JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 52, No. 2, Apr. 2017, pp. 401–425

doi:10.1017/S0022109017000096

Upper Bounds on Return Predictability

Dashan Huang and Guofu Zhou*

Abstract

Can the degree of predictability found in data be explained by existing asset pricing models? We provide two theoretical upper bounds on the R^2 of predictive regressions. Using data on the market portfolio and component portfolios, we find that the empirical R^2s are significantly greater than the theoretical upper bounds. Our results suggest that the most promising direction for future research should aim to identify new state variables that are highly correlated with stock returns instead of seeking more elaborate stochastic discount factors.

I. Introduction

In the past 4 decades, financial economists and investors have found hundreds of economic variables that can predict stock returns. Examples include short-term interest rate (Fama and Schwert (1977)), dividend yield (Fama and French (1988)), earnings-to-price ratio (Campbell and Shiller (1988)), term spreads (Campbell (1987)), book-to-market ratio (Kothari and Shanken (1997)), inflation (Campbell and Vuolteenaho (2004)), corporate issuing activity (Baker and Wurgler (2000)), consumption-to-wealth ratio (Lettau and Ludvigson (2001)), stock volatility (French, Schwert, and Stambaugh (1987)), investor sentiment (Huang, Jiang, Tu, and Zhou (2015)), and short interest (Rapach, Ringgenberg, and Zhou (2016)), among others. The evidence of return predictability has

*Huang (corresponding author), dashanhuang@smu.edu.sg, Singapore Management University; Zhou, zhou@wustl.edu, Olin Business School, Washington University. An earlier version of the paper was circulated in Nov. 2012 under the title "What Is the Maximum Return Predictability Permitted by Asset Pricing Models?" We are grateful to Rui Albuquerque, Andy Chen, Huafeng (Jason) Chen, Felipe Cortes, Phil Dybvig, Frank Fabozzi, Wayne Ferson, Hung-Gay Fung, Feng Gao, Ohad Kadan, Fang Liu, Hong Liu, Fernando Lopez, Cesare Robotti, Anjan Thakor, Ngoc-Khanh Tran, Jun Tu, Chishen Wei, Liuren Wu, Gaiyan Zhang, Joe Zhang, and seminar/conference participants at the 2012 China International Conference in Finance, 2012 Financial Management Association meetings, 2013 National University of Singapore Annual Risk Management Conference, the City University of Hong Kong, Singapore Management University, the University of Cincinnati, the University of Missouri–St. Louis, and Washington University in St. Louis, and especially to Stephen Brown (the editor) and Seth Pruitt (the referee) for extensive comments that significantly improved the paper. Huang acknowledges that the study was funded by the Singapore Management University through a research grant (C207/MSS15B005) from the Ministry of Education Academic Research Fund Tier 1.

led to the development of new asset pricing models, such as the habit formation model (Campbell and Cochrane (1999)), the long-run risks (LRR) model (Bansal and Yaron (2004)), and the rare disaster model (Barro (2006), Wachter (2013)). Although these models allow for time-varying expected returns and explain various stylized facts, such as the high volatility of the market, it is unclear whether they can explain the degree of predictability found in the data.

This article provides two model-driven upper bounds on the predictive R^2, assuming that a class of asset pricing models is true. To the best of our knowledge, we are the first to apply such bounds to various asset models, including the habit formation model, the LRR model, and the rare disaster model. Using data on the market portfolio and component portfolios, we find that the theoretical bounds are violated. This implies that these three models cannot explain the degree of predictability in the data. Moreover, all models, under certainty regulation conditions, cannot explain the degree of predictability so long as they rely on the same state variables.

Ross (2005), (2015) is the pioneer of providing a bound on return predictability. His bound is applicable to all asset pricing models under the no-arbitrage condition. Because his bound is so general, it is necessarily large. For example, Ross's bound on the market portfolio is about 20% with monthly data. If a variable generates an R^2 of 25%, the degree of predictability cannot be explained by any rational asset pricing model. In practice, however, no predictor can yield an R^2 of more than 20%. In fact, the best predictor cannot generate an R^2 consistently exceeding 5% at a monthly frequency (see, e.g., Rapach and Zhou (2013) for a recent survey of stock return predictability). Zhou (2010) provides an improvement on Ross's bound, but it is still large in many cases.

In contrast, our bounds are much tighter than the existing bounds and are binding in most of our applications. To derive our bounds, we consider only asset pricing models that use a fixed set of state variables. Therefore, our bounds are a function of the state variables. In addition, we derive the bounds by using the conditional information of the predictor, following similar insight from Hansen and Richard (1987). Our article is in the same spirit as Bekaert and Liu (2004), who show how conditional information can be used to optimally tighten Hansen and Jagannathan's (1991) original volatility bound. In short, by focusing on state variables of asset pricing models and by using conditional information of the predictor, we are able to refine Ross's (2005), (2015) bound substantially.

To apply our bounds, we consider a smaller set of asset pricing models, all of which use the same state variables x. For this set of models, the bounds on the monthly R^2 can be much smaller than 20%. For instance, suppose we find the upper bound to be 2%. If one predictor has an R^2 of 3% in an empirical setting, we can claim that all asset pricing models with the same state variables, x, cannot explain the degree of predictability in the data.

It is worthwhile to point out that our bounds are semiparametric in the sense that no knowledge of the specific functional form of the stochastic discount factor (SDF) of an asset pricing model is needed to compute the bounds and carry out the inference. In this regard, they are similar to Hansen and Jagannathan's (1991) original volatility bound. The difference is that the former bound focuses on predictability, whereas the latter bound focuses on the volatility of the SDF.

Kirby (1998) is the first to propose a joint test that relates the SDF to the R^2 of predictive regressions. However, his approach is parametric, as it requires the specific functional form of the SDF and the estimation of all the model parameters. In contrast, the functional form and estimation of parameters are unnecessary in the application of our bounds. For example, our bounds only require the same state variables used in the LRR models of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012), among other models. Once the bounds are violated, this is an indication to reject all these models and potential new models that are based on the same set of state variables.

Although Ross's (2005) bound is developed in a frictionless market (as is typically done with standard asset pricing models and other bounds, such as the variance bounds of Hansen and Jagannathan (1991) and Bakshi and Chabi-Yo (2012)), we explore the role of market frictions on R^2 bounds. Following Nagel (2013), we augment the SDF in the frictionless case with an additional state variable that captures different notions of transaction costs, such as the trading costs of Acharya and Pedersen (2005), the funding liquidity of Brunnermeier and Pedersen (2009), or the leverage constraint of Adrian, Etula, and Muir (2014). When the liquidity factor of Pástor and Stambaugh (2003) and the leverage factor of Adrian et al. are used as proxies for transaction costs, the proposed bounds implied by some of the well-known asset pricing models become larger, as might be expected. However, they are still lower than the predictive R^2s found in the data. Hence, accounting for transaction costs or market frictions does not help the aforementioned three major models to explain the degree of predictability found in the data.

The rest of the article is organized as follows: Section II provides two upper bounds on the predictive R^2 based on maximum risk aversion or market Sharpe ratio. Section III presents the data and econometric method. Section IV reports the empirical results for common predictors and some of the well-known asset pricing models, and Section V concludes.

II. Bounds

In this section, we show that the SDF of a rational asset pricing model imposes a constraint on the predictive regression, suggesting that the predictive R^2 cannot be arbitrarily large. An asset pricing model can potentially explain return predictability found in the data if its theoretical upper bound is above the empirical R^2.

A. Return Predictability

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

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

where r_{t+1} is the excess return and z_t is a predictive variable known at the end of period t. The degree of predictability is measured by the regression R^2,

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

When $R^2 > 0$, r_{t+1} can be forecasted by z_t. Otherwise, z_t is not a predictor of r_{t+1}. Welch and Goyal (2008) and Harvey, Liu, and Zhou (2016) provide the references for hundreds of predictors.

B. Bounding R^2 with Maximum Risk Aversion

An important aspect of an asset pricing theory tells us the degree of predictability that is possible. Intuitively, the degree of predictability of an asset owning a risk premium cannot be close to 1. If so, the risky asset is too predictable and one can easily arbitrage between this asset and the risk-free asset. In practice, the R^2 allowed by asset pricing models is much smaller than 1 for monthly and quarterly data.

An asset pricing model typically implies that, as shown in Cochrane (2005), the price of any asset is uniquely determined by a Euler equation, and hence its return must satisfy

$$\mathrm{E}[m(x_{t+1})r_{t+1}|I_t] = 0, \tag{3}$$

where $m(x_{t+1})$ is the normalized SDF with state variables x_{t+1} such that $\mathrm{E}(m(x_{t+1}))=1$, r_{t+1} is the return on the asset in excess of the risk-free rate, and I_t is the information set available to investors up to time t.

Because z_t belongs to I_t, we condition on z_t and rewrite (3) as

$$\mathrm{E}(r_{t+1}|z_t) = -\mathrm{cov}(m(x_{t+1}), r_{t+1}|z_t). \tag{4}$$

It is interesting to examine the implication of return predictability via this relation. If an asset pricing model can explain the predictability, the right-hand side of equation (4) must generate the same magnitude of time variation as $\alpha + \beta z_t$ in expression (1). There are two approaches to test this implication. The first is to specify both $m(\cdot)$ and x_{t+1}. The second is to develop bounds by specifying x_{t+1} alone, and this is the approach we take.

Kirby (1998) is the first to use the first approach by specifying $m(\cdot)$ and x_{t+1} for testing equation (4), and finds that the power utility, the Abel (1990) habit utility, the Epstein and Zin (1991) recursive utility, the conditional capital asset pricing model (CAPM), and the Fama and French (1993) 3-factor model cannot explain the degree of size portfolio predictability. However, the weakness of the first approach is that it is unable to determine whether the failure of a model in explaining return predictability is due to the misspecification of $m(\cdot)$ or x_{t+1}.

Ross (2005), (2015) is the first to use the second approach. In the spirit of Hansen and Jagannathan (1991), he provides an upper bound on the predictive R^2 in expression (2). Under some fairly general conditions, an investor's utility can be bounded by a constant relative risk aversion (CRRA) utility with risk aversion γ. Assume for simplicity that the investor with the maximum γ has the market portfolio as his wealth that is lognormally distributed. Then the maximum return predictability in the economy is

$$R^2 \leq \bar{R}^2_{\mathrm{Ross}} = \gamma^2\sigma^2(r_{\mathrm{mkt}}), \tag{5}$$

where $\sigma^2(r_{\mathrm{mkt}})$ is the variance of the market portfolio r_{mkt}. This bound does not depend on any specification of the SDF, $m(\cdot)$, and is, therefore, nonparametric. If

the maximum risk aversion is 10, Ross's bound will be around 20% given that the monthly volatility of the market portfolio is about 4.5%. As mentioned earlier, a bound of 20% is too large to be violated by any data.

