A} - \rho e_A - \delta e_B = 0, \tag{3.15}$$

$$\frac{\partial w_A}{\partial e_B} + \frac{\partial w_B}{\partial e_B} - \rho e_B - \delta e_A = 0. \tag{3.16}$$

Third, for any given compensation w_B chosen by principal B, principal A chooses w_A so that the IR constraint binds and induces the manager to choose the effort level that maximizes her expected payoff. This amounts to choosing an output (e_A, e_B) which maximizes the bilateral payoff of the coalition he forms with the manager:

$$\begin{aligned} \max_{w_A} \quad & E_{\varepsilon_A}[r_A] - w_A(e_A, e_B) \\ \text{s.t.} \quad & w_A(e_A, e_B) + w_B(e_A, e_B) - C(e) \geq 0, \end{aligned}$$

where $w_B(r_B)$ is taken for granted optimal with respect to fund B.

Compared with the contract in standalone management, a new term $w_B(e_A, e_B)$ occurs in SBS management. When principal A chooses a compensation scheme, he has to consider the effect of $w_B(e_A, e_B)$ on the manager's effort choice. To solve this problem, principal A's offer will make the manager choose e to maximize

$$E_{\varepsilon_A}[r_A] + w_B(e_A, e_B) - C(e).$$

Suppose the compensation scheme $w_B(e_A, e_B)$ is piecewise differentiable. The first-order conditions for principal A is

$$\begin{aligned} \sigma + \frac{\partial w_B}{\partial e_A} - \rho e_A - \delta e_B &= 0, & (3.17) \\ \theta\sigma + \frac{\partial w_B}{\partial e_B} - \rho e_B - \delta e_A &= 0. & (3.18) \end{aligned}$$

Similarly, the problem for principal B is

$$\begin{aligned} \max_e \quad & E_{\varepsilon_B}[r_B] - w_B(e_A, e_B) \\ \text{s.t.} \quad & w_A(e_A, e_B) + w_B(e_A, e_B) - C(e) \geq 0, \end{aligned}$$

where $w_A(e_A, e_B)$ is again taken for granted optimal with respect to principal A. The first-order conditions for principal B is

$$\begin{aligned} \theta\sigma + \frac{\partial w_A}{\partial e_A} - \rho e_A - \delta e_B &= 0, \\ \sigma + \frac{\partial w_A}{\partial e_B} - \rho e_B - \delta e_A &= 0. \end{aligned}$$

Solving these six equations (3.15) - (3.19) generates (3.13).

The second part of proposition 1 shows how the two fund principals trade off between two types of contractual externalities, the positive spillover effect and the negative substitutable effect. This is reflected in the optimal efforts. On the one hand, the positive spillover effect, measured by θ, suggests the manager exerting more efforts on both funds. On the other hand, however, the negative substitute effect implies that the manager will exert less efforts. In particular, when $\delta \leq \rho\theta$, the substitute effect is relatively small, each principal increases incentives to let the manager exert more effort to her own fund relative to that in standalone management. In contrast, when $\delta > \rho\theta$ is large, the substitute effect is so strong to offset the spillover effect. As a result, each principal reduces the incentive and just enjoys the spillover performance.

The joint expected payoff between the two principals is

$$\Pi^{P,fb}(A,B) = \frac{(1+\theta)^2}{\rho+\delta}\sigma^2.$$

Proposition 5 *Under complete information, the symmetric, differential equilibrium payoffs are*

$$\left[\frac{\Pi^{P,fb}(A,B)}{2}, \frac{\Pi^{P,fb}(A,B)}{2}, 0\right].$$

Moreover, when $\delta \leq \rho(2\theta+\theta^2)$, SBS is beneficial. Otherwise, it is detrimental.

As discussed in Introduction, SBS may generate multiple equilibria. Proposition 5 presents the unique symmetric, differential equilibrium outcome such that the two principals equally distribute the total surplus. Actually, any proportional assignment between the two principals will implement an efficient outcome. This implies the challenge in SBS management that, while each principal can separately determine the incentives in w_A and w_B, they need to jointly determine the compensations satisfying the manager's participation condition.[11]

The equilibrium efforts and payoffs in Propositions 1 and 2 are the same with the case when the two principals are merged to one and SBS reduces to the traditional one principal one agent with two tasks contract. The main reason is that there is

[11] Martimort and Stole (2003) exclusively discuss the multiplicity of equilibria in common agency with complete information and pure strategy.

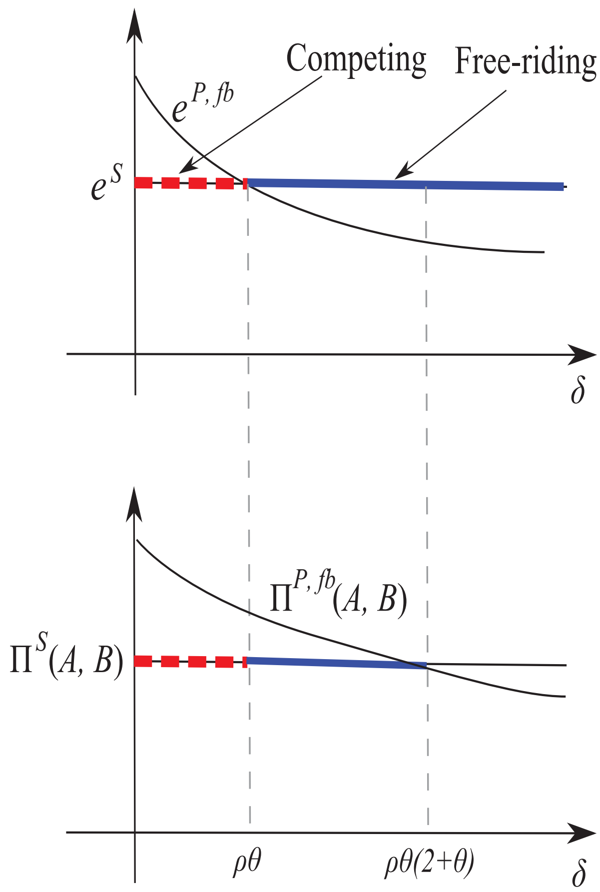

Figure 3.2: Comparison of efforts and payoffs with SBS and standalone management.

