

7512518-portfolio-holdings-and-investment-opportunity-set

SUMMARY: This study estimates the time series of the investment opportunity set (invop) using 13F holdings data. The invop portfolio represents the difference between the market portfolio and the invop portfolio, which is assumed to be composed of two systematic portfolios: the market portfolio and the invop portfolio. The study allows for heterogeneity in risk aversion among institutions, which results in different loadings on the two portfolios. The study estimates the shock to the invop portfolio in each quarter using the realized return of the invop portfolio and uses a rolling average of the shocks as the estimate for the set. The estimate is positively related to future market excess returns and Treasury yield, and negatively related to the variance of Treasury yield and the covariance between the market excess return and Treasury yield. The study also explores whether the estimate alone can predict future market excess returns and dividend growth rates.

METHOD: The study uses 13F holdings data to estimate the invop portfolio, which is the difference between the market portfolio and the invop portfolio. The invop portfolio's loadings on the market and invop portfolios are estimated using a rolling average of the shocks to the invop portfolio. The shocks are derived from the realized return of the invop portfolio. The study uses a numerical procedure to estimate the invop portfolio and the shocks. The invop portfolio is composed of four components: expected market excess returns, expected risk-free rates, the variance of expected risk-free rates, and the covariance between expected market excess returns and expected risk-free rates.

KEY FINDINGS:

- The estimate of the investment opportunity set is positively related to future market excess returns and Treasury yield.
- The estimate is negatively related to the variance of Treasury yield and the covariance between the market excess return and Treasury yield.
- The estimate predicts future market excess returns.
- The dividend yield, the estimate, and Treasury yield together predict future dividend growth rates.
- The estimate alone has out-of-sample predictability for future market excess returns and dividend growth rates.

IMPLICATION FOR TRADING: The estimate of the investment opportunity set can be used as a market timing strategy to improve the Sharpe ratio. Practitioners can use the estimate to identify periods of high and low investment opportunities, allowing them to adjust their portfolio allocations accordingly. Additionally, the estimate can be combined with other variables, such as the dividend yield and Treasury yield, to predict future dividend growth rates. This information can help investors make more informed investment decisions and potentially improve their returns.

Portfolio Holdings and Investment Opportunity Set

Youngmin Kim *

September 23, 2026

Abstract

I estimate the time series of investment opportunity set using the 13F holdings data. I assume that an institution's portfolio consists of two systematic portfolios: the market portfolio and the investment opportunity set mimicking (invop) portfolio. I allow heterogeneity in risk aversion among institutions, which generate their heterogeneous loadings on the two portfolios. I estimate the shock to the set in each quarter by using the realized return of the invop portfolio. I use a rolling average of the shocks as the estimate for the set. The estimate is positively related to future market excess returns and Treasury yield, and negatively to the variance of Treasury yield and the covariance between the (-) market excess return and Treasury yield. This pattern is consistent with investors' intertemporal hedging behavior. The estimate predicts future market excess returns. The dividend yield, the estimate and Treasury yield together predict future dividend growth rates. A market timing strategy based on the estimate improves the Sharpe ratio.

*Purdue University Mitch Daniels School of Business, kim4151@purdue.edu.

1 Introduction

The capital asset pricing model (CAPM) (Sharpe (1964) and Lintner (1965)) is commonly used in the finance industry. However, a big assumption of the CAPM is that investors only care about the right next return but not the ones afterwards. This behavior can be generated if their risk aversion is 1 or the investment opportunity set is constant over time. Either of the two might not be true in the real world. For example, risk averse investors such as insurance companies and pension funds could have a risk aversion greater than 1. Market risk premium is time-varying, high in bad time and low in good time. Risk-free investment opportunities are time-varying too. Treasury yields were low after the 2008 financial crisis, but it became higher since the inflation surge in 2021.

The intertemporal capital asset pricing model (ICAPM) (Merton (1973), Campbell et al. (2018)) allow the investment opportunity set to vary over time. Therefore, in the model, the investors consider all the returns over the future, not just the next one. As a result, the market return and the investment opportunity set turn out to be the factors that investors consider in their portfolio formation, and therefore the two factors determine the risk premia of risky assets. Unfortunately, the set is unobservable, and this paper tries to estimate it using the 13-F holdings data. It is worthwhile to estimate it because it summarizes the attractiveness of upcoming investment opportunities as a number.

The idea is that in the ICAPM model of this paper, the observed portfolio of an institution is a combination of the market portfolio and the investment opportunity set (invop) portfolio. Low risk aversion institutions have a positive loading on the invop portfolio, and high risk aversion ones have a negative loading. Through a numerical procedure with some assumptions, I net out the market portfolio portion and obtain the invop portfolio. I estimate the shock to the set as the realized return of the invop portfolio because its weight is the regression coefficient of the set on the excess return vector. I use a rolling average of the shocks as the estimate of the set.

In the model, the investment opportunity set consists of 4 components: (1) expected market excess returns, (2) expected risk-free rates, (3) the variance of expected risk-free rates, (4) the covariance between the expected market excess returns and the expected risk-free rates. The set is positively related to the first two and negatively to the last two. If the first two are high, it means that the future risky and risk-free investment

opportunities are attractive due to high expected returns. If the last two are high, it means that the future investment opportunities are risky, thereby decreasing the set. Note that if the risky and risk-free opportunities move in the same direction, it increases the volatility of the overall opportunities. I find this relationship in data through the regression of the realized future market excess return on the set estimate, Treasury 10-year yield, the variance of Treasury 10-year yield and the covariance between the market excess return and Treasury 10-year yield. I use the (-) realized market return as the proxy of expected future market excess returns when estimating the covariance.

I explore whether the set estimate alone can predict the future market excess return. I run the regression of future market excess return on the set estimate. I test this for the horizons from 1-quarter to 4-quarter. For the in-sample forecast regression covering from 1980 to 2025, I find that the set estimate predicts the future market excess return with an R-squared of 2.5%, 6.2%, 13.6% for 1-quarter, 2-quarter and 4-quarter, respectively. The dividend yield predicts it with an R-squared of 1.7% , 4.0% and 4.2%, respectively. The payout yield which includes dividends, repurchase and issuance predicts it with an R-squared of 1.2%, 3.2% and 6.3%. I test the out-of-sample predictability by using the period of 1980-1999 as the training sample and 2000-2025 as the test sample. The set estimate has similar out-of-sample predictability, but the dividend yield and the payout yield have weaker out-of-sample predictability.