To make the bound more stringent in practice, we use information about x_{t+1} and conditional information about predictor z_t to sharply improve the bound, without specifying a particular functional form of $m(\cdot)$. For easy applications, we state our first bound as follows:

Proposition 1. Suppose x_{t+1} is a K-dimensional vector of state variables in the SDF satisfying certain distributional assumptions, such as following an elliptical distribution. Combined with Ross's (2005) assumptions, we have

$$R^2 \quad \leq \quad \bar{R}^2_{\text{RA}} \quad = \quad \phi^2_{x,rz}\gamma^2\sigma^2(r_{\text{mkt}}), \tag{6}$$

where

$$\phi^2_{x,rz} \quad = \quad \rho^2_{x,rz}\frac{\text{var}[r_{t+1}(z_t-\mu_z)]}{\text{var}(r_{t+1})\text{var}(z_t)}, \tag{7}$$

in which μ_z is the unconditional mean of z_t and $\rho^2_{x,rz}$ is the squared multiple correlation defined as

$$\rho^2_{x,rz} \quad = \quad \frac{\text{cov}[x_{t+1},r_{t+1}(z_t-\mu_z)]'\,\text{var}^{-1}(x_{t+1})\text{cov}[x_{t+1},r_{t+1}(z_t-\mu_z)]}{\text{var}[r_{t+1}(z_t-\mu_z)]}. \tag{8}$$

Proof. See Appendix Section 1.

Proposition 1 provides a benchmark to evaluate whether an asset pricing model can explain the degree of predictability found in the data. If an asset pricing model generates an upper bound of 2%, which is larger than an $R^2=1\%$ from the data, the model can potentially explain the degree of predictability. However, if the data yield an R^2 of 3%, it is impossible for the model to explain the degree of predictability. As the bound is free of the functional form of $m(\cdot)$, all asset pricing models with the same state variables x cannot explain the degree of predictability. Therefore, research should aim to identify new state variables to develop a model that explains the time-varying expected returns of the asset.

There are three terms in bound (6). The first term, $\phi^2_{x,rz}$, can be broken down further into two terms as expression (7). The first term of expression (7) is the squared multiple correlation between the state variables and the scaled return, which is usually very small for common state variables and predictors. Because z_t is in the time t information set, $r_{t+1}(z_t-\mu_z)$ can be interpreted as a position of $z_t-\mu_z$ units of investment in r_{t+1}. The second term of expression (7) can be interpreted loosely as a standardized variance, although the unit here is not 1, but between 1 and 3 instead (see Appendix Section 3). Then, the value of $\phi^2_{x,rz}$ will be largely determined by the first term of expression (7).[1] If the state variables have a zero multiple correlation with the scaled asset return, which is generally close to the multiple correlation with the unscaled asset return, the SDF $m(x_{t+1})$ will be

[1] z_t may be replaced by any function $f(z_t)$. It is an open question whether the function $f(z_t)$ that maximizes the predictability will also optimize the bound.

uncorrelated with the asset return, and thus it will not be able to price the asset properly and explain the predictability.

The second term of bound (6), γ, is generally assumed to be less than 10, as argued by Mehra and Prescott (1985). Ross (2005) uses the insurance premium to explain that a value of 5 is large enough. Barro and Ursúa (2012) suggest that "a γ [risk aversion] of 6 seems implausibly high." Empirically, Guiso, Sapienza, and Zingales (2017) find that the average risk aversion increases from 2.87 before the 2008 crisis to 3.28 after the collapse of the financial market. Paravisini, Rappoport, and Ravina (2017) estimate the risk aversion from investors' financial decisions and find that the average risk aversion is 2.81 with a median of 1.62. It should be noted that among the models to be examined below, the risk aversions of the LRR and rare disaster models are constant, whereas the risk aversion of the habit formation model is time varying, low in good times and high in bad times. However, all of them, although manifested in different utility setups, should imply a not too large γ. In this article, following Mehra and Prescott (1985) and others, we set the maximum risk aversion to be 10 in our empirical applications.

The last term of bound (6) is the variance of the market portfolio, which is used as the proxy of the variance of the total wealth growth of the marginal investor who has the maximum risk aversion. Although we are aware that total wealth consists of financial wealth and human capital, such as labor income, the choice of focusing only on financial wealth reduces the hurdle for an asset pricing model to satisfy the R^2 bound, as the volatility of financial wealth is much larger than that of human capital. For example, at an annual frequency, Bansal, Kiku, Shaliastovich, and Yaron (2014) show that the volatility of the market portfolio is 5 times as large as that of the labor income growth.

It is worth emphasizing that the R^2 bound (6) depends only on the state variables of the SDF, the maximum risk aversion γ, and Ross's (2005) assumptions. It does not depend on the functional form of $m(\cdot)$. For example, although Bansal and Yaron (2004) and Bansal et al. (2012) assume different degrees of persistence in the consumption volatility, Proposition 1 treats them as the same, as the two models share the same state variables. Consequently, we are no longer concerned about how to estimate the parameters of the models. We can apply the same bound test to them with some moments information on the state variables.

Bound (6) is a substantial improvement over the bound of Ross (2005), which is made possible because we exploit the information of x_{t+1} in $m(\cdot)$ and z_t that predicts r_{t+1}. Comparing bound (6) with bound (5), we improve Ross's bound by introducing the term $\phi^2_{x,rz}$. In applications, $\phi^2_{x,rz}$ is often less than 10%, implying that we improve the bound by 10 times or more.

Zhou (2010), based on Kan and Zhou (2006), provides the following upper bound:

$$R^2 \quad \leq \quad \rho^2_{x,m_0}\gamma^2\sigma^2(r_{\text{mkt}}),$$

where m_0 is the minimum variance SDF in Hansen and Jagannathan (1991) and ρ_{x,m_0} is the multiple correlation between the state variables x and m_0. Although there is no analytical relation between $\phi^2_{x,rz}$ and ρ^2_{x,m_0}, unreported results reveal that $\phi^2_{x,rz}$ is almost always smaller than ρ^2_{x,m_0}, and often much smaller. Hence, our bound provided here is generally much tighter.

C. Bounding R^2 with Market Sharpe Ratio

Instead of using maximum risk aversion, the predictive R^2 can be alternatively bounded above by the market Sharpe ratio. Ross (1976) shows that the market Sharpe ratio is closely related to the volatility of SDF, which implies that extremely high Sharpe ratios are unlikely to persist. With this insight, Cochrane and Saá-Requejo (2000) use the market Sharpe ratio to bound option prices when there are either market frictions or nonmarket risks. Here, we use it to bound predictability.

If there is no arbitrage opportunity, the volatility of any SDF must satisfy the following constraint:

$$\sigma(m_{t+1}) \quad \leq \quad h \times \mathrm{SR}(r_{\mathrm{mkt}}), \tag{9}$$

where h is a free parameter chosen by the marginal investor and $\mathrm{SR}(r_{\mathrm{mkt}})$ is the market Sharpe ratio. Although Cochrane and Saá-Requejo (2000) suggest the choice of $h=2$ to rule out "good deals" (arbitrage opportunities), we choose $h=4$ to make it easier for an asset pricing model to satisfy the no-arbitrage constraint.

To bound R^2 in terms of the market Sharpe ratio $\mathrm{SR}(r_{\mathrm{mkt}})$, we have the following:

Proposition 2. Using the same assumptions on the state variables x_{t+1} in Proposition 1 and constraint (9), the predictive R^2 is bounded above as

$$R^2 \quad \leq \quad \bar{R}^2_{\mathrm{SR}} \quad = \quad \phi^2_{x,rz} \times h^2 \times \mathrm{SR}^2(r_{\mathrm{mkt}}). \tag{10}$$

Proof. See Appendix Section 2.

Bound (10) is similar to bound (6). It is also semiparametric and easy to compute. Depending on the preference, one may choose to use either bound or both. Economically, though, the maximum risk aversion γ in bound (6) should have a close relation with h to ensure the absence of arbitrage. In later applications, we show that the bounds with the choice of $\gamma=10$ and $h=4$ are numerically close to each other.

It is worth noting that our bounds have an interesting implication on cross-sectional return predictability. In the finance literature, a large number of studies find that return predictability exists and varies over portfolios sorted by market capitalization (Ferson and Harvey (1991)), book-to-market ratio (Ferson and Harvey (1991)), industry (Ferson and Harvey (1991)), and volatility (Han, Yang, and Zhou (2013)). Propositions 1 and 2 suggest that the maximum predictability of an asset is likely determined by its squared multiple correlation with the state variables of the SDF. Given the maximum risk aversion or the parameter excluding the arbitrage opportunities, an asset tends to be more predictable if it has a greater multiple correlation with the state variables, regardless of the functional form of $m(\cdot)$. This suggests the need to identify new state variables to explain cross-sectional return predictability or anomaly.

In summary, bounds (6) and (10) provide a simple test of whether a class of asset pricing models can explain the degree of predictability found in the data. They highlight the fact that the state variables in the SDF are the key factor. Therefore, if an asset pricing model with state variables x fails to explain the predictability, new state variables $y \neq x$ may explain it as long as y have a greater multiple

correlation with the asset return. This insight may help explain why Savov (2011) finds that garbage is a good proxy of consumption in explaining asset prices, as it is more volatile and more correlated with stock returns than standard consumption measures.

A remaining question is to find out how measurement errors in x affect the R^2 bound. In Appendix Section 4, we show that if the state variables x suffer from measurement errors, such as $\hat{x} = x + u$, where u are the measurement errors and are uncorrelated with x and r, the squared multiple correlation $\rho^2_{\hat{x},rz}$ is less than $\rho^2_{x,rz}$. This implies that if a given set of state variables fails to explain the degree of predictability, their proxies without additional information about the returns will fail too.

D. Impact of Market Frictions

Our bounds are derived, like many other bounds in the literature, such as those of Hansen and Jagannathan (1991) and Bakshi and Chabi-Yo (2012), under the assumption that the market is frictionless and investors can trade freely without constraints. In practice, however, there are various market frictions that can make some profitable opportunities hard to arbitrage, hence leading to return predictability. This implies that the R^2 upper bound may have to be set higher if market frictions are incorporated.

Market frictions can be the liquidity cost in Acharya and Pedersen (2005), the funding liquidity of Brunnermeier and Pedersen (2009), or the leverage constraint of Adrian et al. (2014). Nagel (2013) reviews asset pricing models with market frictions and shows that the SDF in a frictionless market can be augmented with a factor Λ_t that captures the state of transaction costs:

$$\tilde{m}_{t+1} \quad = \quad m_{t+1}\frac{\Lambda_t}{\Lambda_{t+1}}.$$

Let $\Delta\omega_{t+1} = \log(\Lambda_{t+1}/\Lambda_t)$ and $\tilde{x}_{t+1} = (x'_{t+1}, \Delta\omega_{t+1})'$. Then, we can rewrite $\tilde{m}_{t+1}$ as

$$\tilde{m}_{t+1} \quad = \quad \tilde{m}(\tilde{x}_{t+1}). \tag{11}$$

In this way, a higher $\Delta\omega_{t+1}$ means a higher transaction cost, and an asset paying well in the state of higher $\Delta\omega_{t+1}$ earns a lower expected return. Bounds (6) and (10) can be adjusted easily by including $\Delta\omega_{t+1}$ in the state variables, which is in the spirit of Nagel (2013), who shows how to incorporate liquidity into the standard Epstein–Zin–Weil SDF and leverage constraint into the CAPM SDF as a new state variable, respectively.