information asymmetric and the two compensations are public to all players. In this case, the goal of the two principals is to offer compensation w_A and w_B to solve the following problem:

$$\begin{aligned} \max_{w_A, w_B} \quad & E_{\varepsilon_A}[r_A] + E_{\varepsilon_B}[r_B] - w_A - w_B \\ \text{s.t.} \quad & w_A + w_B - C(e_A, e_B, \delta) \geq 0. \end{aligned}$$

Propositions 2 shows how the two externalities affect the two principals' equilibrium payoffs. In particular, when $\delta \leq \rho\theta(2+\theta)$, SBS management is better off due to the spillover effect. Figure 3.2 show how the equilibrium efforts and payoffs change over the manager's type δ. When $\delta < \rho\theta$, the spillover effect is stronger and the two principals choose to compete by increasing incentives. As a results, the equilibrium efforts are larger than that in standalone management. When $\delta > \rho\theta$, the two principals ride on the spillover performance by reducing incentives. Hence the managers exerts less efforts. The lower panel of figure 3.2 even the two principals ride on the other's contract, the joint performance still could be better once $\delta < \rho\theta(2+\theta)$. This is because the spillover effect can intensify with each other. The better performance of fund A generates a spillover effect on fund B. In turn, the increased performance in fund B delivers a reverse spillover performance on fund A.

One can intuitively relate the substitute parameter δ to the manager's ability. As a result, I have the following implication.

Implication 1 *Fund managers with lower abilities are more likely to be confined in managing one fund; on the other hand, managers with higher ability tend to do SBS management.*

This prediction is consistent with the work of Nohel, Wang and Zheng (2010) and Agarwal and Ma (2012) who document that SBS is more likely to be delegated to well-performing and skilled managers. Particularly, Nohel, Wang and Zheng (2010) find that the outperformance on the mutual fund side is driven by those who began their careers as mutual fund managers, which implies that the experience or skill is a key factor for SBS delegation. Agarwal and Ma (2012) show that a one-standard-deviation increase in the four-factor fund alpha can lead to an increase of 12.5% in the probability of managers switching to SBS management from standalone management.

Implication 2 *SBS managers is more likely to work harder than standalone managers.*

This is a new prediction unexplored but interesting in the SBS literature. The intuition is that competition between the principals raises effort above the level of standalone management. Since the effort exerted on one fund has a positive externality on the other fund, each principal pays at the margin too much for the effort from the manager. In any equilibrium, a principal does not choose effort below the standalone management level because of the threat that the manager will work harder for the other fund. A less degree of competition is obtained when the manager has little incentive to substitute one effort against the other. This occurs when both principals offer the same compensation around the equilibrium. Finally, the spillover effect has a second-order effect on the fund principals. This suggests that the spillover effect intensifies with each other. For example, when the effort e_A is increased, the performance on fund A will improves, generating a positive spillover effect on fund B. In turn, the better performance on fund B will indirectly generates a second-order spillover effect on fund A.

Before ending this section, some one may be interested in the optimal compensation of SBS managers. With the symmetric, differentiable assumption, the optimal compensation is hence

$$w_A^{P,fb} = w_B^{P,fb} = \frac{1}{2}C(e_A^{P,fb}, e_B^{P,fb}, \delta) = \frac{(1+\theta)^2}{2(\rho+\delta)}\sigma^2,$$

where I use the fact that the manager has zero expected payoff at the equilibrium according to proposition 2.

3.4.2 Private contracting

When the two funds are from two different fund families, the two principals propose contracts noncooperatively and each principal can only contract with the manager on effort e_i, i.e., $w_i = w_i(e_i)$.[12] The reason for this noncooperation may be that principal i does not have the auditing rights or monitoring technologies to observe

[12] He and Xiong (2012) give an example that principal cannot observe the effort exerted on another market even the contract is centralized due to market segmentation.

e_j exerted on fund j. This will result in some interesting results. First, there are multiple symmetric, differential equilibria. Second, free-riding will becomes central. Third, the equilibrium payoff of SBS management still could be larger than that in standalone management.

Since principal A pays only for effort e_A, the compensation from fund A is a function of effort e_A regardless of e_B. This is, given the compensation from fund B, principal A privately chooses the pair (e_A, w_A) to maximize her expected payoff,

$$\begin{aligned} \max_{w_A, e_A} \quad & \sigma(e_A + \theta e_B) - w_A(e_A) \\ \text{s.t.} \quad & w_A(e_A) + w_B(e_B) - C(e_A, e_B, \delta) \geq 0. \end{aligned}$$

The difference between private and public contracting is that principal A cannot directly contract on the effort e_B. However, this does not mean that the compensation w_A has no effect on e_B. Actually, an indirect externality still exists. Principal A can vary the incentive w_A to change the marginal cost on fund B so that indirectly affects w_B.

Given fund B's equilibrium contract, define the indirect utility function of principal A as

$$v(e_A) = \max_{\tilde{e}_B} \ \sigma(e_A + \theta \tilde{e}_B) + w_B(\tilde{e}_B) - C(e_A, \tilde{e}_B, \delta),$$

where $w_B(e_B)$ satisfies the first-order condition of the manager

$$\frac{\partial w_B}{\partial e_B} = C_2(e_A, e_B, \delta) = \rho e_B + \delta e_A,$$

which characterizes the manager's choice of e_B, implying

$$\frac{\partial e_B}{\partial e_A} = \frac{\delta}{w_B'' - \rho}. \tag{3.19}$$

The first-order condition with respect to principal A is

$$\sigma\left[1+\theta\frac{\partial e_B}{\partial e_A}\right]+w'_B\frac{\partial e_B}{\partial e_A}=(\rho e_A+\delta e_B)+\left[\rho e_B+\delta e_A\right]\frac{\partial e_B}{\partial e_A},$$

i.e.,

$$\sigma\left[1+\theta\frac{\partial e_B}{\partial e_A}\right]=\rho e_A+\delta e_B.$$

Together with (3.19), I have

$$\sigma\left[1+\frac{\theta\delta}{w''_B-\rho}\right]=\rho e_A+\delta e_B. \tag{3.20}$$

The necessary conditions for a symmetric, differential equilibrium are that the Hessian of the symmetric problem calculated at the equilibrium is semi-negative definite:

$$\begin{aligned} w''_B-\rho &\leq 0, && (3.21)\\ (w''_B-\rho)^2 &\geq \delta^2. && (3.22)\end{aligned}$$

I use these two necessary local concavity conditions to derive the boundaries of the equilibrium sets.