To further validate the invop estimate, I check whether it can predict the future dividend growth rates when combined with dividend yield and Treasury 10-year yield, which is an implication of the Campbell and Shiller (1988) decomposition. I regress the sum of realized future dividend growth rates on the dividend yield, the invop estimate and Treasury 10-year yield. I test this for a horizon from 2-quarter to 20-quarter. I find that the three variables predict the sum of the future dividend growth rates with R^2 ranging from 2% to 28% with the correct signs as the Campbell and Shiller (1988) decomposition implies. This dividend predictability result strengthens the empirical relevance of the invop estimate as the estimate for future investment opportunities which include the expected market risk premiums.

To illustrate that the set estimate captures the future risky investment opportunities, I consider a market timing strategy which increases the loading when the set estimate is high and decreases it when the estimate is low. For the annual re-balancing strategy, the Sharpe ratio improves about 40% relative to the buy-and-hold strategy. This im-

provement is also obtained even if I construct the market timing strategy one quarter after the 13-F holdings date to account for the 13-F data release date. The quarterly and bi-annual re-balancing strategies also improve the Sharpe ratio.

The rest of the paper is organized as follows. In Section 2, I introduce the ICAPM model I use in this paper. In Section 3, I go over data and estimation of the investment opportunity set. In Section 4, I show that the set estimate is related to macroeconomic variables as implied by the model. I also show that the set estimate can predict future market return for many horizons and illustrate the predictability by proposing a market timing strategy based on the estimate. In Section 5, I conclude the paper.

2 Model

In this section, I introduce an ICAPM model that I apply to the 13-F holdings data. The model characterizes the optimal portfolio choice of institutions, and I assume that the institutions in 13-F data construct their portfolios as the model implies.

I assume that investors maximize the expectation on the constant relative risk aversion (CRRA) utility on their terminal portfolio value, and they have an investment horizon of h .

$$V_t = \max_{\{\alpha_t, \dots, \alpha_{t+h-1}\}} E_t \left[\frac{W_{t+h}^{1-\gamma}}{1-\gamma} \right] \quad (1)$$

They are allowed to trade every period. There is no limit in shorting and leverage. γ is the relative risk aversion parameter. Investors have a heterogeneous risk aversion. W_{t+h} is the portfolio value at $t+h$. α_t is the portfolio choice at t . There are N assets, so α_t is a $(N, 1)$ vector.

I express Eq (1) in a recursive manner.

$$V_t = \max_{\alpha_t} E_t[V_{t+1}] \quad (2)$$

$$V_{t+h-1} = \max_{\{\alpha_{t+h-1}\}} E_{t+h-1} \left[\frac{W_{t+h}^{1-\gamma}}{1-\gamma} \right] \quad (3)$$

I first derive the analytical formula for V_{t+h-1} to get an idea of the structure of the value function. Following Campbell and Viceira (1999), the log portfolio return at $t+h$, $r_{p,t+h}$

and the optimal portfolio choice at $t + h - 1$, α_{t+h-1}^* can be approximated as:

$$r_{p,t+h} = \alpha_{t+h-1}^* \left(r_{t+h} - r_{f,t+h} \mathbf{1} + \frac{1}{2} \sigma_{t+h-1}^2 \right) - \frac{1}{2} \alpha_{t+h-1}^{*,T} \Sigma_{t+h-1}^{-1} \alpha_{t+h-1}^* + r_{f,t+h} \quad (4)$$

$$\alpha_{t+h-1}^* \approx \frac{1}{\gamma} \Sigma_{t+h-1}^{-1} \mu_{t+h-1} \quad (5)$$

r_{t+h} is the $(N, 1)$ log return vector. $r_{f,t+h}$ is the log risk-free return. σ_{t+h-1}^2 is the $(N, 1)$ variance vector. μ_{t+h-1} is the expectation of the $(N, 1)$ excess return vector, R_{t+h}^e . Σ_{t+h-1} is the (N, N) covariance matrix of R_{t+h}^e .

By plugging Eq (5) to Eq (3), I obtain the following analytical formula for V_{t+h-1}

$$V_{t+h-1} = \frac{1}{\gamma} \exp \left((1 - \gamma) W_{t+h-1} + \underbrace{\left(\frac{1 - \gamma}{2\gamma} \right) \mu_{t+h-1}^T \Sigma_{t+h-1}^{-1} \mu_{t+h-1}}_{s_{t+h-1}} + (1 - \gamma) r_{f,t+h} \right) \quad (6)$$

Eq (6) implies that V_{t+h-1} is high if (1) the wealth is high, (2) the expected excess return or the Sharpe ratio of the mean-variance portfolio is high and (3) the risk-free rate is high. In other words, V_{t+h-1} is proportional to the risky (s_{t+h-1}) and risk-free ($r_{f,t+h}$) investment opportunities.

To more deeply understand the determinants of V_t , I derive the analytical formula of V_{t+h-2} with one more backward induction. First, I plug Eq (6) to Eq (2).

$$\begin{aligned} V_{t+h-2} &= \max_{\alpha_{t+h-2}} E_{t+h-2} [V_{t+h-1}] \\ &= \max_{\alpha_{t+h-2}} E_{t+h-2} \left[\frac{1}{\gamma} \exp \left((1 - \gamma) W_{t+h-1} + \left(\frac{1 - \gamma}{2\gamma} \right) s_{t+h-1} + (1 - \gamma) r_{f,t+h} \right) \right] \end{aligned} \quad (7)$$

The optimal portfolio at $t + h - 2$ turns out to be

$$\alpha_{t+h-2}^* \approx \frac{1}{\gamma} \Sigma_{t+h-2}^{-1} \left[\mu_{t+h-2} + \left(\frac{1 - \gamma}{2\gamma} \right) \phi_{s,t+h-2} + (1 - \gamma) \phi_{f,t+h-2} \right] \quad (8)$$

$\phi_{s,t+h-2} = Cov_{t+h-2} (R_{t+h-1}^e, s_{t+h-1})$, $\phi_{f,t+h-2} = Cov_{t+h-2} (R_{t+h-1}^e, r_{f,t+h})$. Investors at $t + h - 2$ consider not only the mean and variance of the excess returns but also their covariance with future investment opportunities. They consider the future opportunities

because (1) they invest not only at $t + h - 2$ but also at $t + h - 1$ and (2) their utility is based on the terminal portfolio value. They dislike the assets that covary with the future opportunities because those assets make the terminal portfolio value volatile. Therefore, they put less weight on those assets as implied by Eq (8).