Naturally, a concern with expression (11) is whether including the additional variable improves the R^2 bound, because it enters the calculation of $\rho^2_{x,rz}$. In Appendix Section 5, we show theoretically that if the additional variable $\Delta\omega$ is uncorrelated with x, the squared multiple correlation can be decomposed as

$$\rho^2_{\tilde{x},rz} \quad = \quad \rho^2_{x,rz} + \rho^2_{\Delta\omega,rz}.$$

It is clear that when $\Delta\omega$ is correlated with the stock returns, $\rho^2_{\Delta\omega,rz}$ is greater than 0, suggesting that accounting for market frictions will improve the R^2 bound. In

Section IV, we show empirically that $\Delta\omega$ will raise the R^2 bounds as expected, but the increases are quantitatively small. This implies that accounting for market frictions in the three major asset pricing models still cannot explain the return predictability in the data.

III. Data and Econometric Estimation

In this section, we introduce the predictors and state variables used in this article and the econometric framework for testing whether the predictive R^2 is less than the upper bounds.

A. Data

The data are monthly spanning from Jan. 1959 to Dec. 2013. To examine the robustness of our conclusion, we also consider the results with quarterly data over the same period. The excess return of the market portfolio is the gross return on the Standard & Poor's (S&P) 500 index (including dividends) minus the gross return on the risk-free Treasury bill. As discussed by Ferson and Korajczyk (1995), it is more appropriate to use the simple return instead of the continuously compounded return in the context of this article in that the pricing equation states that the expected return is equal to the conditional covariance of the return with the marginal utility of wealth, which depends on the simple arithmetic return of the optimal portfolio. However, if the continuously compounded return is used, the result produces negligible change, and the conclusion is exactly the same.

We choose eight economic variables that display significant abilities in predicting the market portfolio:

1. BM: Book-to-market ratio extracted from Fama–French 25 size and value-sorted portfolio BM ratios with the partial least squares approach (Kelly and Pruitt (2013)).
2. TMS: Term spread between the long-term yield on government bonds and the Treasury bill rate (Welch and Goyal (2008)).
3. CAY: Consumption-to-wealth ratio (Lettau and Ludvigson (2001)).
4. CAY^{MS}: Markov-switching CAY (Bianchi, Lettau, and Ludvigson (2014)).
5. IK: Investment-to-capital ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy (Cochrane (1991)).
6. NOS: New-orders-to-shipments ratio of durable goods (Jones and Tuzel (2013)).
7. OG: Output gap, measured as the deviation of the log of total industrial production from a trend that includes both a linear and a quadratic component (Cooper and Priestley (2009)).
8. CE: Credit expansion, measured as the past quarter change of bank credit to gross domestic product (GDP) (Baron and Xiong (2017)).

It should be noted that although these eight variables are chosen to study whether the proposed bounds can explain the basic pattern of return predictability, it does not necessarily mean that other predictors are not important.

Because CAY, CAY^{MS}, IK, and CE are available at a quarterly frequency, we convert them into monthly frequency when we explore monthly predictability by assigning the most recent quarterly value to each month. For example, the observation of CAY in the first quarter of 2010 is assigned to March, April, and May 2010, respectively. Similarly, when we explore the quarterly predictability of BM, TMS, NOS, and OG, we use the averages of the April, May, and June observations for their second quarterly observations.

To calculate the R^2 upper bounds, we need the consumption growth rate, which is one of the state variables in the consumption-based asset pricing models. Following common practice, we compute the consumption growth rate as the percentage change in the seasonally adjusted, aggregate, real per capita consumption expenditures on nondurable goods and services. We use the monthly and quarterly seasonally adjusted aggregate nominal consumption expenditures on nondurables and services from National Income and Product Accounts (NIPA) Table 2.3.5 and the monthly nominal consumption expenditures from NIPA Table 2.8.5. Population numbers from NIPA Tables 2.1 and 2.6 and price deflator series from NIPA Tables 2.3.4 and 2.8.4 are used to construct the time series of per capita real consumption figures. Finally, data on the component portfolios sorted by size, book-to-market ratio, and industry are taken from Kenneth French's Web site (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

B. State Variables in the SDF

Since Mehra and Prescott (1985), various consumption-based asset pricing models have been developed to explain the equity risk premium puzzle and other features of the data. Among them, the habit formation model, the LRR model, and the rare disaster model are the three most studied, because they all can generate time-varying expected returns and therefore can explain return predictability to some degree in a qualitative way. Hence, we focus on these three models and investigate whether they can explain the degree of predictability quantitatively.

1. Habit Formation Model

The habit formation model assumes that risk aversion is time varying over economic business cycles. The risk aversion is high in economic recessions when investors require a high premium for taking risks, but it is low in economic expansions when investors require a low premium. Countercyclical risk aversion suggests that the risk premium is countercyclical; hence, stock returns are predictable.

Suppose the log consumption growth rate $\Delta c_{t+1} = \log(C_{t+1}/C_t)$ is independently and identically distributed (i.i.d.) and follows a normal distribution:

$$\Delta c_{t+1} = g + \sigma \varepsilon_{c,t+1}, \quad \varepsilon_{c,t+1} \sim \text{i.i.d. N}(0,1), \tag{12}$$

and where g and σ are the mean and volatility of consumption growth, respectively.

Let X_t be the external habit at time t. According to Campbell and Cochrane (1999), the surplus consumption ratio is $S_t = (C_t - X_t)/C_t$, and its logarithm is assumed to follow the following process:

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)(\Delta c_{t+1} - g),$$

where $\phi, \bar{s}$, and g are parameters, and the sensitivity function $\lambda(s_t)$ is given by

$$\lambda(s_t) = \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})}-1, & s_t < \bar{s}+\frac{1}{2}\left(1-\bar{S}^2\right), \\ 0, & s_t \geq \bar{s}+\frac{1}{2}\left(1-\bar{S}^2\right), \end{cases}$$

where $\bar{S} = \sigma\sqrt{\gamma/(1-\phi)}$ is the steady-state surplus consumption ratio.

The SDF of the habit formation model is

$$m_{t+1} = \delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-A} = \delta e^{-A[g+(\phi-1)(s_t-\bar{s})+(1+\lambda(s_t))(\Delta c_{t+1}-g)]}.$$

In this model, s_{t+1} is the only state variable and can be extracted according to Campbell and Cochrane (1999).

2. LRR Model

Because the LRR model uses the low-frequency time series properties of dividends and aggregate consumption, it can simultaneously explain the equity risk premium puzzle, the risk-free rate puzzle, and the high level of market volatility. The key assumptions are that the consumption growth rate and the dividend growth rate follow the joint dynamics as:

$$\begin{aligned} \Delta c_{t+1} &= g+\mu_{c,t}+\sigma_t\varepsilon_{c,t+1}, \\ \mu_{c,t+1} &= \rho_\mu\mu_{c,t}+\psi_c\sigma_t\varepsilon_{\mu,t+1}, \\ \sigma^2_{t+1} &= (1-\nu)\bar{\sigma}^2+\nu\sigma^2_t+\sigma_w\varepsilon_{\sigma,t+1}, \\ d_{t+1}-c_{t+1} &= \mu_{\mathrm{DC}}+\mathrm{DC}_{t+1}, \\ \mathrm{DC}_{t+1} &= \lambda_{\mathrm{DC}}\mu_{c,t}+\rho_{\mathrm{DC}}\mathrm{DC}_t+\psi_{\mathrm{DC}}\sigma_t\varepsilon_{\mathrm{DC},t+1}, \\ \Delta d_{t+1} &= g+(1+\lambda_{\mathrm{DC}})\mu_{c,t}+(\rho_{\mathrm{DC}}-1)\mathrm{DC}_t+\sigma_t\varepsilon_{c,t+1}+\psi_{\mathrm{DC}}\sigma_t\varepsilon_{\mathrm{DC},t+1}, \end{aligned}$$

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

This model has four state variables: the two latent variables $\mu_{c,t}$ and σ^2_t, the observable consumption growth rate Δc_t, and the dividend–consumption ratio DC_t. Constantinides and Ghosh (2011) show that the log price–dividend ratio and log risk-free rate are affine functions of the two latent variables and suggest that one can invert the system and express the two latent state variables as known affine functions of the observable aggregate log price–dividend ratio and log risk-free rate. Hence, the state variables in SDF for the LRR model are $\Delta c_{t+1}, r_{f,t+1}, dp_{t+1}$, and DC_{t+1}.

3. Rare Disaster Model

The rare disaster model revived by Barro (2006) is intended to resolve the equity risk premium puzzle and does not accommodate time-varying expected returns. Wachter (2013) extends Barro to allow for time-varying probability of disasters, thereby generating return predictability.

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

$$\Delta c_{t+1} = \begin{cases} g + \sigma \varepsilon_{c,t+1}, & \text{with probability } 1 - p_t, \\ g + \sigma \varepsilon_{c,t+1} + \log(1-b), & \text{with probability } p_t, \end{cases}$$

where $\varepsilon_{c,t+1}$ is i.i.d. $\mathrm{N}(0,1)$, and p_t is the probability of disasters. The crucial question is to find a variable to proxy for the unobservable p_t. Wachter (2013) considers the rare disaster model in a continuous-time setting and finds 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, we use the consumption growth rate and the dividend–price ratio as the two state variables for the rare disaster model.

C. Wald Test

The parameters needed 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} can be multidimensional. The moment conditions are

$$h(y_{t+1}, \theta) = \begin{pmatrix} y_{t+1} - \mu_y \\ y_{t+1} y'_{t+1} - (\Sigma_y + \mu_y \mu'_y) \end{pmatrix}, \tag{13}$$

where $\mu_y = \mathrm{E}(y_{t+1})$ and $\Sigma_y = \mathrm{cov}(y_{t+1})$. Because the econometric specification in expression (13) is exactly identified, the generalized method of moments estimator of $\theta = (\mu_y, \Sigma_y)$ is the vector of values that set $1/T \sum_{t=1}^{T} h(y_{t+1}, \theta)$ equal to 0, where T is the number of sample observations.

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

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

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

We use the Wald test to evaluate whether $R^2 \leq \bar{R}^2_{\mathrm{RA}}$ or $\bar{R}^2_{\mathrm{SR}}$, which is equivalent to a one-sided test for $f(\theta_{\mathrm{RA}}) = 0$ or $f(\theta_{\mathrm{SR}}) = 0$, where θ_{RA} and θ_{SR} are the moment parameters used in $f(\theta_{\mathrm{RA}}) = R^2 - \bar{R}^2_{\mathrm{RA}}$ and $f(\theta_{\mathrm{SR}}) = R^2 - \bar{R}^2_{\mathrm{SR}}$. Let Σ_{RA} and Σ_{SR} be the corresponding covariances of θ_{RA} and θ_{SR}. The Wald statistic is

$$W_{\mathrm{RA}} = T f(\hat{\theta}_{\mathrm{RA}}) \left[\frac{df}{d\theta_{\mathrm{RA}}} \hat{\Sigma}_{\mathrm{RA}} \frac{df}{d\theta_{\mathrm{RA}}} \right]^{-1} f(\hat{\theta}_{\mathrm{RA}}) \xrightarrow{d} \chi^2(1)$$

for the bound with maximum risk aversion and

$$W_{\mathrm{SR}} = T f(\hat{\theta}_{\mathrm{SR}}) \left[\frac{df}{d\theta_{\mathrm{SR}}} \hat{\Sigma}_{\mathrm{SR}} \frac{df}{d\theta_{\mathrm{SR}}} \right]^{-1} f(\hat{\theta}_{\mathrm{SR}}) \xrightarrow{d} \chi^2(1)$$

for the bound with market Sharpe ratio.