Substituting (3.21) into (3.20) gives

$$(\rho+\delta)e\leq\sigma, \tag{3.23}$$

which implies that the equilibrium efforts $(e_A^{NP,fb}, e_B^{NP,fb})$ have an upper bound such as

$$e_A^{NP,fb}=e_B^{NP,fb}\leq\frac{\sigma}{\rho+\delta}.$$

On the other hand, from (3.22), one has

$$e_A^{NP,fb} = e_B^{NP,fb} \geq \frac{1-\theta}{\rho+\delta}\sigma.$$

Now I am ready to present the following proposition.

Proposition 6 *Under complete information, any effort* $e^{NP,fb} \in [\frac{1-\theta}{\rho+\delta}\sigma, \frac{1}{\rho+\delta}\sigma]$ *can be a symmetric, differentiable equilibrium. Moreover, in all such equilibria, the agent gets zero rent:*

$$U(\delta) = w_A + w_B - C(e_A, e_B, \delta) = 0.$$

The proof is a special adaption of proposition 1 in Martimort and Stole (2003), and the requirements for the existence of equilibrium are apparently satisfied.

By using a nonlinear differential compensation, fund A restricts not only the manager's equilibrium effort e_A but also the behavior of the manager around this equilibrium. This extra control of the manager's behavior off-the-equilibrium path changes the degree of the fund principals' competition. Choices offered by one principal that are not taken in equilibrium constrains the other principal from inducing the manager to exert a different effort.

Since the SBS manager can always substitute away effort for fund A against effort for fund B, each principal will not pay at the margin as much as that in standalone management. Instead, they ride on the effort bought by the other principal which can generate a positive spillover effect. In equilibrium, the manager thus decreases the effort for each principal with respect to a situation where the principals would have competed. In the extreme case, the maximum degree of free-riding is obtained when the efforts are perfect substitutable. This responds to the equilibrium ($e^{NP,fb} = \frac{(1-\theta)\sigma}{\rho+\delta}$) when both principals offer a flattest nonlinear compensations around the equilibrium. On the other hand, the maximal degree of competition is obtained when each principal provides the steepest contract that is independent of the other one. In this case, no principal can increase incentive anymore ($e^{NP,fb} = \frac{\sigma}{\rho+\delta}$). By varying the slope of the out-of-equilibrium transfer-effort pairs, any effort between the extremes can be implemented.

From proposition (6), in any equilibrium, each fund principal does not choose effort level larger than that in stanalone management except $\delta = 0$. A fair question is whether this kind of free-riding equilibra are detrimental or not. To answer this question, let us calculate the joint payoff between the two principals

$$\Pi^{NP,fb}(A,B) = 2\sigma(1+\theta)e^{NP,fb} - (\rho+\delta)(e^{NP,fb})^2,$$

which is increasing over $[\frac{1-\theta}{\rho+\delta}\sigma, \frac{1}{\rho+\delta}\sigma]$. Hence,

$$\frac{1+2\theta-\theta^2}{\rho+\delta}\sigma^2 \leq \Pi^{NP,fb}(A,B) \leq \frac{1+2\theta}{\rho+\delta}\sigma^2.$$

Proposition 7 *Under complete information, SBS management with private contracting is better off relative to that of standalone management when $\delta \leq \rho\theta(2-\theta)$.*

The proof is simple since when $\delta \leq \rho\theta(2-\theta)$, $\frac{1+2\theta-\theta^2}{\rho+\delta}\sigma^2 \geq \frac{\rho^2}{\rho}$, i.e, the joint payoff with SBS is larger than that in standalone management. Figure 3.3 shows the difference of the equilibrium efforts and payoff between the public and the private contracting. With public contracting, the equilibrium is unique, and $\delta_1 = \rho\theta$ and $\delta_2 = \rho\theta(2+\theta)$ are the thresholds when the equilibrium efforts and payoffs are larger than that in standalone management. When the contracting is private, however, there are infinite equilibriums. Figure 3.3 shows the minimum equilibrium efforts and payoffs. Interestingly, the equilibrium efforts are always less than that in standalone management. This is, due to unobservable contract, each principal provides less incentive to the manager but rides on the other principal's contracting to enjoy the spillover effect. This is reflected in the lower panel of figure 3.3. When $\delta < \delta_3 = \rho\theta(2-\theta)$, the joint payoff with SBS management is larger than that in standalone management.

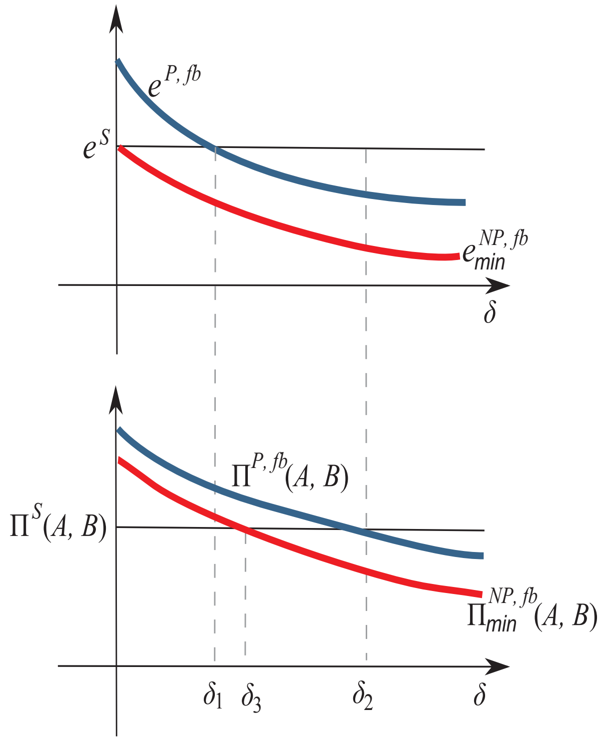

Figure 3.3: Comparison of efforts and payoffs with public and private contracting. $\delta_1 = \rho\theta$, $\delta_2 = \rho\theta(2+\theta)$, and $\delta_3 = \rho\theta(2-\theta)$. $e_{min}^{NP,fb}$ is the minimum equilibrium efforts and $\Pi_{min}^{NP,fb}$ is the minimum equilibrium payoff in SBS management.