I plug Eq (8) to Eq (7) and derive the analytical formula.

$$\begin{aligned}
V_{t+h-2} = & \frac{1}{\gamma} \exp \left[(1 - \gamma) W_{t+h-2} \right. \\
& + (1 - \gamma) \left(\frac{1}{2\gamma} \right) (s_{t+h-2} + E_{t+h-2}[s_{t+h-1}]) + (1 - \gamma) (r_{f,t+h-1} + E_{t+h-2}[r_{f,t+h}]) \\
& + \left(\frac{1 - \gamma}{2\gamma} \right) \frac{1 - \gamma}{(2\gamma)^2} \sigma_{s,t+h-2}^2 + (1 - \gamma) \left(\frac{1 - \gamma}{2\gamma} \right) \sigma_{f,t+h-2}^2 + (1 - \gamma) \left(\frac{1 - \gamma}{2\gamma^2} \right) \sigma_{s,f,t+h-2}^2 \\
& \left. + \frac{(\gamma - 1)^2}{2\gamma^2} \underbrace{\mu_{t+h-2}^T \Sigma_{t+h-2}^{-1} \phi_{s,t+h-2}}_{\text{risky invop rp}} + \frac{(\gamma - 1)^2}{\gamma} \underbrace{\mu_{t+h-2}^T \Sigma_{t+h-2}^{-1} \phi_{f,t+h-2}}_{\text{risk free invop rp}} \right]
\end{aligned} \tag{9}$$

I drop small magnitude terms in the exponent function of the RHS of Eq (9) to make this backward induction process tractable.¹

$$\begin{aligned}
V_{t+h-2} \approx & \frac{1}{\gamma} \exp \left[(1 - \gamma) W_{t+h-2} \right. \\
& + (1 - \gamma) \left(\frac{1}{2\gamma} \right) (s_{t+h-2} + E_{t+h-2}[s_{t+h-1}]) + (1 - \gamma) (r_{f,t+h-1} + E_{t+h-2}[r_{f,t+h}]) \\
& \left. + (1 - \gamma) \left(\frac{1 - \gamma}{2\gamma} \right) \sigma_{f,t+h-2}^2 + (1 - \gamma) \left(\frac{1 - \gamma}{2\gamma^2} \right) \sigma_{s,f,t+h-2}^2 \right]
\end{aligned} \tag{10}$$

Eq (10) implies that the investors at $t + h - 2$ who have multiple re-balancing opportunities care not only the portfolio value in the next period, but also the remaining risky, risk-free opportunities, the variance of risk-free opportunities and the covariance between the risky and risk-free opportunities. They like the high expected returns of risky and risk-free opportunities but dislike the high variance and covariance. Note that if the risky and risk-free opportunities covary, the overall opportunities become more volatile, so investors

¹I drop $\sigma_{s,t+h-2}^2$, the variance of risky opportunities because its coefficient is less than order of 1. I drop the risky investment opportunity mimicking portfolio risk premium because it is much smaller than the other terms such as expected returns of the mean-variance portfolios and γ multiplied expected risk free rates. I drop the risk-free opportunity mimicking portfolio risk premium for the same reason.

do not like the covariance.

By repeating this backward induction with keeping large enough terms, I deduce the following approximate analytical formula for V_t

$$\begin{aligned}
V_t \approx \frac{1}{\gamma} \exp \bigg[& (1 - \gamma) W_t \\
& + (1 - \gamma) \left(\frac{1}{2\gamma} \right) (s_t + E_t[s_{t+1}] + E_t[s_{t+2}] + \dots) \\
& + (1 - \gamma) (r_{f,t+1} + E_t[r_{f,t+2}] + E_t[r_{f,t+3}] + \dots) \\
& + (1 - \gamma) \left(\frac{1 - \gamma}{2\gamma} \right) (\sigma_{f,t}^2 + E_t[\sigma_{f,t+1}^2] + \dots) \\
& + (1 - \gamma) \left(\frac{1 - \gamma}{2\gamma^2} \right) (\sigma_{s,f,t}^2 + E_t[\sigma_{s,f,t+1}^2] + \dots) \bigg]
\end{aligned} \tag{11}$$

Note that γ is investor-specific. To derive a systematic factor, I replace some γ s in Eq (11) by $\bar{\gamma}$ which is a constant like 2 or 3. Then, I define the investment opportunity set, k_t as follows:

$$\begin{aligned}
k_t = & \left(\frac{1}{2\bar{\gamma}} \right) (s_t + E_t[s_{t+1}] + E_t[s_{t+2}] + \dots) + (r_{f,t+1} + E_t[r_{f,t+2}] + E_t[r_{f,t+3}] + \dots) \\
& + \left(\frac{1 - \bar{\gamma}}{2\bar{\gamma}} \right) (\sigma_{f,t}^2 + E_t[\sigma_{f,t+1}^2] + \dots) + \left(\frac{1 - \bar{\gamma}}{2\bar{\gamma}^2} \right) (\sigma_{s,f,t}^2 + E_t[\sigma_{s,f,t+1}^2] + \dots)
\end{aligned} \tag{12}$$

Then, V_t can be written in a concise way as follows.

$$V_t \approx \frac{1}{\gamma} \exp [(1 - \gamma) W_t + (1 - \gamma) k_t] \tag{13}$$