IV. Empirical Results

In this section, we compute the R^2 upper bounds and examine whether the three major asset pricing models can explain the degree of predictability found in the data. We investigate both the market portfolio and component portfolio sorted by size, value, and industry.

A. R^2 Bounds on the Market Portfolio

Suppose r_{t+1} in expression (1) is the excess return on the market portfolio. Table 1 reports the predictive regression results, the upper bound of Ross (2005) $\bar{R}^2_{\text{Ross}}$, the multiple correlation coefficient $\rho_{x,rz}$ that determines the improvement of our bounds over Ross's, and the two proposed bounds. The statistical significance is assessed by the Wald statistic that tests the hypothesis that the predictive R^2 is less than the upper bound.

TABLE 1

R^2 Bounds on the Market Portfolio

Table 1 reports the results of regressing the excess market return on one of the predictors in column 1 and the upper bounds of the predictive R^2 derived from three asset pricing models: habit formation model, long-run risks (LRR) model, and rare disaster model. The predictors include the book-to-market ratio (BM), term spread (TMS), consumption–wealth ratio (CAY), Markov-switching CAY (CAY$^{\text{MS}}$), investment–capital ratio (IK), new-orders-to-shipments ratio (NOS), output gap (OG), and credit expansion (CE). The sample period is Jan. 1959–Dec. 2013. $\rho_{x,rz}$ is the multiple correlation defined in column 8. $\bar{R}^2_{\text{Ross}}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{\text{RA}}$ and $\bar{R}^2_{\text{SR}}$ are the proposed bounds. 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% and 5% levels, respectively.

					Habit Formation Model			LRR Model			Rare Disaster Model		
Predictor	β	t-Stat.	R^2(%)	$\bar{R}^2_{\text{Ross}}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{\text{RA}}$(%)	$\bar{R}^2_{\text{SR}}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{\text{RA}}$(%)	$\bar{R}^2_{\text{SR}}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{\text{RA}}$(%)	$\bar{R}^2_{\text{SR}}$(%)
1	2	3	4	5	6	7	8	9	10	11	12	13	14
BM	0.58	3.06	1.42	20.16	0.02	0.01**	0.01**	0.07	0.11**	0.11**	0.05	0.05**	0.05**
TMS	0.24	2.03	0.66	20.16	0.06	0.07**	0.07**	0.10	0.20**	0.20**	0.06	0.08**	0.08**
CAY	0.25	2.90	1.09	20.16	0.06	0.07**	0.08**	0.12	0.24**	0.24**	0.04	0.03**	0.03**
CAY$^{\text{MS}}$	0.49	3.22	1.78	20.16	0.06	0.09**	0.09**	0.12	0.33**	0.33**	0.01	0.00**	0.00**
IK	−1.11	−2.25	0.84	20.16	0.12	0.34**	0.34**	0.17	0.66*	0.67*	0.15	0.48**	0.49**
NOS	−0.20	−3.10	1.64	20.16	0.08	0.15**	0.15**	0.12	0.33**	0.34**	0.07	0.12**	0.13**
OG	−0.09	−3.36	1.71	20.16	0.06	0.07**	0.07**	0.14	0.40**	0.41**	0.12	0.32**	0.32**
CE	−0.32	−2.07	0.75	20.16	0.09	0.20**	0.20**	0.13	0.40**	0.40**	0.08	0.17**	0.18**

Column 1 of Table 1 indicates the predictors. Columns 2, 3, and 4 report the regression slope, t-statistic, and R^2 for each predictor, respectively. Because all the t-statistics are greater than 2 in absolute value, the eight predictors we consider significantly predict the market portfolio. The associated R^2s range from 0.66% for TMS to 1.78% for CAY$^{\text{MS}}$. Together with t-statistics greater than 2, the sizable R^2s suggest that the market portfolio is predictable and that the degree of predictability varies across the predictors, which in turn implies that the upper bound of R^2 should vary across the predictors.

As discussed in the previous section, we assume that the maximum risk aversion is 10 throughout. The upper bound of Ross (2005), $\bar{R}^2_{\text{Ross}}$, is reported in column 5 of Table 1, which has a constant value of 20.16% regardless of the predictor and asset pricing model used. Because the maximum R^2 with respect to the eight predictors is only 1.78%, Ross's bound is satisfied for all predictors and all models. To the best of our knowledge, there is no single predictor that can generate an R^2 as large as 20.16%, or close to it, at a monthly frequency. Therefore, Ross's bound is unable to reject any of the models for explaining the degree of market portfolio predictability.

Columns 6, 9, and 12 of Table 1 report the multiple correlation $\rho_{x,rz}$ between the state variables and the market portfolio, where x are the state variables used in the habit formation model, the LRR model, and the rare disaster model,

respectively. In comparison to Ross's (2005) bound, the proposed bounds have improved by a factor of $1/\phi^2_{x,rz}$. Over our sample period, the standardized variance term $\text{var}(r_{t+1}(z_t - \mu_z))/\text{var}(r_{t+1})\text{var}(z_t)$ of $\phi^2_{x,rz}$ is between 1 and 3, indicating that the improvement of our bounds comes mainly from the multiple correlation $\rho_{x,rz}$. Because the maximum $\rho_{x,rz}$ within columns 6, 9, and 12 is 0.17, our bounds are at least 11 times lower than Ross's.

Columns 7, 10, and 13 of Table 1 report the proposed R^2 bounds by assuming a maximum risk aversion of 10 with respect to the three asset pricing models. As expected, the low-value $\rho_{x,rz}$ drives the R^2 upper bound close to 0 for all three sets of state variables for the habit formation model, the LRR model, and the rare disaster model. All eight predictors display larger R^2s than the theoretical upper bounds. In other words, the three sets of state variables explain a small fraction of predictability found in the data. Therefore, we can conclude that asset pricing models with the same state variables as the habit formation model, the LRR model, and the rare disaster model cannot explain the magnitude of return predictability.

Because there is no consensus on the choice of the level of maximum risk aversion, a value of 10 may be too small for some extremely conservative investors. Columns 8, 11, and 14 of Table 1 report R^2 bounds based on the market Sharpe ratio. As discussed earlier, we set the free parameter h in bound (10) to be 4, which means that if the monthly market Sharpe ratio is 0.11, any strategy with a monthly Share ratio larger than 0.44 will be a violation of the no-arbitrage condition. It is surprising that the upper bounds with this market Sharpe ratio constraint are similar to those with the maximum risk aversion constraint, suggesting that our two bounds complement each other and can be used interchangeably, according to the user's taste.

Although we focus on the most frequently explored monthly frequency of data, it is of interest to see how the results of Table 1 change when the predictability is examined quarterly. Table 2 reports the results with quarterly data over the same sample period. Compared to Table 1, the predictive R^2s increase significantly, with the exception of TMS, and seven of the eight predictors are larger than 2%. Again, CAY^{MS} stands out as the most pronounced predictor with an R^2 of 4.98%. This is consistent with the predictability literature that finds that the longer the forecasting horizon, the greater the degree of predictability. Theoretically, this appears to hold true as Ross's (2005) bound increases to 68.48%, which is much larger than any of the R^2s of the data. Our proposed bounds are greater than before as well. However, because $\rho_{x,rz}$ is less than or equal to 0.21 for almost all the predictors, our bounds still improve Ross's bound by 7 times or more. For the habit formation model, seven predictors generate larger R^2s than the bounds, six of which are significant. For the LRR model, five predictors deliver significantly larger R^2s than the theoretical bounds. The result for the rare disaster model is similar, with six of the eight predictors violating the bounds. In sum, models based on each of the three sets of state variables cannot explain the predictability of the data.

In examining the effects of market frictions on the R^2 bounds, we first consider the liquidity factor constructed by Pástor and Stambaugh (2003) as a proxy for transaction costs. The monthly data span Aug. 1962 to Dec. 2013, and the

TABLE 2

R^2 Bounds on the Market Portfolio with Quarterly Data

Table 2 reports the results of regressing the excess market return on one of the predictors in column 1 and the upper bounds of the predictive R^2 derived from three asset pricing models: habit formation model, long-run risks (LRR) model, and rare disaster model. The predictors include the book-to-market ratio (BM), term spread (TMS), consumption–wealth ratio (CAY), Markov-switching CAY (CAY^{MS}), investment–capital ratio (IK), new-orders-to-shipments ratio (NOS), output gap (OG), and credit expansion (CE). The sample period is 1959Q1–2013Q4. $\rho_{x,rz}$ is the multiple correlation defined in column 8. $\bar{R}^2_{Ross}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ are the proposed bounds. 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% and 5% levels, respectively.

					Habit Formation Model			LRR Model			Rare Disaster Model		
Predictor	β	t-Stat.	R^2(%)	$\bar{R}^2_{Ross}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
1	2	3	4	5	6	7	8	9	10	11	12	13	14
BM	0.01	2.39	2.48	68.48	0.01	0.01**	0.01**	0.09	0.57**	0.56**	0.04	0.14**	0.13**
TMS	0.36	0.80	0.36	68.48	0.03	0.06	0.05	0.10	0.88	0.85	0.08	0.59	0.57
CAY	0.75	2.84	2.91	68.48	0.10	0.60**	0.58**	0.18	2.16**	2.10**	0.10	0.60**	0.59**
CAY^{MS}	1.51	3.44	4.98	68.48	0.09	0.59**	0.57**	0.17	2.34**	2.28**	0.03	0.08**	0.08**
IK	−3.61	−2.48	2.75	68.48	0.20	2.93	2.85	0.19	2.76	2.68	0.11	0.92**	0.89**
NOS	−0.17	−3.14	4.59	68.48	0.09	0.92**	0.89**	0.18	3.51**	3.41*	0.09	0.87**	0.85**
OG	−0.25	−3.21	4.51	68.48	0.10	0.76**	0.74**	0.16	1.94**	1.88**	0.15	1.57**	1.52**
CE	−0.99	−2.14	2.21	68.48	0.15	1.96	1.90	0.21	3.74	3.64	0.16	2.07	2.02

results are reported in Table 3. Ross's (2005) bound increases from 20.16% in the frictionless case to 20.54% in the friction case. The change is so small that it makes no difference in the inference of the degree of predictability. In contrast, the changes for our new bounds using the habit formation model are relatively large. For example, the bounds of using the investment–capital ratio predictor double from 0.34% to 0.78% and from 0.34% to 0.82% for the bounds with the maximum risk aversion rate and the market Sharpe ratio, respectively. However, the bounds are still small compared to regression R^2 values. Indeed, similar to Table 1, the bounds are binding in all cases. Hence, the conclusion is identical to the earlier case of no market frictions.

We consider the leverage constraint of Adrian et al. (2014) as another proxy of market frictions, which can be included in the SDF as an additional

TABLE 3

R^2 Bounds on the Market Portfolio with Liquidity Cost

Table 3 reports the results of regressing the excess market return on one of the predictors in column 1 and the upper bounds of the predictive R^2 derived from three asset pricing models, habit formation model, long-run risks (LRR) model, and rare disaster model, where the Pástor and Stambaugh (2003) liquidity factor is used as the proxy for liquidity cost. The predictors include the book-to-market ratio (BM), term spread (TMS), consumption–wealth ratio (CAY), Markov-switching CAY (CAY^{MS}), investment–capital ratio (IK), new-orders-to-shipments ratio (NOS), output gap (OG), and credit expansion (CE). The sample period is Aug. 1962–Dec. 2013. $\rho_{x,rz}$ is the multiple correlation defined in column 8. $\bar{R}^2_{Ross}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ are the proposed bounds. 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% and 5% levels, respectively.

