Holdings-based Fund Performance Measures: Estimation and Inference

Wayne E. Ferson
University of Southern California and NBER

Junbo L. Wang
Louisiana State University

Feb 06, 2017

PRELIMINARY and INCOMPLETE DRAFT
Abstract

Despite the increasing importance in recent research of fund holdings data and related performance measures, little is known about the statistical properties of the predominant measures. This paper introduces a predictive panel regression framework for holdings-based measures. Existing measures are treated as special cases, facilitating correction for bias, allowing access to the broad set of statistical tools for panel regressions, and providing new economic insights about the measures. Fixed effects isolate an “average alpha effect” in the measures, which is not related to predicting returns, but to the cross-section of the average stock alphas in a benchmark model. This component is large both in simulated strategies and in data on active US equity funds.

---

1 Junbo Wang may be contacted at junbowang@lsu.edu. Wayne Ferson my be contacted at ferson@usc.edu. This paper is developed from and supercedes Wang’s PhD thesis at the University of Southern California, entitled “Can weight-based measures distinguish between informed and uninformed fund managers.” We are grateful to Wayne Chang, Chris Jones, David Solomon, Scott Joslin and Jinchi Lv for comments. This paper has also benefitted from workshops at the Universite’ Laval, the Universities of Southern California, Washington, York and from presentation at the 2016 Northern Finance Association meetings.
1. Introduction

With the advent of widely available data on mutual funds’ holdings and returns, holdings-based measures of fund performance have become increasingly important and used in many studies. However, no coherent framework integrates the various measures, and the literature provides little information about their statistical properties. This paper fills these gaps.

We introduce a predictive panel regression framework to study holdings-based performance measures, where future stock returns are regressed on a fund’s lagged portfolio weights in the stocks. The intuition is that an informed manager’s portfolio weights should predict the future stock returns. Depending on the specification, the slope coefficient on the lagged weights is proportional to the portfolio change measure of Grinblatt and Titman (GT, 1989, 1993), the Characteristic Selectivity measure of Daniel, Grinblatt, Titman and Wermers (DGTW, 1997), the Conditional Weight Measure of Ferson and Khang (CWM, 2002) or the stochastic discount factor (SDF) measure of Ferson and Mo (FM, 2016). Seeing the measures as special cases of the panel regression makes available various statistical tools developed for panels and facilitates the analyses of bias and efficiency. For example, the measures are affected by a lagged stochastic regressor bias, similar to Stambaugh (1999). We adapt two alternative approaches for addressing this bias from Hjalmarsson (2008, 2010), introduce a third differenced instrumental variables approach and evaluate the corrections using simulations.

---

The predictive panel regression approach to holdings-based performance makes it easy to condition the analysis by including control variables in the regression. Given that the number of observations in the panel is roughly the number of stocks multiplied by the number of time periods, a fairly large number of control variables can be handled. We show how random firm effects in the panel can be used to model time-varying conditional alphas and discuss the role of fixed time effects.

Introducing stock fixed effects in the panel regressions removes an “average alpha effect” from the performance measures. The average alpha effect is not related to predicting returns over time, but instead measures the cross-sectional relation between funds’ time-series average portfolio weights in the stocks, and the stocks’ average alphas in the benchmark model. We find that this component often has a large influence on the measures, but its interpretation may be controversial. We describe two alternative views.

On the one hand, if the alphas of individual stocks in the benchmark model are considered to be a misspecification, then the average alpha effect in performance may be a bias to be corrected. We find that an uninformed buy-and-hold strategy often has a large average alpha effect. For example, even though the DGTW characteristic selectivity measure uses a benchmark that matches stocks on size, book/market and momentum, a buy-and-hold strategy has a large average alpha effect. A passive index fund also displays large average alpha effects.

On the other hand, if using knowledge about the cross-section of the average alphas of individual stocks relative to a model is considered to be a form of investment skill, measured performance should consider the effect. The user of the performance measure may not know about the cross-section of alphas or they may be hard to estimate with much
Mutual fund investors are found to reward funds with higher CAPM alphas via larger new money flows (Berk and van Binsbergen (2016), Barber, Huang and Odean, 2016). Some “quant” investment strategies are based on the cross-section of average alphas against benchmark models, and we find large average alpha effects for simulated momentum strategies.

So, one may view the average alpha effect as a bias or as actual skill. Our contribution is to isolate the average alpha effect without necessarily taking a stand on its merits. We find that when sorting funds on the basis of their classical holdings-based estimates of performance, much of the cross-sectional variation in the estimates is driven by the average alpha effects. We study performance in relation to several well-known fund characteristics. Funds with low factor model R-squares are found by Amihud and Goyenko (2013) to have better performance, and we confirm this with our quintile spreads. After extracting the average alpha effect, the extreme quintile differences are not significant. A similar result is found for fund volatility (Jordan and Riley, 2016) and the active weight measure of Doshi, Elkamhi and Simutin (2015). There is little impact of the average alpha effect on the spreads based on the return gap (Kacperczyk, Sialm and Zheng, 2008), as the effect does not vary much across the quintiles. While the average alpha effects are on the same order of magnitude as the fund return spreads in these examples, they bear no strong relation to the active management measures. Without the average alpha effects, the performance differences tend to be closer to zero, and fewer of the spreads achieve statistical significance.

We study the power of the various performance measures to identify informed strategies. We simulate informed strategy weights, calibrating to various levels of accuracy in the
The predictive panel framework makes it easy to address power through feasible generalized least squares, which we find improves the statistical power. (These results are in process.)

The rest of the paper is organized as follows. Section 2 introduces the holdings-based measures of performance that are the focus of the paper. Section 3 presents the predictive panel regression approach and shows how the classical measures are produced as special cases. Incorporating fixed effects in the regressions, we isolate the average alpha effects. We then expand the model, appending an autoregression for the portfolios weights, obtaining a panel version of the predictive system used by Pastor and Stambaugh (2009) to deal with lagged stochastic regressor bias. We address the bias in our context using results from Hjalmarsson (2008, 2010) and a differenced instrumental variables approach. We then discuss additional estimation issues, including ancillary model parameters, standard errors, scaling and efficiency. Section 4 describes the data. Section 5 describes the simulation methods and presents the simulation results. Section 6 presents empirical results for active US equity mutual funds. Section 7 concludes the paper and an Appendix presents ancillary results. This version of the paper is preliminary and incomplete.

2. Holdings-Based Performance Measures

Denote the portfolio weights of a fund in N stocks at time t as $w_t = [w_t^1, \ldots, w_t^N]'$, and denote the next-period returns in excess of a Treasury bill, of the stocks held by the fund as $r_{t+1} = [r_{t+1}^1, \ldots, r_{t+1}^N]'$. Holdings-based performance measures are versions of $cov(w_t, r_{t+1}) = \sum_{i=1}^N cov(w_t^i, r_{t+1}^i)$, the sum of the covariances between the current weights and the future stock returns. Grinblatt and Titman (1993) show that in a model with
normally distributed returns, an agent with nonincreasing absolute risk aversion and an informative signal about future stock returns, will display a positive holdings-based measure. It is important to sum the covariances across the assets, because an agent might overweight some stocks and underweight others to implement an informed portfolio strategy, so the covariances between subsets of the stocks’ returns and their weights might not be positive even when a fund has information about future returns.