Therefore, when the substitute effect is weak, the spillover effect is more important in SBS management.

Someone may wonder that if one can use direct mechanism by restricting attention to singleton contracts. However, singleton contracts cannot include compensations that will not be implemented in equilibrium.

Corollary 1 *When principals are restricted to singleton contracts of the form* $\{w_i, e_i\}$, $e_A^{NP,fb} = e_B^{NP,fb} = \frac{\sigma}{\rho+\delta}$ *is the unique equilibrium of SBS with complete information.*

The proof is similar to proposition 2 of Martimort and Stole (2003). This equilibrium corresponds to the least degree of free-riding. Assume that principal B offers the direct revelation mechanism $\{w_B(\delta), e_B(\delta)\}$, then principal A's problem is

$$\begin{aligned} \max_{w_A, e_A} \quad & \sigma(e_A + \theta e_B) - w_A \\ \text{s.t.} \quad & w_A + w_B(\delta) - C(e_A, e_B, \delta) \geq 0. \end{aligned}$$

The first-order condition yields

$$\sigma = \rho e_A + \delta e_B.$$

In a symmetric equilibrium, I have

$$e_A^{NP,fb} = e_B^{NP,fb} = e^{NP,fb} = \frac{\sigma}{\rho + \delta}.$$

The joint payoff for the two principals at the equilibrium is

$$\Pi^{NP,fb} = 2(1+\theta)\sigma e^{NP,fb} - C(e^{NP,fb}, e^{NP,fb}, \delta) = \frac{1+2\theta}{\rho+\delta}\sigma^2.$$

Corollary 2 *Suppose the direct mechanism is implemented, i.e., the equilibrium efforts are* $e_A^{NP,fb} = e_B^{NP,fb} = \frac{1}{\rho+\delta}\sigma$. *When* $\delta \leq 2\rho\theta$, *SBS is preferred.*

This may be the most counterintuitive result in this paper. Even both principals freely ride on each other's contracts by providing a flatter incentive to the manager, the expected payoff could still be better than standalone management due to the spillover effect. The reason is that principal A freely enjoys the benefit from the performance spillover but does not pay anything for effort e_B bought by principal B, and vice versa.

3.5 Incomplete Information

This section studies a more realistic situation where the manager's type δ is unobservable and noncontractable, but it is common knowledge that δ follows a uniform distribution with distribution function $F(\delta)$ and density $f(\delta)$ over $[0, \bar{\delta}]$. In this framework, I assume that each fund principal chooses a direct compensation scheme consisting of a pair of functions specifying, for any reported type, the compensation to the manager and the effort level in the principal's own fund. It is important to stress that the manager must send a separate report to each fund principal. Although the manager's reports to the two principals could differ, in equilibrium they coincide with the manager's true type (Martimort and Stole, 2009). I assume that the two principals use continuous, piecewise differentiable compensation functions.

3.5.1 Public contracting

I now analyze the case when the two fund principals contract on both e_A and e_B. As in the full information case, principal A's problem is

$$\max_{e_A, e_B, w_A} \quad \int_0^{\bar{\delta}} [\sigma(e_A + \theta e_B) - w_A] f(\delta) d\delta \tag{3.24}$$

$$\text{s.t.} \quad w_A + w_B - C(e_A, e_B, \delta) \geq 0, \tag{3.25}$$

$$(e_A, e_B) \in \text{argmax}_{\tilde{e}_A, \tilde{e}_B} w_A + w_B - C(\tilde{e}_A, \tilde{e}_B, \delta). \tag{3.26}$$

The difference between complete and incomplete information is the incentive compatibility constraint, which says that the manager must obtain a higher compensation from exerting more efforts and therefore, as the equilibrium, the efforts exerted by the SBS manager are optimal with respect to his type δ.

Define the manager's total compensation as

$$U(\delta) = \max_{e_A, e_B} \quad w_A(e_A, e_B) + w_B(e_A, e_B) - C(e_A, e_B, \delta).$$

With envelop theory, the first-order derivative with respect to the type is

$$\dot{U}(\delta) = -C_\delta(e_A, e_B, \delta) = -e_A e_B \leq 0.$$

Let $\bar{\delta}$ be the least skilled SBS manager to whom the two principals will delegate their funds, i.e., $U(\bar{\delta}) = 0$. As in centralized contracting, principal A's objective function (3.24) can be rewritten as

$$\begin{aligned} \Pi(A) &= \int_0^{\bar{\delta}} [\sigma_A(e_A + \theta e_B) - w_A] f(\delta) d\delta \qquad (3.27) \\ &= \int_0^{\bar{\delta}} \left[\sigma_A(e_A + \theta e_B) + w_B - C(e_A, e_B, \delta) - \frac{F(\delta)}{f(\delta)} e_A e_B \right] f(\delta) d\delta. \end{aligned}$$

Define

$$\delta_P = \delta + 2\frac{F(\delta)}{f(\delta)} = 3\delta.$$

With mild assumption that ensures the existence of optimal effort, the first-order condition for principal A is

$$\begin{aligned} \sigma + \frac{\partial w_B}{\partial e_A} - \rho e_A - \delta_C e_B &= 0, \\ \theta\sigma + \frac{\partial w_B}{\partial e_B} - \rho e_B - \delta_C e_A &= 0. \end{aligned}$$

Similarly, the first-order conditions for principal B is

$$\begin{aligned} \theta\sigma + \frac{\partial w_A}{\partial e_A} - \rho e_A - \delta_C e_B &= 0, \\ \sigma + \frac{\partial w_A}{\partial e_B} - \rho e_B - \delta_C e_A &= 0, \end{aligned}$$

and for the manager is

$$\begin{aligned}\frac{\partial w_A}{\partial e_A} + \frac{\partial w_B}{\partial e_A} - \rho e_A - \delta e_B &= 0,\\ \frac{\partial w_A}{\partial e_B} + \frac{\partial w_B}{\partial e_B} - \rho e_B - \delta e_A &= 0.\end{aligned}$$

Solving the above six equations gives the following proposition.