					Habit Formation Model			LRR Model			Rare Disaster Model		
Predictor	β	t-Stat.	R^2(%)	$\bar{R}^2_{Ross}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
1	2	3	4	5	6	7	8	9	10	11	12	13	14
BM	0.67	3.44	1.83	20.54	0.11	0.26**	0.28**	0.08	0.13**	0.13**	0.06	0.07**	0.07**
TMS	0.23	1.93	0.63	20.54	0.07	0.12**	0.12**	0.11	0.24**	0.25**	0.07	0.11**	0.11**
CAY	0.24	2.72	1.01	20.54	0.07	0.08**	0.08**	0.12	0.28**	0.29**	0.04	0.03**	0.03**
CAY^{MS}	0.48	3.08	1.74	20.54	0.12	0.34**	0.36**	0.12	0.36**	0.38**	0.01	0.00**	0.00**
IK	−1.28	−2.41	1.04	20.54	0.18	0.78	0.82	0.17	0.68**	0.71**	0.14	0.46**	0.48**
NOS	−0.19	−2.89	1.55	20.54	0.12	0.39**	0.41**	0.12	0.34**	0.36**	0.07	0.13**	0.14**
OG	−0.10	−3.45	1.98	20.54	0.15	0.48**	0.51**	0.15	0.48**	0.51**	0.12	0.31**	0.32**
CE	−0.36	−2.21	0.93	20.54	0.10	0.24**	0.25**	0.15	0.60**	0.63**	0.11	0.28**	0.29**

state variable. The rationale is that deleveraging indicates deteriorating funding conditions. The data are collected quarterly from 1968Q1 to 2009Q4.[2] Table 4 reports the results. Similar to Pástor and Stambaugh's (2003) liquidity factor, the bounds are generally greater but are still significantly less than the predictive R^2s. Summarizing Tables 3 and 4, although market frictions can make the theoretical R^2 bounds greater, the bounds are still binding in almost all cases. This indicates that even after accounting for market frictions, asset pricing models based on one of the three sets of state variables still have difficulty explaining the magnitude of predictability of the market portfolio.

TABLE 4

R^2 Bounds on the Market Portfolio with Leverage Constraint

Table 4 reports the results of regressing the excess market return on one of the predictors in column 1 and the upper bounds of the predictive R^2 derived from three asset pricing models, habit formation model, long-run risks (LRR) model, and rare disaster model, where the Adrian, Etula, and Muir (2014) leverage factor is used as the proxy for leverage constraint. The predictors include the book-to-market ratio (BM), term spread (TMS), consumption–wealth ratio (CAY), Markov-switching CAY (CAY^{MS}), investment–capital ratio (IK), new-orders-to-shipments ratio (NOS), output gap (OG), and credit expansion (CE). The sample period is Mar. 1968–Feb. 2010. $\rho_{x,rz}$ is the multiple correlation defined in column 8. $\bar{R}^2_{Ross}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ are the proposed bounds. 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% and 5% levels, respectively.

					Habit Formation Model			LRR Model			Rare Disaster Model		
Predictor	β	t-Stat.	R^2(%)	$\bar{R}^2_{Ross}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)	$\rho_{x,rz}$	$\bar{R}^2_{RA}$(%)	$\bar{R}^2_{SR}$(%)
1	2	3	4	5	6	7	8	9	10	11	12	13	14
BM	0.67	3.00	1.75	22.18	0.10	0.22**	0.13**	0.13	0.38**	0.21**	0.11	0.26**	0.15**
TMS	0.26	1.87	0.71	22.18	0.07	0.12**	0.07**	0.13	0.36**	0.20**	0.09	0.19**	0.11**
CAY	0.25	2.80	1.24	22.18	0.08	0.11**	0.06**	0.12	0.27**	0.15**	0.06	0.06**	0.04**
CAY^{MS}	0.51	3.08	2.03	22.18	0.03	0.02**	0.01**	0.11	0.30**	0.17**	0.03	0.02**	0.01**
IK	−1.28	−1.94	0.81	22.18	0.13	0.43**	0.24**	0.16	0.65	0.36**	0.13	0.40**	0.22**
NOS	−0.22	−3.01	2.01	22.18	0.25	1.68	0.95**	0.15	0.59**	0.33**	0.06	0.09**	0.05**
OG	−0.10	−2.53	1.41	22.18	0.17	0.77**	0.43**	0.16	0.68**	0.38**	0.12	0.36**	0.20**
CE	−0.36	−2.07	0.93	22.18	0.22	1.19	0.67*	0.19	0.95	0.53**	0.12	0.39**	0.22**

B. Bounds on Component Portfolio Predictability

In this section, we examine whether the proposed bounds are also binding for component portfolio predictability. Theoretically, our proposed bounds, (6) and (10), should have different values for different portfolios, because they have different multiple correlations with the state variables. Hence, it is an empirical question as to how the bounds vary at the portfolio level.

Tables 5 and 6 report the R^2s and their upper bounds on portfolios sorted by size and value (book-to-market ratio). There are three interesting observations. First, the eight predictors not only predict the market portfolio as shown in Table 1, but also predict all of the component portfolios with sizable R^2s. The degree of predictability is almost the same as the market portfolio. Second, the upper bounds are violated in all cases, with only a few exceptions. Finally, with respect to the three types of asset pricing models, the LRR model performs the best with the largest bounds. Nevertheless, the bounds are still violated in most cases.

Table 7 reports further results on portfolios sorted by industry. For brevity, we report only the first three predictors: BM, TMS, and CAY. Consistent with Ferson

[2]We are grateful to Tyler Muir for making the data available on his Web site, http://faculty.som.yale.edu/tylermuir/researchpapers.html.

TABLE 5
R^2 Bounds on Size Portfolios

Table 5 reports the results of regressing size portfolio returns on one of the predictors in column 1 and the upper bounds of the predictive R^2 derived from three asset pricing models: habit formation model, long-run risks (LRR) model, and rare disaster model. The predictors include the book-to-market ratio (BM), term spread (TMS), consumption–wealth ratio (CAY), Markov-switching CAY (CAY^{MS}), investment–capital ratio (IK), new-orders-to-shipments ratio (NOS), output gap (OG), and credit expansion (CE). The sample period is Jan. 1959–Dec. 2013. $\bar{R}^2_{Ross}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ are the proposed bounds. 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% and 5% levels, respectively.

		Habit Formation Model		LRR Model		Rare Disaster Model	
Predictor	R^2	$\bar{R}^2_{RA}$	$\bar{R}^2_{SR}$	$\bar{R}^2_{RA}$	$\bar{R}^2_{SR}$	$\bar{R}^2_{RA}$	$\bar{R}^2_{SR}$
1	2	3	4	5	6	7	8
Panel A. Small Size Portfolio							
BM	1.26	0.00**	0.00**	0.07**	0.07**	0.01**	0.01**
TMS	0.60	0.08**	0.08**	0.17**	0.17**	0.06**	0.06**
CAY	0.30	0.03	0.03	0.26	0.26	0.03	0.03
CAY^{MS}	0.83	0.03**	0.03**	0.41**	0.41**	0.02**	0.02**
IK	0.36	0.36	0.37	0.51	0.52	0.33	0.34
NOS	1.38	0.21**	0.21**	0.28**	0.28**	0.10**	0.10**
OG	0.70	0.11**	0.11**	0.32**	0.32**	0.22**	0.22**
CE	1.04	0.13**	0.13**	0.44**	0.45**	0.18**	0.18**
Panel B. Median-Size Portfolio							
BM	1.34	0.00**	0.00**	0.11**	0.11**	0.03**	0.03**
TMS	0.70	0.10**	0.10**	0.14**	0.14**	0.05**	0.05**
CAY	0.62	0.06**	0.06**	0.20**	0.20**	0.02**	0.02**
CAY^{MS}	1.30	0.06**	0.07**	0.36**	0.36**	0.01**	0.01**
IK	0.69	0.41	0.42	0.67	0.68	0.48	0.49
NOS	1.47	0.21**	0.21**	0.32**	0.32**	0.07**	0.07**
OG	1.21	0.10**	0.10**	0.37**	0.37**	0.27**	0.27**
CE	0.87	0.19**	0.19**	0.48**	0.49**	0.22**	0.22**
Panel C. Large Size Portfolio							
BM	1.39	0.01**	0.01**	0.12**	0.13**	0.06**	0.06**
TMS	0.60	0.07**	0.07**	0.22**	0.22**	0.10**	0.10**
CAY	1.23	0.08**	0.08**	0.25**	0.26**	0.04**	0.04**
CAY^{MS}	1.83	0.09**	0.09**	0.28**	0.29**	0.00**	0.00**
IK	0.81	0.33**	0.33**	0.65	0.66	0.46**	0.47**
NOS	1.60	0.14**	0.14**	0.34**	0.34**	0.13**	0.14**
OG	1.78	0.06**	0.06**	0.42**	0.43**	0.33**	0.34**
CE	0.67	0.20**	0.21**	0.36**	0.37**	0.15**	0.16**

and Harvey (1991) and Ferson and Korajczyk (1995), industry portfolios are significantly predictable. However, the degree of predictability varies substantially across industries. For example, when BM is the predictor, the most predictable industry is the consumer durable industry, which has an R^2 of 1.82%, greater than the predictability on the market portfolio, whereas the least predictable industry is the energy industry, which has an R^2 of 0.32%. The R^2 upper bounds are binding and have similar magnitudes as the other component portfolios in Tables 5 and 6.

Overall, the results for the component portfolios are similar to those for the market portfolio, and the three sets of state variables have difficulty explaining the magnitude of predictability.

V. Conclusion

In this article, we investigate whether return predictability found in the data is consistent with asset pricing models. To answer this question, we develop two upper bounds on the predictive R^2. Our bounds have improved substantially over the nonbinding bound of Ross (2005), and they provide likely reasons a given

TABLE 6
R^2 Bounds on Value Portfolios

Table 6 reports the results of regressing value portfolio returns on one of the predictors in column 1 and the upper bounds of the predictive R^2 derived from three asset pricing models: habit formation model, long-run risks (LRR) model, and rare disaster model. The predictors include the book-to-market ratio (BM), term spread (TMS), consumption–wealth ratio (CAY), Markov-switching CAY (CAY^{MS}), investment–capital ratio (IK), new-orders-to-shipments ratio (NOS), output gap (OG), and credit expansion (CE). The sample period is Jan. 1959–Dec. 2013. $\bar{R}^2_{Ross}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{RA}$ and $\bar{R}^2_{SR}$ are the proposed bounds. 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% and 5% levels, respectively.