Eq (13) implies that investors' utility is determined by their wealth and the remaining opportunities. I derive the optimal portfolio at $t - 1$ as follows.

$$\alpha_{t-1}^* = \left(\frac{1}{\gamma} \right) \Sigma_{t-1}^{-1} [\mu_{t-1} + (1 - \gamma) \phi_{k,t-1}] \tag{14}$$

$$\phi_{k,t-1} = Cov_{t-1}(R_t^e, K_t).$$

Suppose there are I investors in the economy. Through market clearing, I derive the

following equation on the risk premia of risky assets.

$$\mu_{t-1} \approx \gamma_{t-1} \phi_{m,t-1} + (\gamma_{t-1} - 1) \phi_{k,t-1} \quad (15)$$

$\gamma_{t-1} = \sum_{i=1}^I \left(\frac{W_{i,t-1}}{W_{t-1}} \right) \gamma_i$ is the average risk aversion. I plug Eq (15) to Eq (14) and derive the following expression.

$$\begin{aligned} \tilde{\alpha}_{t-1}^* &= \frac{\alpha_{t-1}^*}{\mathbf{1}^T \alpha_{t-1}^*} \\ &= \left(\frac{1}{1 + b_{t-1} c_{t-1}} \right) \alpha_{m,t-1} + \left(\frac{b_{t-1}}{1 + b_{t-1} c_{t-1}} \right) \alpha_{k,t-1} \end{aligned} \quad (16)$$

$b_{t-1} = (\gamma_{t-1} - \gamma) / \gamma_{t-1}$ captures how much an investor's risk aversion is relatively smaller than the average. $\alpha_{m,t-1} = \Sigma_{t-1}^{-1} \phi_{m,t-1}$ is the market portfolio. $\alpha_{k,t-1} = \Sigma_{t-1}^{-1} \phi_{k,t-1}$ is the invop portfolio. $c_{t-1} = \mathbf{1}^T \alpha_{k,t-1}$ is the sum of the invop portfolio weights.

Eq (16) implies that an investor's risky portfolio is a combination of the market portfolio and the invop portfolio. Low risk aversion investors put a positive loading on the invop portfolio and high risk aversion investors put a negative loading. Market loading depends on both risk aversion and the sensitivity of future opportunities to current realized returns.

A key component in the invop, k_t in Eq (12) is the risky invop represented by the series of expectation on s_t series. s_t is the expected return or the squared Sharpe ratio of the mean-variance optimal portfolio. But, this portfolio is unobservable so I replace it using the risk premia equation, Eq (15).

$$\begin{aligned} s_t &= \mu_t^T \Sigma_t^{-1} \mu_t \\ &= \gamma_t \phi_{m,t}^T \Sigma_t^{-1} \mu_t + (\gamma_t - 1) \phi_{k,t}^T \Sigma_t^{-1} \mu_t \\ &\approx \gamma_t E_t[R_{m,t+1}^e] \end{aligned} \quad (17)$$

I approximate s_t by omitting the second term, the invop risk premium in Eq (17). Then, s_t is the risk aversion multiplied market risk premium. Assume that $\gamma_t \approx \bar{\gamma}$, $\forall t$. Then, I

can express k_t as follows.

$$\begin{aligned}
k_t = & \left(\frac{1}{2} \right) (E_t[R_{m,t+1}^e] + E_t[R_{m,t+2}^e] + E_t[R_{m,t+3}^e] + \dots) \\
& + (r_{f,t+1} + E_t[r_{f,t+2}] + E_t[r_{f,t+3}] + \dots) \\
& + \left(\frac{1 - \bar{\gamma}}{2\bar{\gamma}} \right) (\sigma_{f,t}^2 + E_t[\sigma_{f,t+1}^2] + \dots) + \left(\frac{1 - \bar{\gamma}}{2\bar{\gamma}^2} \right) (\sigma_{s,f,t}^2 + E_t[\sigma_{s,f,t+1}^2] + \dots)
\end{aligned} \tag{18}$$

3 Data and Estimation

In this section, I go over the data I use for empirical analysis and the estimation procedure for the investment opportunity set.

3.1 Data

U.S. Institutional investors whose portfolio include \$100 million or more of 13-F securities should report their long positions on the securities at the end of each quarter. 13-F securities include U.S.-listed equities, exchange-traded funds (ETFs), closed-end funds, certain options, and convertible debt. I obtain this data from Thomson Reuters and Factset. The data is from 1980Q1 to 2025Q4. I use Thomson Reuters' data for 1980Q1-2010Q4 and Factset's data for 2011Q1-2025Q4 because the number of institutions is greater in Thomson Reuters' data until 2010Q4. I only use the holdings on common stocks listed on NYSE, NASDAQ and AMEX. For individual stock data including return, number of outstanding shares and dividends, I obtain data from CRSP. For market return, I obtain data from Ken French's website. I obtain Treasury yields data from FRED.

3.2 Estimation

The goal is to compute k_t , the investment opportunity set at t . To do so, I aim to estimate $\alpha_{k,t-1}$ because $\alpha_{k,t-1}^T R_t$ is roughly the shock to the set at t . Imagine to regress k_t on the return vector, R_t .

$$k_t = \beta_{k,0,t-1} + \underbrace{\phi_{k,t-1}^T \Sigma_{t-1}^{-1}}_{=\alpha_{k,t-1}^T} R_t \tag{19}$$

The regression coefficient vector is the invop portfolio weight vector. Thus, its realized return is roughly the shock to the set assuming that its expected return is small. Unfortunately, $\beta_{k,0,t-1}$ is unobservable, so I use the 5-year rolling average of the shocks as the

estimate.

$$\hat{k}_t = \frac{1}{h} \sum_{\tau=t-(h-1)}^t \underbrace{\alpha_{k,\tau-1}^T R_\tau}_{=\hat{k}_\tau} \quad (20)$$