Proposition 8 *Under incomplete information and public contracting, the symmetric, differential equilibrium efforts are given by*

$$e_A^{P,sb} = e_B^{P,sb} = \frac{1+\theta}{\rho+\delta_P}\sigma = \frac{1+\theta}{\rho+3\delta}\sigma, \tag{3.28}$$

When $\delta \le \frac{1}{3}\rho\theta$, the SBS manager is over incentivized relative to the standalone manager. Otherwise, he is under incentivized.

Comparing with the equilibrium effort under complete information, each principal reduces the manager's efforts under incomplete information to better extract his rent. The reason is that each principal affords the full cost of information disclosure but only enjoys a part of its benefit.

Now I calculate the rent of the SBS manager with type δ as

$$U(\delta) = \int_\delta^{\bar\delta} e_A e_B d\delta = \frac{(\bar\delta-\delta)(1+\theta)^2}{(\rho+3\bar\delta)(\rho+3\delta)}\sigma^2,$$

and the joint expected payoff of the two fund principals as

$$\begin{aligned}\Pi^{P,sb}(A,B) &= \int_0^{\bar\delta}\Big[\sigma_A(e_A+\theta e_B)+\sigma_B(e_B+\theta e_A) - C(e_A,e_B,\delta) - U(\delta)\Big] f(\delta)\delta\\ &= \int_0^{\bar\delta}\Big[2(1+\theta)\sigma e - (\rho+\delta)e^2 - U(\delta)\Big] f(\delta)d\delta\\ &= \sigma^2(1+\theta)^2\left[\frac{4}{9\bar\delta}[\log(\rho+3\bar\delta)-\log(\rho)] - \frac{1}{3(\rho+3\bar\delta)}\right].\end{aligned}$$

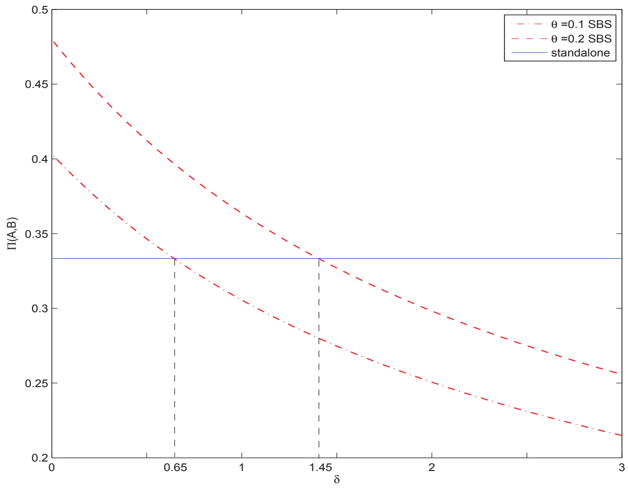

Figure 3.4: Joint payoffs of SBS and standalone management

Define δ_U to be the solution to $\Pi^{P,sb}(A,B) = \frac{\sigma^2}{\rho}$.

Proposition 9 *When $\bar{\delta} \leq \delta_U$, SBS is beneficial.*

While it is difficult to give an analytical formula for δ_U, one can prove that the joint expected payoff of the two principals is monotonically decreasing with respect to $\bar{\delta}$. This suggests that the variation of the SBS manager's type is important in SBS management. When the spillover effect is fixed, the larger the substitute effect, the lower the joint payoff. Suppose $\rho = 3$ and $\sigma = 1$. Figure 3.4 shows how the joint payoff changes over the upper bound of δ. When $\theta = 0.1$, SBS management is better if $\bar{\delta} \leq 0.65$. When $\theta = 0.2$, SBS management is preferred if $\bar{\delta} \leq 1.45$.

3.5.2 Private contracting

Now I consider the non-cooperative case where principal A only contracts on effort e_A and principal B only contracts on effort e_B as in the complete information case.

Given principal B's contract that satisfies the manager's first-order condition,

$$w'_B(e_B) = C_2(e_A, e_B, \delta) \tag{3.29}$$

Principal A's objective function can be written as

$$\begin{aligned}
\max_{U(\delta),e_A} \quad & \int_0^{\bar{\delta}} \Big[\sigma(e_A+\theta e_B)+w_B-C(e_A,e_B)-U(\delta)\Big] f(\delta)d\delta \\
\text{s.t.} \quad & \dot{U}(\delta) = -e_A e_B, \\
& U(\delta) \geq 0, \\
& \dot{e}_A(\delta) \quad \text{is non-increasing.}
\end{aligned}$$

Proceeding as the complete information case, principal A's best response satisfies

$$\sigma_A(1+\theta\frac{\partial e_B}{\partial e_A}) - (\rho e_A + \delta e_B) - \frac{F(\delta)}{f(\delta)}[e_B + e_A\frac{\partial e_B}{\partial e_A}] = 0. \tag{3.30}$$

Differentiating (3.29) with respect to e_A gives

$$\frac{\partial e_B}{\partial e_A} = \frac{\delta}{w''_B - \rho}.$$

Hence, the differential equilibrium equation (3.30) can be rewritten as

$$\sigma_A - (\rho e_A + \delta e_B) - \frac{F(\delta)}{f(\delta)}e_B + [\theta\sigma_A - \frac{F(\delta)}{f(\delta)}e_A]\frac{\delta}{w''_B - \rho} = 0.$$

Symmetric equilibrium implies that

$$\sigma - (\rho + \delta + \frac{F(\delta)}{f(\delta)})e + (\theta\sigma - \frac{F(\delta)}{f(\delta)}e)\frac{\delta}{w'' - \rho} = 0.$$

Proposition 10 *Under incomplete information, any symmetric, differentiable equilibrium with SBS are such that*

$$\frac{\sigma}{\rho + 2\delta} < e^{NP,sb} < \frac{\sigma}{\rho + \delta}$$

over $(0, \bar{\delta}]$. *When* $\delta = 0$, $e = \frac{\sigma}{\rho}$. *The manager with type* δ *has the information rent*

$$U(\delta) = \int_{\delta}^{\bar{\delta}} e^2 dz$$

The proof is similar to proposition 6 of Martimort and Stole (2003).