		Habit Formation Model		LRR Model		Rare Disaster Model	
Predictor	R^2	$\bar{R}^2_{RA}$	$\bar{R}^2_{SR}$	$\bar{R}^2_{RA}$	$\bar{R}^2_{SR}$	$\bar{R}^2_{RA}$	$\bar{R}^2_{SR}$
1	2	3	4	5	6	7	8
Panel A. Low-BM (Growth) Portfolio							
BM	1.73	0.02**	0.02**	0.11**	0.11**	0.05**	0.05**
TMS	0.67	0.07**	0.07**	0.27**	0.28**	0.12**	0.12**
CAY	1.04	0.06**	0.06**	0.24**	0.24**	0.05**	0.05**
CAY^{MS}	1.76	0.08**	0.08**	0.27**	0.27**	0.00**	0.00**
IK	0.73	0.30**	0.31**	0.67	0.68	0.50	0.51
NOS	1.38	0.09**	0.09**	0.26**	0.26**	0.12**	0.13**
OG	1.52	0.05**	0.05**	0.40**	0.40**	0.31**	0.31**
CE	0.71	0.15**	0.15**	0.31**	0.31**	0.15**	0.15**
Panel B. Median-BM Portfolio							
BM	0.96	0.00**	0.00**	0.15**	0.15**	0.05**	0.05**
TMS	0.56	0.09**	0.09**	0.12**	0.12**	0.06**	0.06**
CAY	1.07	0.08**	0.08**	0.16**	0.16**	0.02**	0.02**
CAY^{MS}	1.54	0.07**	0.07**	0.22**	0.23**	0.00**	0.00**
IK	0.74	0.41**	0.41**	0.60	0.61	0.32**	0.33**
NOS	1.92	0.21**	0.22**	0.46**	0.47**	0.11**	0.12**
OG	1.60	0.10**	0.10**	0.46**	0.47**	0.28**	0.28**
CE	0.72	0.28**	0.28**	0.49**	0.50**	0.19**	0.19**
Panel C. High-BM (Value) Portfolio							
BM	1.21	0.00**	0.00**	0.25**	0.25**	0.09**	0.09**
TMS	0.22	0.08	0.08	0.09	0.09	0.02	0.05
CAY	0.21	0.04	0.04	0.16	0.16	0.01	0.01
CAY^{MS}	0.66	0.06**	0.07**	0.31**	0.31**	0.02**	0.02**
IK	0.51	0.38	0.38	0.51	0.52	0.23*	0.23*
NOS	1.34	0.27**	0.28**	0.49**	0.49**	0.12**	0.12**
OG	1.04	0.12**	0.12**	0.44**	0.45**	0.24**	0.25**
CE	0.72	0.19**	0.19**	0.65	0.66	0.10**	0.10**

asset pricing model cannot explain predictability. In forecasting the market portfolio or component portfolios sorted by size, value, and industry, we find that the high predictive R^2s observed in the data almost always exceed the proposed theoretical upper bounds, implying that return predictability cannot be fully explained by any asset pricing models based on three sets of well-known state variables underlying the habit formation model (Campbell and Cochrane (1999)), the LRR model (Bansal and Yaron (2004)), and the rare disaster model (Barro (2006), Wachter (2013)). The reason may be the low correlations between the return(s) and the state variables. This conclusion is unaltered even when market frictions are accounted for.

Although our study focuses on the stock market, it may be useful for studying other asset classes, such as options, bonds, and foreign exchanges, to examine whether the predictability is consistent with existing models in these areas. Moreover, it appears important to extend our bounds to allow for parameter instability and structural breaks. However, this is technically very challenging, as it leads to difficulties similar to those we faced when extending Hansen and

TABLE 7

R^2 Bounds on Industry Portfolios

Table 7 reports the results of regressing industry portfolio returns on the book-to-market ratio (BM), term spread (TMS), or consumption–wealth ratio (CAY) and the upper bounds of the predictive R^2 derived from three asset pricing models: habit formation model, long-run risks (LRR) model, and rare disaster model. The sample period is Jan. 1959–Dec. 2013. $\bar{R}^2_{\text{Ross}}$ is Ross's (2005) bound on the R^2. $\bar{R}^2_{\text{RA}}$ and $\bar{R}^2_{\text{SR}}$ are the proposed bounds. 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% and 5% levels, respectively.

		Habit Formation Model		LRR Model		Rare Disaster Model	
Portfolio	R^2	$\bar{R}^2_{\text{RA}}$	$\bar{R}^2_{\text{SR}}$	$\bar{R}^2_{\text{RA}}$	$\bar{R}^2_{\text{SR}}$	$\bar{R}^2_{\text{RA}}$	$\bar{R}^2_{\text{SR}}$
1	2	3	4	5	6	7	8
Panel A. Predictor: BM							
NODUR	0.50	0.00**	0.00**	0.22**	0.22**	0.10**	0.10**
DURBL	1.82	0.00**	0.00**	0.16**	0.16**	0.08**	0.09**
MANUF	1.34	0.01**	0.01**	0.15**	0.15**	0.08**	0.08**
ENRGY	0.32	0.00**	0.00**	0.11	0.11	0.04	0.04
HITEC	1.30	0.02**	0.02**	0.15**	0.15**	0.06**	0.07**
TELCM	0.67	0.00**	0.00**	0.06**	0.06**	0.02**	0.02**
SHOPS	1.20	0.01**	0.01**	0.15**	0.15**	0.07**	0.07**
HLTH	0.50	0.00**	0.00**	0.06**	0.06**	0.03**	0.03**
UTILS	0.49	0.00**	0.00**	0.20**	0.21**	0.06**	0.06**
OTHER	0.97	0.00**	0.00**	0.08**	0.08**	0.04**	0.04**
Panel B. Predictor: TMS							
NODUR	0.43	0.12*	0.13*	0.24*	0.24*	0.09**	0.09**
DURBL	1.14	0.07**	0.07**	0.14**	0.14**	0.07**	0.07**
MANUF	0.80	0.09**	0.09**	0.16**	0.16**	0.05**	0.05**
ENRGY	0.14	0.04	0.04	0.07	0.07	0.01	0.01
HITEC	0.62	0.04**	0.04**	0.33**	0.34**	0.09*	0.09*
TELCM	0.27	0.12	0.13	0.04	0.04	0.01	0.01
SHOPS	0.50	0.10**	0.10**	0.12**	0.12**	0.09**	0.09**
HLTH	0.01	0.14	0.14	0.26	0.27	0.22	0.22
UTILS	0.26	0.06	0.06	0.17	0.18	0.08	0.09
OTHER	0.37	0.04**	0.04**	0.16	0.16	0.12*	0.12*
Panel C. Predictor: CAY							
NODUR	1.43	0.13**	0.13**	0.27**	0.27**	0.09**	0.09**
DURBL	0.20	0.02	0.02	0.19	0.19	0.05	0.05
MANUF	0.56	0.06**	0.06**	0.15**	0.15**	0.02**	0.02**
ENRGY	0.11	0.02	0.02	0.13	0.14	0.05	0.05
HITEC	0.61	0.02**	0.02**	0.14**	0.14**	0.04**	0.04**
TELCM	1.33	0.11**	0.11**	0.17**	0.17**	0.02**	0.02**
SHOPS	0.42	0.05**	0.05**	0.24	0.24	0.06**	0.06**
HLTH	1.04	0.07**	0.07**	0.26**	0.26**	0.17**	0.18**
UTILS	0.64	0.08**	0.09**	0.29**	0.29**	0.21**	0.22**
OTHER	1.03	0.05**	0.06**	0.23**	0.24**	0.04**	0.04**

Jagannathan's (1991) SDF volatility bound. Although these issues are of interest to study further, we leave them for future research.

Appendix. Proofs

1. Proof of Proposition 1

We prove Proposition 1 in two steps. In the first step, we show that with mild assumptions, the R^2 from the predictive regression $r_{t+1}=\alpha+\beta z_t+\varepsilon_{t+1}$ is bounded above as $R^2\le\phi^2_{x,rz}\text{var}(m(x_{t+1}))$, where $m_{t+1}=m(x_{t+1})$ is a specific normalized SDF such that $\text{E}(m(x_{t+1}))=1$. In the second step, we show that the variance of any SDF can be bounded above by the variance of a CRRA utility's SDF with risk aversion γ.

Step 1. The Euler equation suggests that

$$\text{E}[m(x_{t+1})r_{t+1}|I_t] = 0. \tag{A-1}$$

Let μ_z denote the mean of predictor z. Because $z_t - \mu_z$ is in the information set I_t, we multiply equation (A-1) by $z_t - \mu_z$ on both sides and apply the law of iterated expectations to obtain

$$\mathrm{E}[m_{t+1} r_{t+1}(z_t - \mu_z)] \quad = \quad 0,$$

which can be rearranged as

$$\text{(A-2)} \qquad \mathrm{cov}(r_{t+1}, z_t) \quad = \quad -\mathrm{cov}[m_{t+1}, r_{t+1}(z_t - \mu_z)].$$

Because $\mathrm{cov}(r_{t+1}, z_t) = \mathrm{cov}(r_{t+1}, z_t - \mu_z) = \mathrm{E}[r_{t+1}(z_t - \mu_z)]$, equation (A-2) indicates that the expected excess return with $z_t - \mu_z$ units of investment in the asset r_{t+1} is equal to the negative covariance between the normalized SDF and the realized excess return of the investment, which implies that any dynamic trading strategy that exploits the predictability of r_{t+1} must be priced by the normalized SDF.

In the predictive regression (1), $\beta = \mathrm{cov}(r_{t+1}, z_t)/\mathrm{var}(z_t)$. Combining expression (2) and equation (A-2) gives

$$\text{(A-3)} \qquad \begin{aligned} R^2 \quad &= \quad \frac{\mathrm{var}(\alpha + \beta z_t)}{\mathrm{var}(r_{t+1})} \quad = \quad \frac{\beta^2 \mathrm{var}(z_t)}{\mathrm{var}(r_{t+1})} \quad = \quad \frac{\mathrm{cov}^2(r_{t+1}, z_t)}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)} \\ &= \quad \frac{\mathrm{cov}^2[m_{t+1}, r_{t+1}(z_t - \mu_z)]}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)}. \end{aligned}$$

To the best of our knowledge, Kirby (1998) is the first to test this implication by using a specific functional form of $m(\cdot)$. Instead, we derive semiparametric bounds here that do not require the functional form.