From the definition of covariance we can write:

$$cov(w_t', r_{t+1}) = E(w_t'(r_{t+1} - E(r_{t+1}))) = E((w_t - E(w_t))'r_{t+1}).$$

Thus, holdings-based measures can be computed by demeaning the portfolio weights or the returns (or both). Versions of all three approaches appear in the literature. In place of the expected weight we typically find a benchmark weight, and in place of the expected return we typically find a benchmark return. The measures are estimated as versions of $(1/T)\sum_t w_t' r_{t+1}$, where either the weights or the returns are demeaned with a benchmark, or both are demeaned. The measures that we study in this paper are:

$$GT = \frac{1}{T-\tau} \sum_{t=\tau+1}^{T} ((w_t - w_{t-\tau})'r_{t+1}),$$

$$DGTW = \frac{1}{T} \sum_{t=1}^{T} (w_t'(r_{t+1} - r_{t+1}^{D,t})),$$

$$CWM = \frac{1}{T} \sum_{t=1}^{T} (w_t - w_{bt})'(r_{t+1} - E(r_{t+1}|Z_t)),$$

$$FM = \frac{1}{T} \sum_{t=1}^{T} w_t' r_{t+1} (a-b'r_{B(t+1)}),$$

Equation (2a) is the portfolio change measure of Grinblatt and Titman (1989,1993). This is an example of demeaning the weights. The weight of the fund $\tau$ periods before the
current period proxies for the expected weight. In this measure a manager records performance when the current portfolio $w_i'r_{t+1}$, achieves higher average hypothetical returns than the past portfolio weights, $w_{t-\tau}$, would earn on the same returns.\(^3\)

Grinblatt and Titman (1993) discuss the choice of the lag, $\tau$. If $\tau$ is too small, the past weights might still contain information about the future stock returns, leading to an underestimation of the information in the current weights. If $\tau$ is too large, the portfolio’s risk might change between the two dates and the measure, because it involves no risk adjustment, might be biased. We adopt the same criteria as Grinblatt and Titman (1993). We use $\tau = 4$ with quarterly data and $\tau = 12$ with monthly data.

Equation (2b) is the Characteristic Selectivity measure of DGTW (1997). This is an example of demeaning the stock returns, using the benchmark return vector $\mathbf{r}_{d,t} = [r_{d,t-1}, \ldots, r_{d_N,t}]$. Intuitively, a manager is informed if the stocks selected can beat their DGTW benchmark portfolio returns.\(^4\)

The DGTW benchmark return for each stock is constructed as follows. First, stocks are ranked by firm size and divided into five size groups, with each group having the same number of stocks. Within each size group the stocks are ranked by their market-to-book values, and divided into five market-to-book groups. Finally, in each of the 25 groups, the stocks are sorted by their average returns during the past months ($t-2$ to $t-12$) before the current month $t$, and split into five groups according to their past average returns. This produces 125 stock groups, each containing the same number of stocks. The value-

\(^3\) The returns are hypothetical because the weights are based on, say quarterly, snapshots of the actual fund holdings and the measured stock returns ignore all fund costs.

\(^4\) The DGTW measure is one term in a decomposition of the GT measure. The other terms refer to factor timing and average style exposure (see DGTW, 1997).
weighted returns of the stocks in each of the 125 groups become the DGTW benchmark returns. Each stock is assigned one of the 125 benchmarks based on the closest match to its size, book/market and past return characteristics.

Equation (2c) is the Ferson and Khang (2002) Conditional Weight-based Measure. This is an example of demeaning both the stock returns and the portfolio weights. The benchmark weight vector, $w_{b,t}$, is the actual weight from $\tau$ periods ago, updated with a buy-and-hold strategy: $w_{ibt} = w_{it-\tau} \prod_{j=1,\ldots,\tau} \left[ R_{i,t-\tau+j} / \sum_i w_{i,t-\tau+j-1} R_{i,t-\tau+j} \right]$, where $R_{i,t}$ is the gross (one plus the rate of) return of stock $i$ at the subscripted date. The assumption is that, under the null hypothesis of no ability, the manager is expected to use a buy-and-hold strategy. The stock returns are demeaned using $E(r_{t+1} | Z_t)$, the conditional mean returns given standard lagged instruments, $Z_t$. The conditional expected returns are estimated using regressions of returns on the lagged instruments. (We use the same lagged instruments as in Ferson and Khang in our illustrations.)

The intuition for the CWM is that a truly informed manager should depart from a buy and hold strategy when she can predict returns, over and above their predictability using public information and a simple regression. A fund delivers performance in the CWM when the portfolio’s hypothetical unexpected return (based on the public information) exceeds that of the buy-and-hold benchmark. The hypothesis that the CWM is zero assumes semi-strong form efficient markets in the sense of Fama (1970), giving managers no credit for the mechanical use of the public information in $Z_t$.

Equation (2d) is the Ferson and Mo (2016) SDF-based measure, equivalent to $\alpha_p = E(m_{t+1} r_{pt+1})$, where $r_{pt+1} = w_t' \mathbf{r}_{t+1}$ is the fund’s hypothetical portfolio excess returns. Ferson and Mo assume a linear factor model for the SDF:
where \( r_{Bt+1} \) is a vector of benchmark excess returns. Thus, the FM measure replaces \( r^i_{t+1} \) with the risk-adjusted excess stock returns, \( r^i_{t+1} (a-b'r_{Bt+1}) \). Under the assumption that the alphas \( E(m_{t+1} r^i_{t+1})=0 \), this is an example of demeaning the stock returns. A fund delivers abnormal performance in the FM measure by over-weighting stocks with subsequent high risk-adjusted returns and under-weighting those with low risk-adjusted returns. Ferson and Mo (2016) consider different choices for the benchmark returns, \( r_{Bt+1} \), including the Carhart (1997) four factor model that we use here.

2.1 Interpreting the Performance Measures

It is important to keep in mind that holdings-based performance measures are on a before-cost basis. They are designed merely to capture the information in a fund’s portfolio weights about future stock returns. They do not reflect the returns to investors, who must bear funds’ turnover-related trading costs, expense ratios and the impact of funds’ trading between reporting dates.

The sample estimates of all of the holdings-based measures can be written as versions of:

\[
(1/T) \sum_i \sum_t w^i_t r^i_{t+1},
\]

where either the weights or the returns or both are demeaned with a benchmark. depending on the specifications for the \( r^i_{t+1} \) and \( w^i_t \). In particular, using the weight change, \( w_{it} - w_{it-r} \),
in place of \( w^i \) we obtain the GT measure. Replacing the stock returns with the DGTW benchmark-adjusted returns, we obtain the DGTW measure. Replacing \( w_t \) with \((w_t - w_{bt})\) and \( r_{t+1} \) with \([r_{t+1} - E(r_{t+1}|Z_t)]\), we obtain the CWM. Finally, replacing \( r^i_t \) with \( r^i_t (a-b'r_{Bt+1}) \) we obtain the FM measure.

### 3. The Predictive Panel Regression Approach

Many previous studies have employed panel regression methods in mutual fund research. Typically, the left-hand side variables are measured at the fund level. For example, flow-performance studies going back to Sirri and Tufano (1998), regress the new money flows to mutual funds on measures of fund performance and various fund-level control variables. Several studies put measures of fund performance on the left hand side to investigate its relation to things like fund size (Ferreira, Keswanit, Miguel and Ramos, 2013), fund and industry size (Pastor, Stambaugh and Taylor, 2015) or other variables. In the latter study fixed effects are prominent as they are in our analysis, but these are fund level fixed effects. In contrast, in our approach individual stock returns are on the left hand side of the regression, and a particular fund’s holdings of the stocks are on the right hand side. Thus, we estimate a separate panel regression for each fund. While our approach is different from previous mutual fund studies, some of the same econometric issues arise, as discussed below.