I allow heterogeneity in risk aversion among institutions, which is the source of heterogeneous portfolios among institutions according to the model. But, I take into account that the model is not a perfect description of their portfolio choice, so I view the observed portfolio as the sum of the optimal portfolio implied by the model, Eq (16) plus an idiosyncratic component. To reduce the effect of the idiosyncratic component, I use 10 (turnover quintile $\times$ market coefficient quantile) portfolios + retail portfolio to deduce the invop portfolio. I use turnover and market coefficient as criteria because both are a function of risk aversion which is the source of heterogeneity in the model. I use two criteria instead of one to make the 10 portfolios as much different as possible from each other. The retail portfolio is defined as the residual portfolio. For institution i , I compute the portfolio turnover as follows.

$$Turnover_{i,t-1} = \frac{1}{N} \sum_{n=1}^N |\alpha_{i,n,t-1}^* - \alpha_{i,n,t-2}^*| \quad (21)$$

The market coefficient is the slope coefficient ($\beta_{i,m,1,t-1}$) of the regression of an institution i 's portfolio on the market portfolio.

$$\alpha_{i,n,t-1}^* = \beta_{i,m,0,t-1} + \beta_{i,m,1,t-1} \alpha_{m,n,t-1} + u_{i,n,t} \quad (22)$$

I sum the holdings of institutions that belong to the same group of (portfolio turnover quintile, market coefficient quantile). I use the group portfolios as the base portfolios for estimation.

I re-write Eq (16) for a group g portfolio with an idiosyncratic component, $e_{g,t-1}$ because the idiosyncratic components across institutions in a group might not sum to a zero vector.

$$\tilde{\alpha}_{g,t-1}^* = \underbrace{\left(\frac{1}{1 + b_{g,t-1} c_{t-1}} \right)}_{\text{market loading}} \alpha_{m,t-1} + \underbrace{\left(\frac{b_{g,t-1}}{1 + b_{g,t-1} c_{t-1}} \right)}_{\text{invop loading}} \alpha_{k,t-1} + e_{g,t-1} \quad (23)$$

I re-arrange terms in Eq (23).

$$\alpha_{k,t-1} + \underbrace{\left(\frac{1 + b_{g,t-1}c_{t-1}}{b_{g,t-1}} \right) e_{g,t-1}}_{noise} = \underbrace{\left(\frac{1 + b_{g,t-1}c_{t-1}}{b_{g,t-1}} \right) \left[\tilde{\alpha}_{g,t-1}^* - \left(\frac{1}{1 + b_{g,t-1}c_{t-1}} \right) \alpha_{m,t-1} \right]}_{net\ market\ portion} \quad (24)$$

Eq (24) implies that the inverse invop loading multiplied net-market portion of a group portfolio is equal to the invop portfolio plus a noise. I can obtain this noisy estimate from each group. I assume that every group's idiosyncratic component has the same variance. Then, the noise variance is proportional to the inverse of the invop loading. So, I use the inverse noise variance weighted average as the estimate for the invop portfolio.

$$\hat{\alpha}_{k,t-1} = \frac{1}{G} \sum_{g=1}^G \left(\frac{b_{g,t-1}}{1 + b_{g,t-1}c_{t-1}} \right) \left[\tilde{\alpha}_{g,t-1}^* - \left(\frac{1}{1 + b_{g,t-1}c_{t-1}} \right) \alpha_{m,t-1} \right] \quad (25)$$

The remaining task is to infer $b_{g,t-1}$ and c_{t-1} . To do so, I use the information in the slope coefficient of the regression of the group portfolio on the maket portfolio.

$$\alpha_{g,n,t-1}^* = \beta_{g,m,0,t-1} + \beta_{g,m,1,t-1} \alpha_{m,n,t-1} + u_{g,n,t} \quad (26)$$

Let's imagine to regress $\alpha_{k,n,t-1}$ on $\alpha_{m,n,t-1}$

$$\alpha_{k,n,t-1} = d_{0,t-1} + d_{1,t-1} \alpha_{m,n,t-1} + u_{k,t-1} \quad (27)$$

Using Eq (23), (26) and (27), I obtain the following relationship.

$$\beta_{g,m,1,t-1} = \left(\frac{1}{1 + b_{g,t-1}c_{t-1}} \right) + d_{1,t-1} \left(\frac{b_{g,t-1}}{1 + b_{g,t-1}c_{t-1}} \right) \quad (28)$$

Eq (28) implies that once c_{t-1} and $d_{1,t-1}$ are determined, $b_{g,t-1}$ is determined. I find c_{t-1} and $d_{1,t-1}$ through a numerical procedure with some assumptions. First, I find the ratio, $d_{1,t-1}/c_{t-1}$, that maximizes the correlation among the noisy invop portfolio estimates by each group portfolio. In other words, I compute the RHS of Eq (24) for each group, in this case 11 groups. Then, I compute the (11, 11) correlation matrix and compute the average of the lower half triangle of the matrix. I choose $d_{1,t-1}/c_{t-1}$ that maximizes the average. To pin down the exact values of c_{t-1} and $d_{1,t-1}$, I make two assumptions: $c_{t-1} < 0$ and $\sigma(\hat{\alpha}_{k,t-1}) = f \times \sigma(\alpha_{m,t-1})$. Negative c_{t-1} means that the sum of the invop portfolio weights is negative. This is expected given that stock returns are usually

negatively correlated with Treasury yields because a higher yield increases the discount rate so brings down the stock price. And, negative stock return usually implies a high future expected stock returns as implied by Campbell (1991). I assume that the standard deviation of the invop portfolio weights is a constant times that of the market portfolio weights. The motivation is that the average investment horizon among investors does not change much over time (maybe 1-year +/-), so I reflect this by assuming a constant relationship between the invop portfolio size (in terms of dispersion) and the market portfolio size. In this paper, I assume $f = 4$ with the idea that the average investment horizon is 1-year, so 4 quarters.

4 Investment Opportunity Set and Return Predictability

Following the procedure mentioned in the previous section, I estimate the investment opportunity set, k_t at the end of each quarter from 1985Q2 to 2025Q4. Note that the estimate is the 5-year rolling average of the invop shocks.