To facilitate the understanding of readers, we first consider a simple case when x_{t+1} and $r_{t+1}(z_t - \mu_z)$ are jointly normally distributed conditional on time t. From equation (A-3), we have

$$\text{(A-4)} \qquad \begin{aligned} R^2 \quad &= \quad \frac{\mathrm{cov}^2[m(x_{t+1}), r_{t+1}(z_t - \mu_z)]}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)} \\ &= \quad \frac{\left[\mathrm{cov}(x_{t+1}, r_{t+1}(z_t - \mu_z))' \mathrm{var}^{-1}(x_{t+1}) \mathrm{cov}(m_{t+1}, x_{t+1})\right]^2}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)} \\ &\leq \quad \left[\mathrm{cov}(x_{t+1}, r_{t+1}(z_t - \mu_z))' \mathrm{var}^{-1}(x_{t+1}) \mathrm{cov}(x_{t+1}, r_{t+1}(z_t - \mu_z))\right] \\ &\quad \times \frac{\left(\mathrm{cov}(m_{t+1}, x_{t+1})' \mathrm{var}^{-1}(x_{t+1}) \mathrm{cov}(m_{t+1}, x_{t+1})\right)}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)} \\ &= \quad \rho^2_{x,rz} \mathrm{var}(r_{t+1}(z_t - \mu_z)) \times \frac{\mathrm{cov}(m_{t+1}, x_{t+1})' \mathrm{var}^{-1}(x_{t+1}) \mathrm{cov}(m_{t+1}, x_{t+1})}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)} \\ &\leq \quad \rho^2_{x,rz} \frac{\mathrm{var}(r_{t+1}(z_t - \mu_z))}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)} \mathrm{var}(m_{t+1}) \\ &= \quad \phi^2_{x,rz} \mathrm{var}(m_{t+1}). \end{aligned}$$

To derive equation (A-4), we use Stein's lemma twice as[3]

$$\text{(A-5)} \qquad \mathrm{cov}(m(x_{t+1}), r_{t+1}(z_t - \mu_z)) \quad = \quad \mathrm{cov}(x_{t+1}, r_{t+1}(z_t - \mu_z))\mathrm{E}(m'(x_{t+1})),$$

$$\text{(A-6)} \qquad \mathrm{cov}(m(x_{t+1}), x_{t+1}) \quad = \quad \mathrm{var}(x_{t+1})\mathrm{E}(m'(x_{t+1})),$$

[3]We are very grateful to the referee for this and many other suggestions that improved the paper substantially.

where $m(\cdot)$ is assumed to be differentiable and $\mathrm{E}(|m'(\cdot)|)<\infty$ (see, e.g., Cochrane (2005), p. 164). A more general proof is provided in Furman and Zitikis (2008), where the assumption of the differentiability of $m(\cdot)$ can be relaxed. Inequalities in (A-4) use the Cauchy–Schwarz inequality.

Although normal distribution is generally used in the asset pricing literature, we prove that expression (A-4) can hold in more general cases when r_{t+1} and $x_{t+1}(z_t-\mu_z)$ follow a general distribution, but with an additional assumption of $\mathrm{E}(\varepsilon_{t+1}|x_{t+1})=0$, where ε_{t+1} is the residual in the orthogonal decomposition $r_{t+1}(z_t-\mu_z)=a+bx_{t+1}+\varepsilon_{t+1}$. As discussed in Kan and Zhou (2006), a sufficient condition for this assumption is that the state variables are elliptically distributed (normal distribution is a special case), which seems to fit state variables well. In fact, although technically very complex, one may expand the density function into Taylor series and plug them into the bounds. The contributions of higher moments are likely smaller than the first two moments. Because doubling the bounds does not greatly affect our empirical results, we conjecture that our bounds can be extended by relaxing the assumption. However, we retain this assumption here for simplicity.

Under the assumption $\mathrm{E}(\varepsilon_{t+1}|x_{t+1})=0$, we have

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

Hence,

$$\mathrm{cov}(r_{t+1}(z_t-\mu_z),m(x_{t+1})) \quad = \quad \mathrm{cov}[b'x_{t+1},m(x_{t+1})] \quad = \quad b'\Sigma_{xm},$$

where $\Sigma_{xm}=\mathrm{cov}(x_{t+1},m(x_{t+1}))$. Using the Cauchy–Schwarz inequality generates

$$\begin{aligned}[\mathrm{cov}(r_{t+1}(z_t-\mu_z),m(x_{t+1}))]^2 &= (b'\Sigma_{xx}^{1/2}\Sigma_{xx}^{-1/2}\Sigma_{xm})^2\\ &\le (b'\Sigma_{xx}b)(\Sigma_{xm}'\Sigma_{xx}^{-1}\Sigma_{xm}).\end{aligned} \tag{A-7}$$

Using equation (A-7), equation (A-3) can be bounded above as

$$\begin{aligned}R^2 &= \frac{\mathrm{cov}^2(m_{t+1},r_{t+1}(z_t-\mu_z))}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)}\\ &\le \frac{b'\Sigma_{xx}b}{\mathrm{var}(r_{t+1}(z_t-\mu_z))}\frac{\mathrm{var}(r_{t+1}(z_t-\mu))(\Sigma_{xm}'\Sigma_{xx}^{-1}\Sigma_{xm})}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)}\\ &\le \rho_{x,rz}^2\frac{\mathrm{var}(r_{t+1}(z_t-\mu_z))}{\mathrm{var}(r_{t+1})\mathrm{var}(z_t)}\mathrm{var}(m_{t+1})\\ &= \phi_{x,rz}^2\mathrm{var}(m_{t+1}).\end{aligned} \tag{A-8}$$

From expressions (A-4) and (A-8), we can conclude that, given that an asset pricing model can explain the degree of predictability, the predictive R^2 cannot be arbitrarily large but is bounded above by the variance of the SDF that is derived from the asset pricing model.

Step 2. We show that the variance of SDF $\mathrm{var}(m_{t+1})$ in expressions (A-4) and (A-8) can be bounded further, so that the final R^2 bound does not depend on the specification of the SDF.

Ross (2005) shows that if a utility function $U(w)$ is bounded above in the relative risk aversion by a utility function $V(w)$, that is, the risk aversion of $U(w)$ is less than or equal to that of $V(w)$, then

$$\mathrm{var}(m_U) \quad \le \quad \mathrm{var}(m_V),$$

where m_U and m_V are the corresponding SDFs. Moreover, if $V(w)$ is a CRRA utility function with risk aversion γ ($\gamma\neq 1$), the optimal wealth is the market portfolio and lognormally distributed, such as $\log w\sim\mathrm{N}[\mu(r_{\mathrm{mkt}}),\sigma^2(r_{\mathrm{mkt}})]$, then

$$\mathrm{var}(m_U) \quad \le \quad \mathrm{var}(m_V) \quad = \quad \gamma^2\sigma^2(r_{\mathrm{mkt}}). \tag{A-9}$$

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

Combining expressions (A-4), (A-8), and (A-9), if investors are bounded above by a maximum risk aversion γ, we have the R^2 bound as

$$R^2 \quad \leq \quad \bar{R}^2_{\text{RA}} \quad = \quad \phi^2_{x,rz}\gamma^2\sigma^2(r_{\text{mkt}}).$$

This completes the proof. □

2. Proof of Proposition 2

The proof of Proposition 2 consists of two steps. The first step is the same as that in the proof of Proposition 1, which shows that $R^2 \leq \rho^2_{x,rz}\text{var}(m_{t+1})$. In the second step, to make the absence of arbitrage true, we assume that constraint (9) holds, that is,

$$\text{var}(m_{t+1}) \quad \leq \quad h^2 \times \text{SR}^2(r_{\text{mkt}}).$$

Therefore, we have

$$R^2 \quad \leq \quad \bar{R}^2_{\text{SR}} \quad = \quad \phi^2_{x,rz} \times h^2 \times \text{SR}^2(r_{\text{mkt}}).$$

This completes the proof. □

3. Value of $\text{var}[r_{t+1}(z_t - \mu_z)]/\text{var}(r_{t+1})\text{var}(z_t)$

We show that the numerical value of $\text{var}[r_{t+1}(z_t - \mu_z)]/\text{var}(r_{t+1})\text{var}(z_t)$ is generally between 1 and 3. As a result, the squared multiple correlation $\rho^2_{x,rz}$ in expression (7) is the key parameter in determining the upper bound of R^2.

Without loss of generality, we assume $\mu_z = 0$ and $\text{var}(z_t) = 1$. Hence $\alpha = \text{E}(r_{t+1})$ in expression (1). With the law of total variance,

$$\begin{aligned} \text{var}(r_{t+1}z_t) &= \text{var}[\text{E}(r_{t+1}z_t|z_t)] + \text{E}[\text{var}(r_{t+1}z_t|z_t)|z_t] \\ &= \text{var}((\alpha + \beta z_t)z_t) + \text{E}(z_t^2\text{var}(r_{t+1}|z_t)). \end{aligned}$$

When z_t cannot predict r_{t+1} at all, that is, $\beta = 0$ and $\text{var}(r_{t+1}|z_t) = \text{var}(r_{t+1})$, then

$$\text{(A-10)} \qquad \frac{\text{var}(r_{t+1}z_t)}{\text{var}(r_{t+1})\text{var}(z_t)} \quad = \quad \frac{\alpha^2\text{var}(z_t) + \text{var}(r_{t+1})\text{var}(z_t)}{\text{var}(r_{t+1})\text{var}(z_t)} \quad = \quad 1 + \frac{\text{E}^2(r_{t+1})}{\text{var}(r_{t+1})}.$$

When z_t is a perfect predictor of r_{t+1}, then $\beta^2 = \text{var}(r_{t+1})$ and $\text{var}(r_{t+1}|z_t) = 0$, where we use the assumption that $\text{var}(z_t) = 1$. If z_t follows a normal distribution, we have

$$\text{(A-11)} \qquad \frac{\text{var}(r_{t+1}z_t)}{\text{var}(r_{t+1})\text{var}(z_t)} \quad = \quad \frac{\text{var}[(\alpha + \beta z_t)z_t]}{\text{var}(r_{t+1})\text{var}(z_t)} \quad = \quad 2 + \frac{\text{E}^2(r_{t+1})}{\text{var}(r_{t+1})}.$$

The second term in equations (A-10) and (A-11) is the squared Sharpe ratio of r_{t+1} and is clearly far below 1 at a monthly horizon in practice.

In general, when z_t is not a perfect predictor of r_{t+1}, based on the above analysis, the value of $\text{var}[r_{t+1}(z_t - \mu_z)]/\text{var}(r_{t+1})\text{var}(z_t)$ is between 1 and 3.

4. Effect of Measurement Errors in x on $\rho^2_{x,rz}$

We discuss how measurement errors affect $\rho^2_{x,rz}$, which is a key parameter in determining the R^2 bound. Let $\tilde{r}_{t+1} = r_{t+1}(z_t - \mu_z)$. Suppose the proxy of x_{t+1} is $\hat{x}_{t+1} = x_{t+1} + u_{t+1}$, where u_{t+1} are the measurement errors and are uncorrelated with x_{t+1} and $\tilde{r}_{t+1}$. Without loss of generality, we assume the variance–covariance matrix of u_{t+1} is invertible. For brevity,

we abbreviate the subscript $t+1$ in the sequel. The squared multiple correlation between $\hat{x}$ and $\tilde{r}$ is

$$
\begin{aligned}
\rho^2_{\hat{x},rz} &= \frac{\mathrm{cov}(\hat{x},\tilde{r})'\mathrm{var}^{-1}(\hat{x})\mathrm{cov}(\hat{x},\tilde{r})}{\mathrm{var}(\tilde{r})} \\
&= \frac{\mathrm{cov}(x,\tilde{r})'\mathrm{var}^{-1}(\hat{x})\mathrm{cov}(x,\tilde{r})}{\mathrm{var}(\tilde{r})} \\
&= \frac{\mathrm{cov}(x,\tilde{r})'\mathrm{var}^{-1}(x)\mathrm{cov}(x,\tilde{r})}{\mathrm{var}(\tilde{r})} \\
&\quad - \frac{\mathrm{cov}(x,\tilde{r})'\mathrm{var}^{-1}(x)\left[\mathrm{var}^{-1}(x)+\mathrm{var}^{-1}(u)\right]^{-1}\mathrm{var}^{-1}(x)\mathrm{cov}(x,\tilde{r})}{\mathrm{var}(\tilde{r})} \\
&\leq \rho^2_{x,rz}.
\end{aligned} \tag{A-12}
$$

Equation (A-12) uses the assumption that u are uncorrelated with $\tilde{r}$. It goes on to use the assumptions that u are uncorrelated with x and that the variance–covariance matrix var(u) is invertible, and it uses the Woodbury matrix identity that $(A+UCV)^{-1} = A^{-1} - A^{-1}U(VA^{-1}U+C^{-1})^{-1}VA^{-1}$, where A, U, C, and V all denote matrices of the correct sizes. In our case here, $A = \mathrm{var}(x)$, $C = \mathrm{var}(u)$, and U and V are identity matrices.