We start with the simplest case of our predictive panel regression framework. Specifically, for each fund manager, assuming that the portfolio contains \( N \) stocks and exists for \( T \) periods, the regression is:
In this regression, \( r_{i+1}^t \) is the excess return of stock \( i \) at time \( t+1 \), and \( w_t^i \) is the weight of the stock held at time \( t \). The slope coefficient \( \beta \) captures the ability of a manager’s weights to predict future stock returns; thus, the estimated \( \beta \) should be positive when the manager is informed. Because the coefficient \( \beta \) is not stock-specific, the panel regression sums across the stocks in a fund’s portfolio, precisely as specified by the holdings-based measures. For illustrative purposes, we initially make the usual assumption that the regression errors are mean zero and uncorrelated with the right hand side variables.

The OLS slope coefficient estimator in the panel regression (5) is:

\[
\hat{\beta} = \frac{1}{T} \sum_i \sum_t (r_{i+1}^t w_t^i) / \left( \frac{1}{T} \sum_i \sum_t (w_t^i w_t^{i'}) \right)
\]  

Thus, the numerator is the same as Equation (4), which includes all of the holdings-based measures as special cases.

3.1 Introducing Fixed Effects

Equation (5) is unrealistic for stock returns, because under the null hypothesis that \( \beta=0 \), it implies that all the stocks have the same expected return equal to zero. We therefore introduce individual stock fixed effects, and the panel regression model becomes:

\[
r_{i+1}^t = a^i + \beta w_t^i + \varepsilon_{i+1}^t.
\]
The model is estimated by introducing a vector of $N$ stock dummy variables, which take the value of one if the return belongs to stock $i$ and zero otherwise.\(^5\) The coefficient of the dummy variable for stock $i$ is $a_i$, the fixed effect of stock $i$. Under the null hypothesis ($\beta = 0$), the fixed effect is the expected return of the stock.

With fixed effects in the model, the Frish-Waugh (1933) theorem shows that slope coefficient estimator is the same as that obtained by subtracting the time-series means from each of the left and right-hand side variables, and running the regression on the demeaned variables with no intercept. This is the “within” group estimator (where the group is the time-series of the returns for a given stock). Thus, the numerator of the OLS slope estimator of (7) is equal to:

$$\hat{\beta}_{w,\text{num}} = \Sigma_i (1/T) \Sigma_t (r_{i,t+1} - \bar{r}_i)(w_{i,t} - \bar{w}_i),$$

(8)

where $\bar{r}_i = (1/T) \Sigma_t r_{i,t+1}$ and $\bar{w}_i = (1/T) \Sigma_t w_{i,t+1}$. The within estimator captures the average across the stocks, of the time-series covariances between the weights and the future returns.

The numerator of the within estimator is an example of a holdings-based measure that

\(^5\) We also consider a panel regression with a common intercept: $r_{i,t} = a + \beta w_i + \epsilon_{i,t}$. With a common intercept in the regression, the Frish-Waugh (1933) theorem shows that slope coefficient estimator is the same as that obtained by subtracting the overall time-series and cross-sectional means from the left and right-hand side variables, and running the regression on the demeaned variables with no intercept. Thus, the numerator of the pooled OLS slope estimator is equal to:

$$\hat{\beta}_{p,\text{num}} = \Sigma_i (1/T) \Sigma_t (r_{i,t+1} - \bar{r})(w_{i,t} - \bar{w}),$$

where $\bar{w} = (1/NT) \Sigma_i \Sigma_t w_{i,t}$ and $\bar{r} = (1/NT) \Sigma_i \Sigma_t r_{i,t+1}$. We can relate the estimator with the common intercept to the estimator without an intercept using a sum of squares decomposition: $(1/T) \Sigma_i \Sigma_t r_{i,t+1} w_{i,t} = (1/T) \Sigma_i \Sigma_t (r_{i,t+1} - \bar{r})(w_{i,t} - \bar{w}) + (1/T) \Sigma_i \Sigma_t w_{i,t} \bar{r}$. Since the weights sum to 1.0 across the stocks, $\bar{w} = (1/N)$ and the last term reduces to $\bar{r}$. Thus, the numerator of the pooled OLS estimator is equal to the numerator of the baseline estimator in (6), minus the average over time of the return on an equally-weighted portfolio. While the model with a common intercept offers little economic insight, it has been suggested as a solution to finite sample bias (Hjalmarsson, 2008), so we examine it in that context below.
demeans both the returns and the weights, although it is equivalent to de-meaning only the returns or only the weights, according to (8).

3.2 Interpreting Fixed Effects

We can relate the estimator with the stock dummies to the estimator without dummies using: $\sum_i \sum_t w^i_t = \sum_i \sum_t (r^i_{t+1} - \bar{r}^i_t)(w^i_t - \bar{w}^i_t) = \sum_i \sum_t \bar{r}^i_t \bar{w}^i_t$. Thus, stock fixed effects remove from OLS slope numerator the term $(1/T) \sum_i \sum_t \bar{r}^i_t \bar{w}^i_t = \sum_i \bar{r}^i_t \bar{w}^i_t$. To interpret this, assume that stock returns follow a factor model regression on the vector of benchmark excess returns $r^B$:

$$r^i_{t+1} = \alpha_i + \beta_i \bar{r}^B_{t+1} + u^i_{t+1}. \quad (9)$$

This regression is not very restrictive, in that it allows for a stock-specific intercept or alpha. We make the standard regression assumption, $E(u^i_{t+1}) = E(r^B_{t+1} u^i_{t+1}) = 0$. Both the stock and benchmark returns are measured in excess of a short-term Treasury bill return. Note that the DGTW benchmark is a special case of the term $\beta_i \bar{r}^B_{t+1}$, where there are 125 benchmarks and the $\beta_i$ are either zero or 1.0. Note that we use the factor model here only to interpret the fixed effects. It is not used in their estimation.

Using the factor model regression Equation (9), then $\bar{r}^i_t = \alpha_i + \beta_i \bar{r}^B_t$ and:

$$\sum_i \bar{r}^i_t \bar{w}^i_t = \sum_i \alpha_i \bar{w}^i_t + (\sum_i \bar{w}^i \beta_i) \bar{r}^B_t. \quad (10)$$

---

6 Regression (4) is explicitly an unconditional regression, and its specification does not in principal preclude the existence of a conditional model with a time-varying alpha, as discussed below.
The first term of (10) is the *average alpha effect*. The second term is an average “style exposure” effect, which appears only in some versions of the regression (and when it appears we find it to be small). Thus, we see that the panel regression with stock fixed effects removes the average alpha effect from the performance measure. If the time-series sample means of either the left-hand side adjusted returns or the right-hand side adjusted weights is zero for all stocks, the Equation (10) is zero and the estimator with fixed effects is equivalent to the baseline estimator of Equation (5), and there is no average alpha effect.