Figure 1: Time Series of Investment Opportunity Set (k_t)

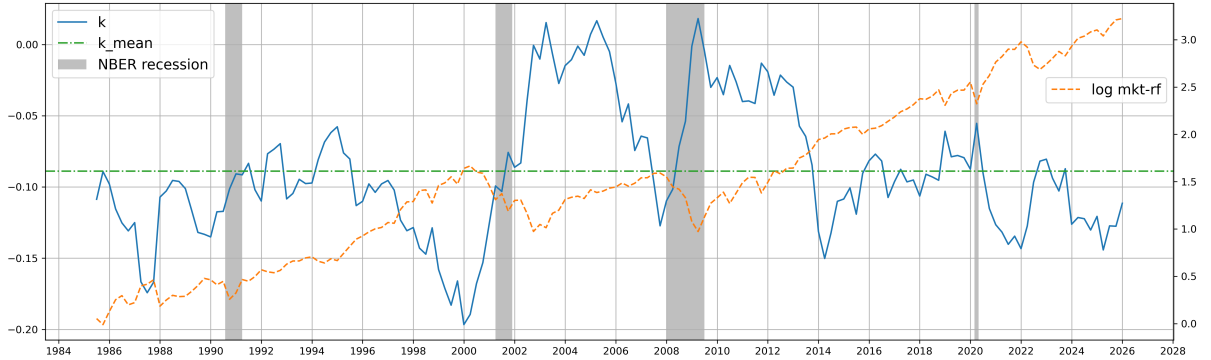


Fig 1 shows the invop and its mean on the left axis and the log cumulative market excess return on the right axis. The shaded regions are NBER recessions. The invop tends to decrease before a big stock price crash and to increase during the crash. For example, for the 2008 financial crisis, it starts to decline from 2005 and keeps declining until the price peak in 2007. Afterwards, it increases during the crash and attains its peak at the price bottom in 2009. A similar pattern is shown for other stock price crashes in 1991, 2000-2002 and 2020. This implies that the invop has a component that predicts

the future market returns, which is consistent with Eq (18).

Eq (18) tells us that the investment opportunity set has information on future market returns, future risk free rates and their covariance. I test this implication by running the regression of the realized future market return on the invop estimate and macroeconomic variables that represent the invop elements. For the expected risk-free rates, I use Treasury 10-year yield. For the variance series of the expected risk-free rates, I use rolling 12-month variance of Treasury 10-year yields. For the covariance between the expected market returns and expected risk-free rates, I use the rolling 12-month covariance between the (-) realized market return and Treasury 10-year yield. The idea is that market return is inversely correlated with the future expected market returns, as implied by Campbell (1991). To summarize, I run the following regression. Note the data frequency is quarterly.

$$R_{m,t+1,t+h}^e = \beta_0 + \beta_k k_t + \beta_f t10y + \beta_{fv} Var_t(t10y) + \beta_{mfv} Cov_t(-R_m, t10y) \quad (29)$$

Table 1: Future Market Return and Macro Variables

	$R_{m,t+1}^e$	$R_{m,t+1,t+2}^e$	$R_{m,t+1,t+4}^e$	$R_{m,t+2}^e$	$R_{m,t+2,t+3}^e$	$R_{m,t+2,t+5}^e$
	(1)	(2)	(3)	(4)	(5)	(6)
const	0.053*** (0.014)	0.102*** (0.025)	0.211*** (0.040)	0.047*** (0.013)	0.099*** (0.022)	0.210*** (0.040)
k	0.279** (0.122)	0.665*** (0.225)	1.377*** (0.384)	0.364*** (0.126)	0.747*** (0.228)	1.360*** (0.395)
t10y	-0.002 (0.003)	-0.002 (0.006)	-0.008 (0.009)	-0.000 (0.003)	-0.002 (0.005)	-0.008 (0.008)
var(t10y)	0.017 (0.029)	0.061 (0.046)	0.174** (0.067)	0.043 (0.028)	0.107** (0.047)	0.166** (0.064)
cov(- R_m , t10y)	0.452 (1.037)	1.191 (2.570)	2.680 (3.365)	0.812 (1.684)	1.839 (2.417)	3.369 (2.692)
Observations	158	158	158	158	158	158
R^2	0.029	0.078	0.196	0.051	0.120	0.195

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 1 shows the results for future 1-quarter, 2-quarter and 4-quarter returns. Stan-

dard errors are computed by the Newey and West (1987)) estimator with the maximum lag of 4. The signs of the coefficients are consistent with the model implication, Eq (18), assuming that $\bar{\gamma} > 1$. The coefficients for the invop estimate and the variance of Treasury 10-year yield are significant for most specifications. The ones of Treasury 10-year yield and the covariance are not significant. The R^2 increases as the forecast horizon increases, which is expected because the invop estimate is about all the future opportunities, not just the next quarter opportunity. This forecast power of the invop survives even if I consider the returns 1-quarter after the quarter of the invop estimate. I consider these specifications because the 13-F holdings data is released about 45 days after the holdings date. This result implies that a big portion of the information contained in the invop estimate is about the opportunities from a quarter from now.

I compare the forecasting power of the invop estimate with that of the dividend yield (Campbell and Shiller (1988)) and the payout yield (Boudoukh et al. (2007)), which are well-known market predictors. I compute payout as dividend + share repurchase - share issuance. I include all U.S. common stocks in NYSE, NASDAQ and AMEX.