The inequality at the end of (A-12) uses the fact that the second term of the equation preceding it is nonnegative. With a mild assumption that var(u) is positively definite, the second term of that equation is always positive and less than $\rho^2_{x,rz}$. The more noise in the proxy of x, the larger the effect of var(u), and the closer the second term to $\rho^2_{x,rz}$. In the extreme case when each element of u has an infinite variance, $\mathrm{var}^{-1}(u)$ approaches a 0 matrix, and the second term of that equation approaches $\rho^2_{x,rz}$, leading to $\rho^2_{\hat{x},rz}$ toward 0. Therefore, the low R^2 bound may be due to the measurement errors that attenuate the measured $\rho^2_{x,rz}$.

5. Incremental Value of an Additional State Variable on $\rho^2_{x,rz}$

In Section II.C, we consider the effect of market frictions on the R^2 bound by including an additional variable, $\Delta\omega_{t+1}$, in the vector of state variables x_{t+1}. Denote by $\tilde{x}_{t+1} = (x'_{t+1}, \Delta\omega_{t+1})'$ the new state variable vector, and let $\tilde{r}_{t+1} = r_{t+1}(z_t - \mu_z)$, as in Appendix Section 4. Without loss of generality, we assume that $\Delta\omega_{t+1}$ is uncorrelated with x_{t+1}. Hence, the variance–covariance matrix of $\tilde{x}_{t+1}$ is

$$
\mathrm{var}(\tilde{x}_{t+1}) = \begin{bmatrix} \mathrm{var}(x_{t+1}) & 0 \\ 0 & \mathrm{var}(\Delta\omega_{t+1}) \end{bmatrix}.
$$

The squared multiple correlation between $\tilde{x}_{t+1}$ and $\tilde{r}_{t+1}$ is

$$
\begin{aligned}
\rho^2_{\tilde{x},rz} &= \frac{\mathrm{cov}[\tilde{x}_{t+1},\tilde{r}_{t+1}]'\mathrm{var}^{-1}(\tilde{x}_{t+1})\mathrm{cov}[\tilde{x}_{t+1},\tilde{r}_{t+1}]}{\mathrm{var}[\tilde{r}_{t+1}]} \\
&= \frac{\mathrm{cov}(x_{t+1},\tilde{r}_{t+1})'\mathrm{var}^{-1}(x_{t+1})\mathrm{cov}(x_{t+1},\tilde{r}_{t+1})}{\mathrm{var}(\tilde{r}_{t+1})} \\
&\quad + \frac{\mathrm{cov}(\Delta\omega_{t+1},\tilde{r}_{t+1})'\mathrm{var}^{-1}(\Delta\omega_{t+1})\mathrm{cov}(\Delta\omega_{t+1},\tilde{r}_{t+1})}{\mathrm{var}(\tilde{r}_{t+1})} \\
&= \rho^2_{x,rz} + \rho^2_{\Delta\omega,rz} \\
&\geq \rho^2_{x,rz}.
\end{aligned} \tag{A-13}
$$

Therefore, accounting for market frictions will improve the R^2 bound.

References

Abel, A. "Asset Prices under Habit Formation and Catching Up with the Joneses." *American Economic Review*, 80 (1990), 38–42.

Acharya, V., and L. H. Pedersen. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics*, 77 (2005), 374–410.

Adrian, T.; E. Etula; and T. Muir. "Financial Intermediaries and the Cross-Section of Asset Returns." *Journal of Finance*, 69 (2014), 2557–2596.

Baker, M., and J. Wurgler. "The Equity Share in New Issues and Aggregate Stock Returns." *Journal of Finance*, 55 (2000), 2219–2257.

Bakshi, G., and F. Chabi-Yo. "Variance Bounds on the Permanent and Transitory Components of Stochastic Discount Factors." *Journal of Financial Economics*, 105 (2012), 191–208.

Bansal, R.; D. Kiku; I. Shaliastovich; and A. Yaron. "Volatility, the Macroeconomy, and Asset Prices." *Journal of Finance*, 69 (2014), 2471–2511.

Bansal, R.; D. Kiku; and A. Yaron. "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices." *Critical Finance Review*, 1 (2012), 183–221.

Bansal, R., and A. Yaron. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, 59 (2004), 1481–1509.

Baron, M., and W. Xiong. "Credit Expansion and Neglected Crash Risk." *Quarterly Journal of Economics*, forthcoming (2017).

Barro, R. J. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics*, 121 (2006), 823–866.

Barro, R. J., and J. F. Ursúa. "Rare Macroeconomic Disasters." *Annual Review of Economics*, 4 (2012), 83–109.

Bekaert, G., and J. Liu. "Conditional Information and Variance Bounds on Pricing Kernels." *Review of Financial Studies*, 17 (2004), 339–378.

Bianchi, F.; M. Lettau; and S. C. Ludvigson. "A Markov Switching cay." Working Paper, New York University (2014).

Brunnermeier, M. K., and L. H. Pedersen. "Market Liquidity and Funding Liquidity." *Review of Financial Studies*, 22 (2009), 2201–2238.

Campbell, J. Y. "Stock Returns and Term Structure." *Journal of Financial Economics*, 18 (1987), 373–399.

Campbell, J. Y., and J. H. Cochrane. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107 (1999), 205–251.

Campbell, J. Y., and R. Shiller. "Stock Prices, Earnings, and Expected Dividends." *Journal of Finance*, 43 (1988), 661–676.

Campbell, J., and T. Vuolteenaho. "Inflation Illusion and Stock Prices." *American Economic Review*, 94 (2004), 19–23.

Cochrane, J. H. "Production-Based Asset Pricing and the Link between Stock Returns and Economic Fluctuations." *Journal of Finance*, 46 (1991), 209–237.

Cochrane, J. H. *Asset Pricing*. Princeton, NJ: Princeton University Press (2005).

Cochrane, J. H., and J. Saá-Requejo. "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets." *Journal of Political Economy*, 108 (2000), 79–119.

Constantinides, G. M., and A. Ghosh. "Asset Pricing Tests with Long-Run Risks in Consumption Growth." *Review of Asset Pricing Studies*, 1 (2011), 96–136.

Cooper, I., and R. Priestley. "Time-Varying Risk Premiums and the Output Gap." *Review of Financial Studies*, 22 (2009), 2801–2833.

Epstein, L., and S. Zin. "Substitution, Risk Aversion, and Temporal Behavior of Consumption and Asset Returns II: An Empirical Analysis." *Journal of Political Economy*, 99 (1991), 263–286.

Fama, E., and K. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22 (1988), 3–25.

Fama, E., and K. French. "Common Risk Factors in the Returns on Stock and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.

Fama, E., and G. W. Schwert. "Asset Returns and Inflation." *Journal of Financial Economics*, 5 (1977), 115–146.

Ferson, W., and C. Harvey. "The Variation in Economic Risk Premiums." *Journal of Political Economy*, 99 (1991), 385–415.

Ferson, W., and R. A. Korajczyk. "Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?" *Journal of Business*, 68 (1995), 309–349.

French, K. R.; G. W. Schwert; and R. F. Stambaugh. "Expected Stock Returns and Volatility." *Journal of Financial Economics*, 3 (1987), 3–29.

Furman, E., and R. Zitikis. "Weighted Risk Capital Allocations." *Insurance: Mathematics and Economics*, 43 (2008), 263–269.
Guiso, L.; P. Sapienza; and L. Zingales. "Time Varying Risk Aversion." *Journal of Financial Economics*, forthcoming (2017).
Han, Y.; K. Yang; and G. Zhou. "A New Anomaly: The Cross-Sectional Profitability of Technical Analysis." *Journal of Financial and Quantitative Analysis*, 48 (2013), 1433–1461.
Hansen, L. P., and R. Jagannathan. "Implications of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy*, 99 (1991), 225–262.
Hansen, L. P., and S. F. Richard. "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." *Econometrica*, 55 (1987), 587–613.
Harvey, C. R.; Y. Liu; and H. Zhou. "... and the Cross-Section of Expected Returns." *Review of Financial Studies*, 29 (2016), 5–68.
Huang, D.; F. Jiang; J. Tu; and G. Zhou. "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns." *Review of Financial Studies*, 28 (2015), 791–837.
Jones, C. S., and S. Tuzel. "New Orders and Asset Prices." *Review of Financial Studies*, 26 (2013), 115–157.
Kan, R., and G. Zhou. "A New Variance Bound on the Stochastic Discount Factor." *Journal of Business*, 79 (2006), 941–961.
Kelly, B., and S. Pruitt. "Market Expectations in the Cross-Section of Present Values." *Journal of Finance*, 68 (2013), 1721–1756.
Kirby, C. "The Restrictions on Predictability Implied by Rational Asset Pricing Models." *Review of Financial Studies*, 11 (1998), 343–382.
Kothari, S. P., and J. Shanken. "Book-to-Market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis." *Journal of Financial Economics*, 44 (1997), 169–203.
Lettau, M., and S. Ludvigson. "Consumption, Aggregate Wealth, and Expected Stock Returns." *Journal of Finance*, 56 (2001), 815–849.
Mehra, R., and E. Prescott. "The Equity Premium Puzzle." *Journal of Monetary Economics*, 15 (1985), 145–161.
Nagel, S. "Empirical Cross-Sectional Asset Pricing." *Annual Review of Financial Economics*, 5 (2013), 167–199.
Paravisini, D.; V. Rappoport; and E. Ravina. "Risk Aversion and Wealth: Evidence from Person-to-Person Lending Portfolios." *Management Science*, 63 (2017), 279–297.
Pástor, Ľ., and R. F. Stambaugh. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 111 (2003), 642–685.
Rapach, D. E.; M. C. Ringgenberg; and G. Zhou. "Short Interest and Aggregate Stock Returns." *Journal of Financial Economics*, 121 (2016), 46–65.
Rapach, D. E., and G. Zhou. "Forecasting Stock Returns." In *Handbook of Economic Forecasting, Vol. II*, G. Elliott and A. Timmermann, eds. Amsterdam: Elsevier (2013), 328–383.
Ross, S. A. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory*, 13 (1976), 341–360.
Ross, S. A. *Neoclassical Finance*. Princeton, NJ: Princeton University Press (2005).
Ross, S. A. "The Recovery Theorem." *Journal of Finance*, 70 (2015), 615–648.
Savov, A. "Asset Pricing with Garbage." *Journal of Finance*, 66 (2011), 177–201.
Wachter, J. A. "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 68 (2013), 987–1035.
Welch, I., and A. Goyal. "A Comprehensive Look at the Empirical Performance of Equity Premium Performance." *Review of Financial Studies*, 21 (2008), 1455–1508.
Zhou, G. "How Much Stock Return Predictability Can We Expect from an Asset Pricing Model?" *Economics Letters*, 108 (2010), 184–186.