Special cases of (10) apply to the different holdings-based measures that replace $w^i_t$ or $r^t_{t+1}$ with the appropriate adjusted variables. Equation (10) is zero for the CWM because the sample mean of the predictive regression residuals for each stock is zero. However, in the other classical measures the sample means of the adjusted variables are not zero, and the average alpha effect appears. In the GT measure we replace the weight $w^i_t$ with the weight change, $w^i_t - w^i_{t-1}$. Then the first term of (10) captures a cross-sectional relation between average weight changes with alphas and the second term of (10) is an average "style factor exposure change." In the DGTW measure we replace the stock returns $r^t_{t+1}$ with their benchmark adjusted returns, so $r^t$ becomes the DGTW alpha and (10) reduces to the average alpha effect. In the FM measure we replace $r^t_{t+1}$ with $r^t_{t+1} (a - b'^t r^t_{B,t+1})$, and on the assumption that the alphas of the benchmarks themselves are zero (used to identify the parameters $a$ and $b$ as discussed below), the second term of (10) is zero and we have only the average alpha effect.\(^7\)

\(^7\) The adjusted left-hand side variable for the FM example is $y^t_{it+1} = r^t_{t+1} (a - b'^t r^t_{B,t+1})$, and $E(y^t_{it+1}) = E((a + b'^t r^t_{B,t+1} + u^t_{it+1}) m^t_{it+1}) = E(m) a$, when $m$ is linear in $r^t_{B,t+1}$ and $r^t_{B,t+1}$ has zero alpha. Thus, the left-hand side of (10) becomes the mean of $m$ multiplied by the average alpha effect.
3.3 Interpreting the Average Alpha Effect

The average alpha effect is $\Sigma_i w_i^\alpha$, capturing a cross-sectional relation between the average weights and the fixed alphas. Evidently, the average alpha effect does not reflect the ability to predict returns over time. While the ability to predict future returns is usually the focus of investment performance measures -- and this is the basic idea of the original Grinblatt and Titman (1989, 1993) theory that motivates the measures -- the average alpha effect is a pure cross-sectional effect. As such, its interpretation may be debated.

One view is that the average alpha effect is simply a bias that should be removed. If the weights are properly demeaned, then all the $w_i^\alpha$ and the average alpha effects are zero. One can also view nonzero alphas in the factor model as misspecification of the factor model. Cremers, Petajisto and Zitzewitz (2012) show that passive indexes like the Russell 2000 have nonzero alphas in the standard factor models used in fund performance evaluation, and we find that a passive index fund has large, average alpha effects. We show that clearly uninformed strategies, like buy-and-hold, can have large average alpha effects. This happens because over time, buy-and-hold weights drift toward larger holdings in the stocks with higher expected returns. To the extent that the average alphas are correlated with the expected returns in the cross-section, there will be a cross-sectional relation between average portfolio weights and alphas in a buy-and-hold strategy, and thus, an average alpha effect.

Assuming constant intercepts in the regression model, it should be easy to estimate the alphas using past data on the stock and benchmark returns. In this case the weak-form efficient market hypothesis of Fama (1970) says that managers should not be credited with abnormal performance for using these alphas. A similar argument holds in a conditional
model like the CWM, replacing the fixed alphas in (4) with conditional alphas that may depend on lagged public information, $Z_t$. The argument then is that managers should get no credit for using readily available public information about the alphas, under the semi-strong form efficient markets hypothesis.\(^8\)

On the other hand, users of the performance measures might take the benchmark model as given, and not know about the cross-section of stocks’ average alphas relative to the model. Many behavioral studies provide evidence, and argue that stock market prices may reflect seemingly naïve expectations. When the benchmark models must be estimated in real time using past data, there may be a lot of uncertainty about the alphas for a future period. “Quant” portfolio strategies are often based in part on average alphas relative to benchmark models. If the user of the performance measure conditions on the benchmark model being true, then the ability to detect and trade on the average alphas relative to the model is a form of investment ability. This might be considered as a form of long-run security selection. The average alpha effect says that the fund on average, overweights the high average alpha stocks and underweights the low alpha stocks, relative to the risk model.

One can argue that those funds that understand and can use the cross-section of average alphas relative to a model, have a skill that should be measured and even rewarded. Indeed, Berk and VanBinsbergen (2016) and Barber, Huang and Odean (2016) find that mutual fund investors reward funds with positive CAPM alphas by sending them more money to manage. One of our contributions is to isolate the average alpha effect so that it

---

\(^8\) Consider the model of Christopherson et al. (1998) featuring both time-varying conditional alphas and betas: $\gamma_{t+1} = a(Z_t) + b(Z_t')\tilde{r}_{t+1} + v_{t+1}$, where $a(Z_t)$ is a time-varying conditional alpha. Then, the average alpha effect, extracted by stock level fixed effects as shown below, becomes now $\Sigma_i \left\{ (1/T)\sum_t a_i(Z_t) \right\}' \left( (1/T)\sum w_{it} \right)$ and its interpretation is essentially unchanged. The average style effect described below now includes the covariance between the time-varying betas and future benchmark returns.
may be separately evaluated and considered as a component of performance or not, depending on the preferences of the user of the performance measure.

### 3.4 Extracting the Average Alpha Effect

We argue that the difference between a classical holdings-based performance estimator and an unbiased estimate in a fixed-effects model can serve as a good proxy of the average alpha effect (plus, in the case of the GT measure, the average style effect). Suppose that we estimate the classical regression of (5) when the true model has stock fixed effects as in (7). The OLS estimator of (5) can then be written as:

\[
\hat{\beta} = \beta + \sum_i \sum_t a_i w_{it}^t / \sum_i \sum_t w_{it}^t + \sum_i \sum_t w_{it}^t \varepsilon_i / \sum_i \sum_t w_{it}^t, \tag{11}
\]

where \(\beta\) is the true value of the slope in (7) and the \(\varepsilon_i\) are the residuals of (7). Hjalmarsson's (2010) evidence and Panel B of our Table 2 below suggest that the sample mean of the term involving the residuals is close to zero. The fixed effects produce the middle term of (11), which is evaluated using the factor model by substituting \(a_i = \alpha_i + b_i' r_{t} - \beta w_i\). The expected values of the first two of these \(a_i\) terms are the average alpha effect and the average style effects described above. Denote those as \([AA+AS]\). The expression in (11) implies:

\[
E \hat{\beta} = [AA+AS] + \beta [1 - (1/T) \sum_i \sum_t \bar{w}_i / (1/T) \sum_i \sum_t w_{it}^2] \tag{12}
\]

\[
= [AA+AS] + \beta [1 - (1/T)^2 \left( \sum_i w_{it}^2 + \sum_i \sum_{i\neq t} w_{it} w_{it} \right) / (1/T) \sum_i \sum_t w_{it}^2],
\]

\[
= [AA+AS] + \beta [1 - (1/T)(1+2\rho/(1-\rho))],
\]
where the last line uses an AR(1) approximation for the weights (Equation (14) below). If the final term in (12) is close to $\beta$ it justifies using the original estimator minus the unbiased estimator for the model with fixed effects to proxy for $[AA+AS]$. As $T$ gets large the approximation is exact, but it is likely to be close in realistic finite samples. If $\rho = 0.9$ and $T = 200$, the term multiplying $\beta$ is about 0.905.