In the following, I calculate the range of the joint expected payoff. If $e = \frac{\sigma}{\rho+2\delta}$,

$$U(\delta) = \frac{(\bar{\delta} - \delta)\sigma^2}{(\rho + 2\delta)(\rho + 2\bar{\delta})}.$$

So

$$\begin{aligned} \Pi^{NP,sb}(A, B) &> \int_0^{\bar{\delta}} \Big[2(1+\theta)\sigma e - C(e, e, \delta) - U(\delta) \Big] f(\delta) d\delta \\ &= \sigma^2 (1 + 2\theta) \frac{\log(\rho + 2\bar{\delta}) - \log(\rho)}{2\bar{\delta}}. \end{aligned} \tag{3.31}$$

On the other hand, if $e = \frac{\sigma}{\rho+\delta}$,

$$U(\delta) = \frac{(\bar{\delta} - \delta)\sigma^2}{(\rho + \bar{\delta})(\rho + \delta)}.$$

Hence,

$$\begin{aligned} \Pi^{NP,sb}(A, B) &< \int_0^{\bar{\delta}} \Big[2(1+\theta)\sigma e - C(e, e, \delta) - U(\delta) \Big] f(\delta) d\delta \\ &= \sigma^2 \left[\frac{2\theta[\log(\rho + \bar{\delta}) - \log(\rho)]}{\bar{\delta}} + \frac{1}{\rho + \bar{\delta}} \right]. \end{aligned} \tag{3.32}$$

Define δ^U to be the solution to

$$\frac{(1 + 2\theta)[\log(\rho + 2\bar{\delta}) - \log(\rho)]}{2\bar{\delta}} = \frac{1}{\rho}. \tag{3.33}$$

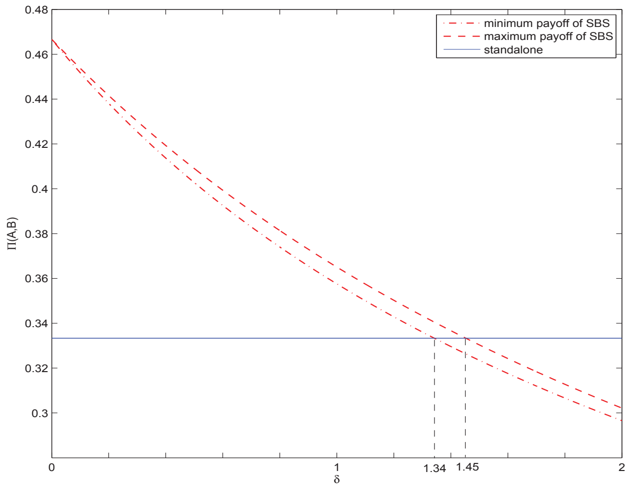

Figure 3.5: Joint payoffs of SBS and standalone management

Proposition 11 *Under incomplete information, SBS with non-cooperative contracting is beneficial when* $\bar{\delta} \leq \delta^U$.

The proof is simple since both (3.31) and (3.32) are decreasing with respect to $\bar{\delta}$. This proposition says that when $\bar{\delta} < \delta^U$, the worst-case equilibrium payoff with SBS management is larger than that in standalone management. Figure 3.5 shows how the equilibrium payoffs change in terms of the upper bound δ at which the two principals are indifferent between SBS management and standalone management. Since there are infinite equilibra, this figure shows the minimum and maximum payoffs for a specific $\bar{\delta}$. Suppose $\rho = 3$, $\theta = 0.2$ and $\sigma = 1$. When $\bar{\delta} < 1.34$, the payoff with the worst-case equilibrium ($e_A^{NP,sb} = e_B^{NP,sb} = \frac{\sigma}{\rho+2\delta}$) is larger than that in standalone management.

Implication 3 *SBS is more likely to happen in fund families.*

The intuition of this implication is that, if the two funds from one fund family, the principals are more likely to cooperate with each other. On the other hand, to understand the basic externality across funds that leads to a private contracting outcome different from the public contracting, one needs to describe how a change in one compensation affects the choice of the other one. Under asymmetric information, there is a tradeoff between increasing the marginal efficiency of the effort allocation and reducing the informational rent left to the manager. If fund A distorts its compensation further downward, this makes the manager choose a higher effort to fund B. Since fund A's compensation is distorted downward, a marginal increase of incentive in fund B is beneficial for fund B. On the other hand, under private contracting, principal A cannot observe the effort exerted on fund B, and hence she does not pay it but can enjoy the spillover effect. With this direct positive externality, each principal will ride on the other principal's contract until the maximal free-riding equilibrium is achieved.

With my theoretical results, one can easily to explain the conflicting, empirical evidence of SBS management. Recall that in Cici, Gibson, and Moussawi (2010), they define SBS managers to be *firms* who simultaneously manage mutual funds and hedge funds, and find that SBS mutual fund underperform unaffiliated funds. This is more likely to correspond to the private contracting case for SBS management. Regarding a firm/fund family as an SBS manager, investors invested in one fund do not even know those investors who invest in the other fund managed by the same firm/family. In this case, it is impossible for one principal to observe the effort exerted on the other fund. As a result, the contracts are private, and the final fund performance may be worse due to more distortions. Instead, in Nohel, Wang and Zheng (2010), they define an SBS manager as an identified person who simultaneously manage two funds. With this definition, they find that SBS mutual funds outperform unaffiliated funds. In this case, SBS funds are more likely to belong to one fund family.