Figure 2: Invop, Div Yield and Payout Yield

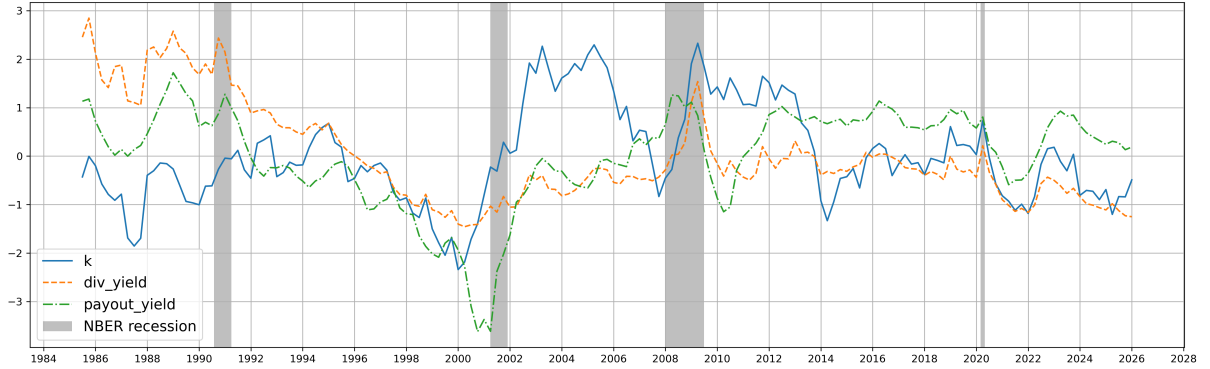


Figure 2 plots the time series of the standardized version of the three predictors. They tend to move in the same direction relative to their own recent values. For example, they tend to increase during the four recessions in the sample. They all decrease between 1994 and 2000. However, the magnitude of changes are different among the three so that the levels diverge from each other for some periods. For example, between 1984 and 1990, the invop estimate is much lower than the other two. Between 2002 and 2006, the invop estimate is much higher than the other two. This difference in level is the main source of the heterogeneous forecasting power among them.

Table 2: Future Market Return and Predictors

	$R_{m,t+1,t+4}^e$	$R_{m,t+1,t+4}^e$	$R_{m,t+1,t+4}^e$	$R_{m,t+2,t+5}^e$	$R_{m,t+2,t+5}^e$	$R_{m,t+2,t+5}^e$
	(1)	(2)	(3)	(4)	(5)	(6)
const	0.211*** (0.035)	-0.004 (0.066)	0.079*** (0.023)	0.208*** (0.036)	0.021 (0.068)	0.077*** (0.023)
k	1.346*** (0.382)			1.320*** (0.387)		
div yield		4.922* (2.843)			3.596 (2.893)	
payout yield			2.098* (1.163)			2.206** (1.101)
Observations	158	158	158	158	158	158
R^2	0.136	0.042	0.063	0.132	0.023	0.070

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 2 compares the three predictors' forecasting power on future 1-year market excess return through regression. Standard errors are computed by the Newey and West (1987) estimator with the maximum lag of 4. I consider the holding period from the next quarter and the next next quarter. For both cases, the invop estimate shows the best forecasting power. It attains an R^2 about 13%. The dividend yield and the payout yield attains an R^2 between 2% and 7%. All three predictors have a positive sign, which implies that they forecast future returns in the correct direction.

It is natural that the dividend yield has less predictive power than the invop estimate because the dividend yield is the sum of future expected market returns and future dividend growth rates. In other words, the dividend yield is a noisy predictor of future market returns. This implies that the dividend yield, the invop and Treasury yield together can predict the future dividend growths. The Campbell and Shiller (1988) decomposition shows this relationship.

$$\begin{aligned}
E_t \left[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \right] &\approx -\frac{\kappa}{1-\rho} - \log \left(\frac{D_t}{P_t} \right) + E_t \left[\sum_{j=0}^{\infty} \rho^j r_{m,t+1+j} \right] \\
&\approx -\frac{\bar{\kappa}}{1-\rho} - \frac{D_t}{P_t} + E_t \left[\sum_{j=0}^{\infty} \rho^j R_{m,t+1+j}^e \right] + E_t \left[\sum_{j=0}^{\infty} \rho^j r_{f,t+1+j} \right]
\end{aligned} \tag{30}$$

In Eq (30), I replace the log dividend yield by the dividend yield and decompose the expected market returns into expected market excess returns and expected risk free rates. To empirically test this specification, I regress the sum of future realized YoY quarterly dividend growth rates on the dividend yield, the invop estimate (k) and Treasury 10-year yield.

Table 3: Future Realized Div Growth Rates and Predictors

	$\sum_{j=0}^1 \Delta d_{t+1+j}$	$\sum_{j=0}^3 \Delta d_{t+1+j}$	$\sum_{j=0}^7 \Delta d_{t+1+j}$	$\sum_{j=0}^{11} \Delta d_{t+1+j}$	$\sum_{j=0}^{15} \Delta d_{t+1+j}$	$\sum_{j=0}^{19} \Delta d_{t+1+j}$
const	0.267*** (0.089)	0.550*** (0.172)	1.086*** (0.291)	1.526*** (0.431)	1.825*** (0.550)	2.057*** (0.542)
div yield	-6.425 (4.307)	-12.111 (8.976)	-14.499 (12.193)	-14.601 (16.434)	-15.884 (22.503)	-23.891 (24.496)
k	1.049 (0.662)	2.266* (1.216)	5.272*** (1.730)	7.233*** (2.041)	6.718*** (2.569)	3.974 (2.922)
t10y	0.016 (0.013)	0.029 (0.025)	0.034 (0.033)	0.031 (0.039)	0.017 (0.052)	0.006 (0.062)
Observations	143	143	143	143	143	143
R^2	0.080	0.117	0.224	0.278	0.227	0.137

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3 summarizes the regression results for horizons from 2-quarter to 20-quarter. For a horizon h , the Newey and West (1987) estimator with the maximum lag of h was used for standard errors. The dividend yield negatively predicts and the invop and Treasury 10-year yield positively predict the sum of future realized dividend growth rates, which is consistent with the decomposition, Eq (30). The R^2 ranges between 8% and 28%. It rises as the forecast horizon increases up to 12-quarter and then declines. This dividend predictability result further provides empirical support for the invop estimate (k) as the estimate for future investment opportunities which include the risky opportunities (i.e., expected market excess returns).

I illustrate the forecasting power of the invop estimate on future market excess returns by testing the performanc of a market-timing strategy. The strategy adjusts its loading on the market portfolio based on the market risk premium estimate. To account for the 13-F release date, I construct the portfolio after a quarter from the holdings date. I estimate the market risk premium based on the two models: (1) I use only the invop

estimate as the predictor, (2) I use the entire set of 4 predictors in Table 1. I use the period from 1980Q1 to 1999Q4 as the train sample and the period from 2000Q1 to 2025Q4 as the test sample.