3.3 Conditioning the Model

The CWM illustrates conditioning the model. In this example the left-and side variable $r_{t+1}^i$ is replaced with $r_{t+1}^i - E(r_{t+1}^i|Z_t)$, where the conditional expectation for stock $i$ is estimated by a regression as $\delta_i^i Z_t$, where $\delta_i$ is a regression coefficient vector for stock $i$ and $Z_t$ is the lagged public information, including a constant. We can in principle estimate all of the parameters simultaneously by including the lagged variables on the right hand side of the panel regression. The modified regression is:

$$r_{t+1}^i = \delta_i^i Z_t + \beta w_t^i + \epsilon_{t+1}^i.$$  \hspace{1cm} (13)

Since $Z_t$ includes a constant, the stock fixed effects are included in the $\delta_i^i Z_t$ term, which is now an example of a random firm effect.

By the Frish-Waugh Theorem, the slope coefficient estimator for $\beta$ in the panel regression (13) is the same as what you get from first regressing the returns on the lagged $Z_t$ and the weights on $Z_t$ as seemingly-unrelated regressions and then regressing the residuals from the return regression on the residuals from the weight regression. The
residuals from the return regression are estimates of $r_{it+1}^t - \text{E}(r_{it+1}^t|Z_t)$, and the weight residuals are the time $t$ weights purged of their conditional expectations given the public information at time $t$. The coefficient numerator then precisely captures the idea of the CWM, and we use this approach.

It is common in panel regressions to include stock-specific control variables, $x_{it}$ as additional explanatory variables. Consider a regression in which $a'x_{it}$ replaces $\delta_i'Z_t$ in (13), where $a$ is an L-vector of coefficients. In a typical panel regression, the coefficients are the same for each stock. Using the Frish-Waugh theorem, it is easy to show that the numerator of the coefficient in this panel regression is equal to the numerator in the base line regression of (5), plus the panel regression coefficient vector of the $w_{it}$ regressed on the $x_{it}$, multiplied by $(1/T) \sum_i \sum_t x_{it} w_{it}$. In this sense, conditioning the regression on stock-specific variables results in a performance estimator that removes the information in the weights about the stock-specific control variables, and finds any performance from the remaining information in the weights.

One might think to condition the panel regression on time by including time dummies in the regression. By the Frish-Waugh theorem, the OLS slope coefficient in the regression with time dummies is the same as what is obtained by subtracting the cross-sectional mean values from both the left and right-hand side variables, and running the regression with the demeaned variables and no intercept. The cross-sectionally demeaned returns are simply the returns net of the return for an equally-weighted portfolio of the stocks. Since the weights in a fund’s portfolio sum to 1.0 across the stocks, the cross-sectional mean of the weights is simply $(1/N)$. Thus, the numerator of the basic panel regression is equal to the numerator of the regression with time dummies, less the time-series average return of an
equally weighted portfolio of the left hand side stock returns. The impact of time fixed
effects on the slope $\beta$ in this problem is exactly the same as the impact of including a
common intercept in the panel regression.

3.3 Standard Errors

The basic panel regression, with fixed effects or other considerations, is essentially the
model examined in detail by Petersen (2009), except that the portfolio weight predictor
variables here are endogenous and may be persistent. Petersen focusses on the panel
standard errors for OLS slope coefficient estimators under various models for the error
terms and various clustering strategies. This machinery may be applied to holdings-based
performance measures. The literature on holdings-based measures has not emphasized its
standard error estimators. DGTW use the time-series variance of the CS measures at each
period, $CS_t = (w'_t (r_{t+1} - r'_{t+1}))$, to compute standard errors. Ferson and Khang (2002) use
GMM-derived standard errors for their CWM. Seeing the measures as the result of a panel
estimation allows future research to compute reliable standard errors for holdings-based
measures using results from panel econometrics.

We defer a detailed study of standard error estimators to future work, using standard
panel methods to compute the standard errors in this study. Under the null hypothesis that
the slope is zero the serial dependence in the portfolio weights do not affect the error terms
of the panel regressions. Since stock returns are highly correlated at a point in time, but
have little time-series correlation, we cluster the standard errors by time.
3.4 Lagged Stochastic Regressor Bias

Our approach is to estimate holdings-based performance measures through predictive panel regressions. Predictive panel regressions are affected by a bias due to lagged stochastic regressors, studied by Hjalmarsson (2008, 2010). This bias, known as the “Stambaugh bias,” is related to the persistence of the predictor variable, in this case the portfolio weight, and to the correlation of its future innovations with those of the stock returns. The Stambaugh bias is examined in a time-series context by Stambaugh (1999), Pastor and Stambaugh (2009), Amihud and Hurvich (2004) and Amihud et al. (2008, 2010), among others. We examine three alternative methods to address the bias. The first two are proposed by Hjalmarsson. These include a parametric bias correction (Hjalmarsson, 2008) and a recursively demeaned instrumental variables approach, (Hjalmarsson, 2010). Our third approach is a differenced instrumental variables approach.

To capture the fact that the fund’s portfolio weights are highly autocorrelated, the parametric bias correction assumes that they follow an $AR(1)$ model:

$$w_{t+1}^i = \gamma^i + \rho w_t^i + v_{t+1}^i.$$  \hfill (14)

Equations (7) and (14) form a panel “predictive system” (e.g., Pastor and Stambaugh, 2009).

The lagged stochastic regressor bias in the panel regression with fixed effects may be understood as follows. Consider the numerator of the least squares dummy variable (the within) estimator, written with only the weights demeaned:

$$\hat{\beta}_{w,num} = (1/T)\Sigma t \Sigma i \ r_{i+1} \ (w_t^i - (1/T)\Sigma t w_t^i).$$  \hfill (15)
Substituting from Equation (7) for the returns yields:

\[
E(\hat{\beta}_{w, num} - \beta_{num}) = -E[(1/T)\sum_i \sum_{t} \epsilon_{i,t+1} (1/T)\sum_{t'=t+1} w_{i,t'}].
\] (16)

Equation (16) shows that the lagged stochastic regressor bias arises if \( \epsilon_{i,t+1} \) is correlated with \( \sum_{t'=t+1} w_{i,t'} \). The implicit demeaning of the variables, because of the intercepts in the regression, introduces the future weights. Under buy-and-hold, for example, we would expect a positive correlation between \( \epsilon_{i,t+1} \) and \( w_{i,t+1} \), as a positive return shock increases the portfolio weight, and therefore a negative bias results. The larger is the serial correlation in the weights, the larger is the expected bias as the shocks at time \( t+1 \) accumulate in the future weights.

Hjalmarsson (2008) proposes a correction for the Stambaugh bias. He assumes a local-to-unity structure for the autoregressive parameter in (14), \( \rho = 1 - c/T \). He also makes the strong assumption that the errors in (7) are independent across stocks. (We relax this assumption in our simulations.) He uses sequential limit theory (first, fix \( N \) and let \( T \) go to infinity, then let \( N \) go to infinity) to obtain several results. First, perhaps surprisingly, when there are no fixed effects in the model there is no Stambaugh bias and the pooled OLS estimator of Equation (5) is consistent. Averaging across the stocks has the effect of washing out the lagged stochastic regressor bias. However, with fixed effects the estimator in (7), with stock dummy variables, is biased as discussed above.