Implication 3 can be directly used to test outsourcing mutual fund management. It is popular for mutual fund families to offer individual funds with various investment styles and fee structures, which can increase the competitive advantage and result in more stable inflows since investors are more likely to switch to funds inside a family when they want to access different styles. However, mutual fund families face challenges since they may lack the expertise in hiring and evaluating managers in styles other than their core ones. To reduce management costs, most of fund families outsource some funds to external advisors who help manage the funds and obtain the management fees. This suggests that investors may suffer from conflicts of interest faced by their fund advisors. Chen, Hong, Jiang and Kubik (2012) find that outsourcing funds underperform in-house funds due to firm boundaries make it difficult to extract performance from an outsourced relationship. However, they do not investigate the performance difference between outsourced funds since an outside manager can manage funds from different fund families, thereby the manager may favor funds from one family over those from another family.

3.6 Extension

So far, I have focused on the case where the SBS manager does not have any ownership in the funds he manages. With this implicit assumption, the decentralized compensation schemes distort the optimal efforts in the case of information asymmetry, especially when the two principals are competing with each other. In this section, I consider an extension for future research that may relieve the distortions of SBS.

3.6.1 Sequential SBS Management

Let us now consider a sequential timing SBS. That is, the two fund principals contract with the manager sequentially. Implicit under simultaneous common agency is that principals obtain an expected fund performance better than their reservation/standalone performance. Principals choose to participate because they earn more than standalone management. This imposes some conditions on the way they can share contributions of SBS manager.

With sequential SBS, the difficulty could be that Fund A offers compensations to the manager which are so low that inducing the manager's participation may become too costly for Fund B. To avoid such issues, Fund B should have the option of choosing standalone management, breaking down the SBS management. This might be captured by introducing a constraint stipulating that Fund A designs his own mechanism taking into account that Fund B should chooses SBS too.[13]

Given $w_A(r_A)$ offered by Fund A, optimal compensation between Fund B and the manager leads to choose an indirect utility of Fund B which maximizes

$$V_B(\delta) = \max_e \ \sigma_B(e_B + \theta e_A) + w_A - \frac{\rho}{2}(e_A^2 + e_B^2) - \delta e_A e_B - \frac{F(\delta)}{f(\delta)} e_A e_B \tag{3.34}$$

With the Envelope theorem, we have

$$V_B'(\delta) = -\left(1 + \frac{d}{d\delta}\left(\frac{F(\delta)}{f(\delta)}\right)\right) e_A e_B. \tag{3.35}$$

Using revealed preferences arguments and (3.34) yields the monotonicity condition

$$e_A'(\delta) \leq 0, \quad e_B'(\delta) \leq 0.$$

Because Fund B may not participate SBS, he could decide to give up the SBS with type δ if the expected utility of having that type manager is worse. To avoid this situation, the following constraint must hold

$$V_B(\delta) \geq \frac{\sigma_B^2}{2\rho} = \Pi^{S,fb}(B), \tag{3.36}$$

which is the expected utility with standalone management.

[13]Here is an example. Suppose the compensation is linear, $w_A = a_A + b_A r_A$ and $w_B = a_B + b_B r_B$. Then at the equilibrium, we can specify the optimal b_A, b_B, and $a_A + a_B$, the specific a_A and a_B can not be specified. That is also a feature with common agency that has multiple equilibrium outcomes or multiple equilibra.

Acting as a Stackelberg leader, principal A anticipates this continuation of the contracting game and solves

$$\max_{e,V_2} \int_0^{\bar{\delta}} \left[\sigma_A(e_A+\theta e_B)+\sigma_B(e_B+\theta e_A)-\frac{\rho}{2}(e_A^2+e_B^2)-[\delta+\frac{F(\delta)}{f(\delta)}]e_A e_B - V_2(\delta)\right] f(\delta) d\delta$$

subject to (3.35) and (3.36).

From (3.35) and the fact that (3.36) binds at $\bar{\delta}$ only, we get

$$V_2(\delta)=\int_\delta^{\bar{\delta}}\left(1+\frac{d}{d\delta}\left(\frac{F(\delta)}{f(\delta)}\right)e_A e_B\right)d\delta+\frac{\sigma_B^2}{2\rho},$$

which is decreasing if $\frac{d}{d\delta}\left(\frac{F(\delta)}{f(\delta)}\right)\geq 0$ and $\frac{d^2}{d\delta^2}\left(\frac{F(\delta)}{f(\delta)}\right)\geq 0$.

Then the above optimization problem can be rewritten as

$$\max_{e,V_2} \int_0^{\bar{\delta}} \left[\sigma_A(e_A+\theta e_B)+\sigma_B(e_B+\theta e_A)-\frac{\rho}{2}(e_A^2+e_B^2)-\left[\delta+\frac{F(\delta)}{f(\delta)}\left(2+\frac{d}{d\delta}\left(\frac{F(\delta)}{f(\delta)}\right)\right)\right]e_A e_B\right] f(\delta) d\delta$$

Define

$$\delta_S=\delta+\frac{F(\delta)}{f(\delta)}\left(2+\frac{d}{d\delta}(\frac{F(\delta)}{f(\delta)})\right)$$

$$e_A^{Sq} = \frac{\rho(\sigma_A+\theta\sigma_B)-\delta_S(\sigma_B+\theta\sigma_A)}{\rho^2-\delta_S^2}, \quad (3.37)$$

$$e_B^{Sq} = \frac{\rho(\sigma_B+\theta\sigma_A)-\delta_S(\sigma_A+\theta\sigma_B)}{\rho^2-\delta_S^2}. \quad (3.38)$$

Now we can obtain some flexible results. When $\frac{d}{d\delta}\left(\frac{F(\delta)}{f(\delta)}\right)\geq 1$, the distortion is larger than that of SBS with non-cooperation. When $0\leq\frac{d}{d\delta}\left(\frac{F(\delta)}{f(\delta)}\right)<1$, the distortion is larger than that of SBS with cooperation. Otherwise, the distortion is less than the case of SBS with cooperation.