Table 4: Market Timing

	$R_{m,t+2,t+5}^e$	$R_{m,t+2,t+5}^e$	$R_{m,t+2,t+5}^e$	$R_{m,t+2,t+5}^e$
	(1)	(2)	(3)	(4)
const	0.351*** (0.078)	0.416*** (0.148)	-0.051 (0.058)	-0.109* (0.061)
k	2.201*** (0.740)	2.161*** (0.692)		
t10y		-0.014 (0.016)		
var(t10y)		0.070** (0.032)		
Cov($-R_m$, t10y)		2.897 (2.030)		
RP (k)			0.715*** (0.217)	
RP (k + 3 other)				0.866*** (0.209)
Sample	Train	Train	Test	Test
Observations	54	54	99	99
R^2	0.187	0.250	0.165	0.242

Note: *p<0.1; **p<0.05; ***p<0.01

The first two columns of Table 4 shows the model estimation results on the train sample of the two models. The last two columns show the performance of the models' market risk premium estimates in predicting realized future 1-year return in the test

sample. The second model with more predictors explain the realized future return better than the first model in the train sample, as expected. The second model has a higher R^2 , and the signs of the coefficients are consistent with the ICAPM model, Eq (18). Both model estimates well predict the future returns in the test sample. They attain an R^2 of 17% and 24%, respectively. The coefficients are 0.72 and 0.87, respectively.

I consider the following market timing strategy.

$$\begin{aligned} R_{s,t+2,t+h-1}^e &= w_t R_{m,t+2,t+h-1}^e \\ w_t &= c \times \mu_{m,t} \end{aligned} \tag{31}$$

$R_{s,t+2,t+h-1}$ is the h -quarter holding strategy return (i.e., re-balancing every h quarter) covering from the beginning of period $t+2$ to the end of period $t+h-1$. I construct the portfolio at the beginning of period $t+2$, not $t+1$ to account for the 13-F release dates. w_t is the loading on the market portfolio. The loading is a constant times the market risk premium estimate which is computed using the parameters estimated on the train sample. I compute the Sharpe ratios of the two strategies during the test sample.

Table 5: Market Timing Sharpe Ratio

Portfolio Formation Timing (test sample: 2000Q1 - 2025Q4)				
	End of Q2	End of Q3	End of Q4	End of Q1
SR (buy-hold)	0.45	0.48	0.49	0.44
SR (k)	0.65	0.82	0.70	0.53
SR (k + 3 other)	0.67	0.86	0.75	0.54
Improvement (k) (%)	43%	71%	42%	21%
Improvement (k + 3 other) (%)	48%	80%	52%	24%

Table 5 compares the Sharpe ratios of the two market timing strategies and the buy-hold strategy for $h = 4$ (i.e., annual rebalancing). For robustness, I consider all four possible re-balancing timings and compute the Sharpe ratio. For all re-balancing timings, the Sharpe ratio significantly increases. In terms of percentage improvement, it ranges from 20% to 80%. The average improvement for the univariate (invop) model is 44%. The one for the multivariate (invop + 3 other) is 51%. This implies that the invop estimate contains the information on future investment opportunities which include future market risk premiums.

5 Conclusion

In this paper, I develop an estimate of the investment opportunity set. It is a number that describes the attractiveness of the future risky and risk-free investment opportunities, reflecting both expected returns and volatility. The set is an important systematic variable for portfolio choice and risky premia, as implied by many ICAPM models. However, in my knowledge, there were not many attempts to estimate it as a single number. Most of the literature decompose it to multiple elements and estimate them with a vector autoregression framework, as in Campbell et al. (2018). I fill this gap by proposing a way to estimate the set as a number using the 13-F holdings data.

The idea is that under the ICAPM model of this paper, an institution's portfolio choice is a combination of the market portfolio and the investment opportunity mimicking (invop) portfolio. Through numerical procedures with some assumptions, I extract the invop portfolio by netting out the market portfolio portion from the observed portfolio. And, I use the realized return of the invop portfolio as the shock to the set. Then, I use the 5-year rolling average of the shock as the estimate.

To validate the estimate is really about the invop, I test the main implication of the ICAPM model, which is return predictability of the invop. In the model, the invop is a sum of expected future market excess returns, expected future risk-free rates and their variance. The invop estimate can predict the future market excess returns on its own and better when combined with other predictors (Treasury 10-year yield and variance). The predictability is stronger than the one by the dividend yield and the payout yield. I also find that the invop estimate, the dividend yield and Treasury 10-year yield together can predict the sum of future realized dividend growth rates, which is consistent with the Campbell and Shiller (1988) decomposition. I also propose a market timing strategy to illustrate the invop's forecasting power on future market excess returns. The strategy significantly improves the Sharpe ratio relative to the buy-and-hold strategy.

I believe that the ICAPM framework is a strong foundation that can explain investors' portfolio choice and risky premia. It is worthwhile to develop the ICAPM framework more on both theoretical and empirical side. I hope that this paper contributes to the ICAPM literature with this portfolio holdings based estimate of the investment opportunity set.

References

- Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael Roberts. 2007. "On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing." *Journal of Finance* 62 (2).
- Campbell, John Y. 1991. "A Variance Decomposition for Stock Returns." *Economic Journal* 101 (405).
- Campbell, John Y., Stefano Giglio, Christopher Polk, and Robert Turley. 2018. "An intertemporal CAPM with stochastic volatility." *Journal of Financial Economics* 128 (2).
- Campbell, John Y. and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1 (3).
- Campbell, John Y. and Luis M. Viceira. 1999. "Consumption and Portfolio Decisions when Expected Returns are Time Varying." *Quarterly Journal of Economics* 114 (2).
- Lintner, John. 1965. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *The Review of Economics and Statistics* 47 (1).
- Merton, Robert C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5).
- Newey, Whitney K. and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3).
- Sharpe, William F. 1964. "Capital Asset Price: A Theory Of Market Equilibrium Under Conditions Of Risk." *Journal of Finance* 19 (3).