Hjalmarsson (2008) proposes a bias-corrected estimator for the fixed effects case. Letting \( R_{i,t+1} = r_{i,t+1} - \bar{r}_i \) and \( W_i = w_i - \bar{w}_i \), the corrected estimator is:
\[ \hat{\beta}_c = \Sigma_i \Sigma_t (R_{i,t+1} W_{it} - NT \operatorname{Cov}(\epsilon, v) \theta(c)) / \Sigma_i \Sigma_t (W_{it} W_{it}'), \quad (17) \]

where \( \operatorname{Cov}(\epsilon, v) = (1/N) \Sigma_i \operatorname{Cov}(\epsilon_{it}, v_{it}) \) is consistently estimated from the OLS residuals of the panel predictive regression system. He finds that the estimator is relatively insensitive to how the residuals are estimated. The crucial parameter is \( c \) in \( \theta(c) = -(e^c - c - 1)/c^2 \). He proposes to estimate the parameter \( c \) as \( T(1-\rho_{\text{pool}}) \), where \( \rho_{\text{pool}} \) is the pooled estimator of \( \rho \) with no intercept in the regression (14). Equation (17) is the first bias corrected estimator that we evaluate in our setting.

Hjalmarsson (2010) notes that with a common intercept in the model, the pooled OLS estimator is consistent and asymptotically normal. He also proposes a forward recursively-demeaned estimator for the fixed effects case to address its bias. Let \( w_{d,it} = w_{it} - [1/(T-t+1)] \sum_{s=t}^{T} w_{is} \) and \( r_{d,it} = r_{it} - [1/(T-t+1)] \sum_{s=t}^{T} r_{is} \) be the forward demeaned variables. The estimator is:

\[ \hat{\beta}_{id} = \Sigma_i \Sigma_t (r_{d,it+1} W_{it}) / \Sigma_i \Sigma_t (w_{id,it} W_{it}'), \quad (18) \]

Note that the original weight, without any demeaning, is used as an instrument in (18). The intuition is that the forward demeaned returns are independent of the lagged weights. The recursive demeaning induces a moving average term into the errors of the demeaned model, which should be reflected in the panel equivalent of HAC standard errors. We use this approach. While the recursive demeaning gives up some data and thus some efficiency, Hjalmarsson (2010) finds that it effectively controls the lagged stochastic regressor bias,
and our results below confirm this.

3.5 A Differenced IV Estimator

Our third approach is a differenced estimator. Take the difference between equation (7) and the same equation \( \tau \) periods before, and the fixed effects cancel out:

\[
r_{t+1}^{i} - r_{t+1-\tau}^{i} = \beta(w_{t}^{i} - w_{t-\tau}^{i}) + \epsilon_{t+1}^{i} - \epsilon_{t+1-\tau}^{i}.
\] (19)

The classical difference estimator is

\[
\hat{\beta}_{\text{diff}} = \frac{\sum_{t} \sum_{i} (w_{t}^{i} - w_{t-\tau}^{i})(r_{t+1}^{i} - r_{t+1-\tau}^{i})}{\sum_{t} \sum_{i} (w_{t}^{i} - w_{t-\tau}^{i})^{2}}.
\] (20)

We show in the Appendix that the classical difference estimator in Equation (20) suffers a Stambaugh bias for \( \tau \geq 2 \). We evaluate the classical difference estimator for \( \tau = 1 \) below. We also use the lagged weight difference as an instrumental variable, forward difference the returns and weights, and define the differenced IV estimator as:

\[
\hat{\beta}_{\text{div}} = \frac{\sum_{t} \sum_{i} (w_{t}^{i} - w_{t-\tau}^{i})(r_{t+1}^{i} - r_{t+1+\tau}^{i})}{\sum_{t} \sum_{i} (w_{t}^{i} - w_{t-\tau}^{i})(w_{t}^{i} - w_{t+\tau}^{i})}.
\] (21)

The Appendix shows that the differenced IV estimator is consistent, and we evaluate its finite sample bias below.

3.6 Scaling

The panel slope coefficients are proportional to the various holdings-based performance measures. Thus, the coefficients will be zero under the null hypothesis that
the performance measures are zero. In order to test the null hypothesis of zero performance, the panel slope coefficients may be used directly. However, the holdings-based measures are set in economically meaningful units. For example, the FM measure is a certainty equivalent excess return for the SDF used in the model, meaning that the alpha is like an extra risk-free return earned by the fund. Thus, we scale the panel slope coefficients to the scale of the original measures. This is a simple matter of multiplying the left-hand side variables by the total variance of the weights, such as \((1/T) \sum_i \sum_t w_i^t \) in the base case of Equation (7). For the within estimator we use the de-meaned weights, and the appropriate weight variances for the other cases.

3.7 Efficiency

The classical performance measures correspond to Ordinary Least Squares (OLS) estimators in the panel regression framework. However, the errors in the regression are unlikely to satisfy the standard OLS assumptions. The standard deviations vary across stocks. Weighted least squares (WLS) estimators, normalizing the returns and weights by the standard deviations of the stock returns, can improve efficiency. We also consider a feasible GLS approach based on a simple error components model.

In our WLS measure we divide the stock returns and the portfolio weights by the standard deviations of the regression error terms, and use these adjusted returns and weights to conduct OLS estimation. For simplicity we use the factor model residuals (where the DGTW benchmark adjusted returns are a special case) to proxy for the error terms in the panel regression. The Appendix A.2 provides details.

For our feasible GLS approach, the factor model regression provides a natural
factor structure that allows us to capture the covariation across stocks through the common factors. We estimate the factor model regression (4) for each stock as a seemingly unrelated regression model, assuming this gives us consistent residuals and slope coefficient estimates. We make the simplifying assumptions that the residuals in the factor model are uncorrelated across stocks and over time, but may have different variances across stocks, and use their sample variances to model the diagonal part of the covariance matrix. This allows us to model the full covariance matrix as an invertible matrix, whose inverse square root matrix can then be used to transform the left and right-hand side variables. We run OLS and the various difference estimators on the transformed variables in panel regressions. The Appendix A.2 provides details.

3.7 Ancilliary Parameter Estimation

The FM model requires the estimation of the parameters $(a,b)$, in the SDF, $a-b'\mathbf{r}_{t+1}$. These parameters are estimated in a separate step and the standard errors of the performance measures are adjusted to reflect their sampling variability. In a first step the parameters $(a,b)$ are estimated using only the benchmark excess returns and the short-term Treasury return in an exactly identified GMM system. This delivers the GMM covariance matrix for the parameter estimates, $V(a,b)$, and guarantees that the benchmark returns have zero alphas in the model. In the second step the panel regression is run, replacing the left hand side returns, $r_{it+1}$, with $y_{it+1} = r_{it+1} (a-b'\mathbf{r}_{t+1})$, plugging in the first-step estimates for $(a,b)$. We can express the panel estimator for the slope as a function of the estimates for $(a,b)$. Using the delta method we derive an approximate standard error for the panel model slope that accounts for the estimation errors in the parameters $(a,b)$. The details are
provided in the Appendix.

4. Data

We obtain quarterly holdings data for mutual funds from Thomson-Reuters. The period of the holdings data is from 1980 to 2012. There are in total 3596 equity funds in this sample. To avoid some standard biases, we employ several screening methods to filter the holdings data. We exclude data before 1984 in most of our analyses, since Fama and French (2010) show that there is a selection bias in data before 1984. Moreover, since most of the research papers using holdings data focus on US equity funds, we also exclude other types of mutual funds. Evans (2010) discusses an incubation bias in fund performance measures, and following his suggestions, we exclude observations before the reported date of fund organization or when a fund has a TNA (total net asset value) of less than 15 million dollars.