3.6.2 SBS with ownership

Khorana et al. (2007) find that mutual fund performance improves by about three basis points for each basis point of managerial ownership. This motivates the consideration of the case when the manager has a proportion of α ownership in both funds.[14] I only consider the case of cooperation contracting. In this case, principal A's objective function is

$$\max_{e} \int_{0}^{\bar{\delta}} \left[\sigma(e_A + \theta e_B) + w_B - C(e_A, e_B, \delta) - (1-\alpha)\frac{F(\delta)}{f(\delta)} e_A e_B \right] f(\delta) d\delta.$$

Define

$$\delta_O = \delta + 2(1-\alpha)\frac{F(\delta)}{f(\delta)} = (3-2\alpha)\delta,$$

which implies the optimal efforts as

$$e_A^O = e_B^O = \frac{(1+\theta)\sigma}{\rho + \delta_O} = \frac{(1+\theta)\sigma}{\rho + (3-2\alpha)\delta}. \tag{3.39}$$

Obviously, ownership can reduce the effort distortion with information asymmetry, and as a result, it may then be nearly optimal relative to the case with full information.

3.7 Conclusion

In this paper, I present a common agency model to explore the benefits and costs of SBS when one manager is simultaneously managing multiple funds with different incentives. In contrast to the traditional centralized contracting, the SBS contract is decentralized and explicitly considers the competition between the two fund principals. This explains the empirical puzzle why the SBS performance is mixed.

[14] Recall that the objective here is to explore the effect of ownership on the optimal efforts and the case of different ownerships in the two funds is beyond the scope of this paper.

3.8 References

Admati, A.R., and P. Pfleiderer. 1997. Does it all add up? Benachmarks and the compensation of active portfolio managers. *Journal of Business* 70: 323–350.

Agarwal, V., and L. Ma. 2012. Managerial multitasking in the Mutual fund industry. Working paper, Georgia State University.

Agarwal, V., N.M. Boyson, and N.Y. Naik. 2009. Hedge funds for retail investors? An examination of hedged mutual funds. *Journal of Financial and Quantitative Analysis* 44: 273–305.

Bernheim, D.B., and M.D. Whinston. 1986a. Common agency. *Econometrica* 54: 923–942.

Bernheim, D.B., and M.D. Whinston. 1986b. Menu auctions, resource allocation, and economic influence. *Quarterly Journal of Economics* 101: 1–32.

Bhattacharya, S., and P. Pfleiderer. 1985. Delegated portfolio management. *Journal of Economic Theory* 36: 1–25.

Brown, D.P., and Y. Wu. 2011. Mutual fund families and performance evaluation. Working Paper, University of Wisconsin-Madison.

Chen, J., H. Hong, and J.D. Kubik. 2012. Outsourcing mutual fund management: firm boundaries, incentives, and performance. *Journal of Finance*, forthcoming.

Cici, G., S. Gibson, and R. Moussawi. 2010. Mutual fund perfroamnce when parent firms simultaneously manage hedge fund. *Journal of Financial Intermediation* 19: 169–187.

Dangl, T., Y. Wu, and J. Zechner. 2008. Market discipline and internal governance in the mutual fund industry. *Review of Financial Studies* 21: 2307–2343.

Dixit, A. 1997. Power of incentives in private versus public organizations. *American Economic Review: Papers and Proceedings* 82: 378–383.

Elton, E.J., M.J. Gruber, and T.C. Green. 2007. The impact of mutual fund family membership on investor risk. *Journal of Financial and Quantitative Analysis* 42: 257–278.

Gaspar, J.M., M. Massa, and P. Matos. 2006. Favoritism in mutual fund families? evidence on strategic cross-fund subsidization. *Journal of Finance* 61: 73–104.

Gavazza, A. 2011. Demand spillovers and market outcomes in the mutual fund industry. *RAND Journal of Economics*, forthcoming.

He, Z., and W. Xiong. 2012. Delegated asset management and investment mandates. *Journal of Financial Economics*, forthcoming.

Holmstrom, B., and P. Milgrom. 1991. Multi-task principal-agent analyses: linear contracts, asset ownership and job design. *Journal of Law, Economics and Organization* 7: 24–52.

Laffont, J.J., and J. Tirole. 1986. Using cost observation to regulate firms. *Journal of Political Economy* 94: 614–641.

Liang, P.J. and L. Nan. 2011. Endogenous precision of performance measures and limited managerial attension. Working Paper, Carnegie Mellon University.

Khorana, A., H. Servaes, and L. Wedge. 2007. Portfolio manager ownership and fund performance, *Journal of Financial Economics* 85: 179–204.

Kolokolova, O. 2011. Strategic behavior within families of hedge funds. it Journal of Banking & Finance 35: 1645–1662.

Maier, N. and M. Ottaviani. 2009. Information sharing in common agency: when is transparency good? *Journal of the European Economic Association* 7: 162–187.

Martimort, D. 1996. Exclusive dealing, common agency, and multiprincipals incentive theory. *RAND Journal of Economics* 27: 1–31.

Martimort, D. 2007. Multi-contracting mechanism design. *Advances in Economics and Econometrics*, Chapter 2, Cambridge University Press.

Martimort, D., and L.A. Stole. 2003. Contractual externalities and common agency equilibria. Advances in Theoretical Economics 3: Article 4.

Martimort, D., and L.A. Stole. 2009. Market participation in delegated and intrinsic common-agency games. *RAND Journal of Economics* 40: 78–102.

Massa, M. 2003. How do family strategies affect fund performance? When performance-maximization is not the only game in town. *Journal of Financial Economics* 67: 249–304.

Nanda, V., J. Wang, and L. Zheng. 2004. Family values and the star phenomenon: strategies of mutual fund families. *Review of Financial Studies* 17: 667–698.

Nohel, T., J. Wang, and L. Zheng. 2010. Side-by-side management of hedge funds and mutual funds. *Review of Financial Studies* 23: 2342–2373.

Peng, L., and A. Röell. 2008. Manipulation and equity-based compensation. *American Economic Review: Papers & Proceedings* 98: 285–290.

Ruenzi, S., and A. Kempf. 2008. Tournaments in mutual fund families. *Review of Financial Studies* 21: 1013–1036.

Sialm, C., and T.M. Tham. 2011. Spillover effects in mutual fund companies. Working paper, University of Texas at Austin.

Yadav, V. 2010. Portfolio matching by multiple-fund managers: effects on fund performance and flows, Working paper, INSEAD.