In addition to the holdings data, we acquire monthly prices and returns of individual stocks from the CRSP monthly stock file. We also obtain the delisting returns to deal with cases when firms go out of the market. The sample contains stocks with at least one month of returns from 1970 to 2012. The DGTW benchmark returns are collected from Russ Wermers’ website, and are then combined with the monthly stock file. For some of our analyses we use daily stock returns and mutual fund returns from CRSP, available starting in 1999 for the mutual funds. We also use daily data on interest rates from the Federal Reserve Data base, available from 1986.

We use holdings and stock prices to construct portfolio quarterly and monthly

---

9 We only select stocks with non missing values of the returns and DGTW benchmark returns. According to Wermers (2004), the stocks selected for the DGTW benchmarks have at least two years of data on book values, returns and market capitalization.
weights for the mutual funds. We first describe the quarterly weights. Let the holdings of
the stocks (measured as the number of shares held) and stock prices at \( t \) ’th quarter be \( \mathbf{h}_t \)
and \( \mathbf{p}_t \), where \( \mathbf{h}_t = [h_{t1}, \ldots, h_{tN}]' \), and \( \mathbf{p}_t = [p_{t1}, \ldots, p_{tN}]' \). The weights of the stocks in the
fund portfolio are \( \mathbf{w}_t = [w_{t1}, \ldots, w_{tN}]' \) where \( w_{it} = \frac{h_{it}p_{it}}{\mathbf{h}_t\mathbf{p}_t} \).

To construct monthly weights, since the holdings data have a lower frequency than
the price data (in general, the mutual funds report their holdings every three or six months,
but CRSP has price data every month), we assume that between the two consecutive
reporting dates the funds keep the same holdings. This assumes that the manager adopts a
buy-and-hold strategy during the periods when we do not observe the holdings data. The
monthly weights at month \( t \) are constructed as the product of a fund’s preceding reported
holdings and the month \( t \) prices of the held stocks. Weights constructed in this way are
used by Kacperczyk, Sialm and Zheng (2008), Busse and Tong (2012), Amihud and
Goyenko (2013) and others. We explore the impact of using quarterly versus monthly
weights on our main results in a robustness section and find little change.

Summary statistics for the data are shown in Table 1. In panel A, for each fund in
the sample we compute the time-series average of its total net assets (TNA) in millions of
dollars, the average number of stocks held, the return gap from Kacperczyk, Sialm and
Zheng (2008) and the active weight as in Doshi et al. (2013). We compute the sample
standard deviations of the reported fund returns, \( \sigma \), and the R-squares of factor model
regressions in the Carhart (1997) four-factor model. We also compute the sample
autocorrelations of the funds’ portfolio weights, \( \rho_j \), averaged across the holdings, at various
lags \( j, j=1, \ldots 5 \). The means, std, max and min are taken across the funds in the sample.
Finally, for each fund we compute the correlation between the errors of the portfolio weight autoregression and the panel regression of future stock returns on the weights, averaged across the holdings. The average correlation is denoted as Error Corr in the table, and the covariance for a particular fund $i$ serves as $\text{Cov}(\varepsilon_i, v_i)$ in the bias-adjusted estimator of Equation (17). The sample period is from 1984 to 2012, and the number of funds is 3596.

The average fund has total assets of $684$ million and holds 114 stocks. As shown by Kacperczyk, Sialm and Zheng (2008), the before expense-ratio reported returns are close to the raw returns estimated using fund weights and stock returns, so the return gaps are small on average. However, the maximum and minimum values for the return gap are very different across funds. The average first order autocorrelation of the funds’ weights is 0.94, so the weights are persistent time-series. This value is slightly below the value where the simulation evidence in Ferson, Sarkissian and Simin (2003) indicates that spurious regression becomes a concern. However, the autocorrelation is high enough to make the lagged stochastic regressor bias a concern.

As a reality check for the AR(1) model assumption on the weight process in our predictive regression system, we estimate the coefficients $\rho_j$ in the following panel regressions:

$$w_{i,t+1} = \gamma^i + \rho_j w_{i,t+1-j} + \nu_{i,t+1},$$

where $j$ runs from 1 to 5. If the weights follow an AR(1) process, we should observe that $\rho_j = \rho_j^i$. Panel A of Table 1 suggests this is a good approximation. The values of $\rho_j^i$, are close to the values of $\rho_j$ (for example, at the mean values, $\rho_1^2 = 0.89$ and $\rho_2 = 0.885$, while $\rho_3 = .839$ and $\rho_3 = .828$).

The last row of Panel A in Table 1 presents statistics for Error Corr, the sample correlations of two errors ($\varepsilon_i^i$ and $\nu_i^i$) in the predictive regression system. While the
average value is small, at -0.004, the values can be either positive or negative. The maximum and minimum values are 0.24 and -0.19, respectively, which suggests that the Stambaugh bias can change signs for different funds. In Panel C of Table 1, we compute the autocorrelation of the first differences in the weights. The first order autocorrelations are all less than zero and small. If the weights are well approximated by a first order autoregressive process, the autocorrelations of the first differences should be a small number.

5. Simulation

We first discuss the effectiveness of our adjustments for the Stambaugh bias. We then use the bias-adjusted measures to study the magnitudes of the performance measures and average alpha effects in controlled simulation environments. We examine the power of the measures, compared with classical returns-based measures. Finally we examine the impact of feasible GLS and WLS estimation on the power.

5.1. The effectiveness of Bias Adjustments

Here we evaluate methods for bias adjusted estimation, using simulations where we bootstrap versions of the predictive system of equations (7) and (14). We report results for various “true” values of the panel regression slope, $\beta$ and the autocorrelations of the portfolio weights, $\rho$. To calibrate the regression slopes we estimate the pooled panel model for $\beta$ using the differenced IV method (which, luckily turns out to have little bias) and sort funds based on the estimated slopes. We select five funds near the 5%, 10%, 50%, 90% and 95% cutoff values, and use the holdings of the randomly-selected funds to calibrate
the simulations. The differenced IV $\beta$ is taken to be the true value in calibrating these simulations. For a given value of the slope $\beta$, the intercept for each stock in Equation (7) is chosen so that the average value of its residuals is zero. Thus, the data generating process features heterogeneous fixed effects across the stocks. The parameter $\rho$ is also allowed to differ across the stocks in these experiments. For a given value of $\rho$ the intercepts in Equation (14) are fixed so that the average value of the residuals for each stock is zero.

In a given simulation experiment we generate the residuals for the equations (7) and (14) by plugging in the calibrated values of $\beta$, $\rho$ and the intercepts for that experiment. We then bootstrap the vector of residuals for all of the stocks together with the weight residual vector for a given month, selecting months randomly with replacement. We build up the simulated weight series recursively. This preserves the serial dependence of the weights, the dependence between the returns and weight innovations, which are important for the Stambaugh bias, and the dependence across the stock returns. The cross-sectional dependence may be important given that the bias adjusted estimator of Hjalmarsson (2008) assumes cross-sectional independence. The number of months for each fund in the simulations is the same as the number of months where the fund exists in the real data. When a stock held by the fund has a missing return, the weight is set to zero.

Panel A in Table 2 presents the results of the simulations to address bias in the estimators. The true values of the slopes are shown in the first column and the estimated values, averaged across 1000 simulation trials, are shown in the other columns. The regressions are scaled as described in Section 3.7, where the dependent variable is the return multiplied by the variance of the weight, so as to deliver measures equal to the numerator of the slope coefficient, and measured as an excess return in decimal fraction
per quarter. The first column shows that the median estimated measure is small, at 0.06% per quarter, but the range of $\beta$ values in the experiments cover -0.3% to almost 0.6% per quarter.

The second column of Panel A shows that the OLS estimator in Equation (6), is expected to be much larger than the slope in the model with fixed effects. The difference is due to the fixed effects which capture the average positive stock returns in our problem. The classical OLS estimator is approximately the coefficient with fixed effects plus the average alpha effects, as previously shown. For example, when the true value of the slope is large and negative, as in the first row of the table, the stock fixed effects must be large and positive to match the average returns, and this shows up as a large positive average alpha effect in the classical estimate.

The OLS estimator with a common intercept (column 3) is also larger than the true coefficient with fixed effects, but by a smaller amount than the classical OLS estimator. Including a single intercept is a step toward estimating fixed effects, but is not an adequate solution to the problem.

Column 4 of Table 2 shows results for the bias adjusted estimator from Hjalmarsson (2008). This estimator allows for fixed effects but still has a small upward biased. This estimator ignores the dependence across stocks in the regression residuals, which our simulations capture.

Columns 5 and 6 of Table 2 present results for the standard least squares dummy variable estimator and the difference estimator for the model with fixed effects. These estimators suffer from the Stambaugh bias and the bias for the slope is negative. According to Equation (14) this says that the average correlation between the stock return residuals
and the future weights is positive, leading to a negative bias. This makes sense, and is what we would expect for a buy-and-hold strategy. Thus, without bias adjustment the standard panel regression estimators with fixed effects would likely be too pessimistic about fund performance.

Columns 7 and 8 of Table 2 present results for the bias-adjusted estimators, the differenced IV estimator (Diff IV) and the Hjalmarsson (2010) estimator shown in Equation (18) (Haj2010). The differenced IV estimator performs pretty well in terms of bias. The largest bias across the five experiments is only 0.03% per month. For Haj2010 the largest bias is only 0.01% per month. Thus, the two estimators that perform the best at removing the Stambaugh bias are differenced or forward demeaned instrumental variables estimators.

Overall, Panel A of Table 2 shows that the forward demeaned estimator of Hjalmarsson (2010) and the differenced IV estimator are the least biased estimators. We therefore concentrate on using these approaches to estimate the models in the next section, where we focus on the impact of the average alpha effects on the various performance measures.

Panel B in Table 2 presents simulations which show that the original measures are close to unbiased when there are no fixed effects in the data generating model. The simulation procedure is the same as before, but we replace the stock returns with either the DGTW adjusted returns, the conditional mean adjusted returns, where the conditional mean is $\delta_i'Z_t$, or the FM benchmark adjusted return, $m_{t+1}r_{t+1}$. The key step for the simulations in Panel B is to set the firm fixed effects equal to zero in Equation (7) in the simulated returns. It is clear that each of the four measures in Panel B are close to their corresponding true values
when there are no fixed effects in the data. This is consistent with the findings of Hjalmarsson (2008, 2010).

The results in Table 2 justify using the differences between the estimated original measures and the approximately unbiased measures with fixed effects as proxies for the average alpha effect (plus any average style effect).

5.2 The importance of Average Alpha Effects in Hypothetical Strategies

Here we simulate three types of strategies for hypothetical funds: (1) a buy-and-hold strategy, (2) a momentum strategy and (3) a random selection strategy, where the manager randomly rebalances her portfolio. These simulations allow us to assess the importance of average alpha effects in strategies that are uninformed or which trade on known patterns in the model alphas, such as momentum. Later we describe simulations in which the manager has investment ability to various degrees, and we use them to evaluate the power of the methods.

5.3 Simulating stock returns in Hypothetical Strategies

The goals here are to closely approximate the true statistical distribution of the stock returns, and to avoid the normality assumption of a Monte-Carlo simulation. In each trial, we bootstrap the stock returns for the same number of months as in the real data. In the real data, most stocks have returns that exist for a number of consecutive months, and then disappear. We wish to replicate this feature in the hypothetical strategies. While most of the strategies we simulate assume that funds hold most of the stocks for a number of consecutive months, consistent with the persistent weights in the real data, some short-
lived stocks have extreme average returns and alphas with respect to the various models. If these stocks were to exist for more periods in the simulation than in the actual data, it will lead to an exaggeration of the average alpha effect in the simulations. Thus, we strive to keep the number of months during which a stock exists the same in these simulations as in the real data. Moreover, the cross-sectional covariance among different stocks should be preserved in the simulations. In order to maintain all these features, the classical bootstrap method that randomly selects the returns of stocks from their time series, as in the previous exercise, is inadequate.

We bootstrap the stock returns with the following method. Consider the DGTW model as an example. The bootstrap is based on the following decomposition:

\[ r^j_t = r^j_{DGTW,t-1} + (r^j_t - r^j_{DGTW,t-1}), \]

where the \( r^j_{DGTW,t-1} \) are the DGTW benchmark returns and the \( r^j_t - r^j_{DGTW,t-1} \) are the DGTW benchmark adjusted returns. (The approach using the other benchmarks is similar.) The first step is to construct a pool of 125 DGTW benchmark returns, \( r^j_{DGTW,t-1} \). We also create another pool of the DGTW benchmark adjusted returns, \( r^j_t - r^j_{DGTW,t-1} \), covering only those periods during which the stock return \( r^j_t \) exists. The bootstrap is conducted with the following procedure: (1) In each month, we select all the benchmark returns at a randomly selected time point from the benchmark pool. Note that we draw all 125 DGTW benchmark returns (or all the benchmark excess returns, \( r_B \)), thus preserving the cross-sectional covariance between the benchmark returns. (2) For each stock that exists at this time point in the real data, we independently bootstrap the corresponding benchmark adjusted return from the pool of adjusted returns for that stock, picking a separate time period at random for the adjusted return. This is consistent with a
regression residual being independent of the regressors in the factor model. If a stock does not exist on the (calendar) date, the adjusted return for that stock is set as missing in the second step. This approach assures that the stock returns exist during the same periods in the simulation as in the real data. The bootstrapped stock return for its nonmissing dates is the summation of its bootstrapped benchmark return and the benchmark-adjusted return. This procedure captures the correlations of the stocks only through their benchmarks. The benchmark-adjusted return captures the alpha of the stock relative to the benchmark, which enables us to examine the average alpha effects.

5.4 Simulating Holdings in Hypothetical Strategies

Under the null hypothesis of no performance, the holdings are uninformed about future stocks returns. The buy-and-hold and random selection strategies fall into this category. The momentum strategy is an example of trading based on a well-known alpha relative to (most) models, and helps us to characterize the average alpha effects that arise under such strategies. To make the strategies realistic, we need to keep the weights between 0 and 1, and we need to keep the number of stocks in the portfolios similar to those in the real data.

5.4.1 Buy-and-hold strategy weights

The weights of the buy-and-hold strategy are simulated as follows: In each trial, we randomly select 1000 stocks and assume that the fund chooses the stocks from this pool. At time $t = 120$ (which corresponds to a fund that starts in 1984 in our simulation), the fund selects stocks and equally weights them. If a stock does not exist, the fund will not
hold it. On average, there are 150 stocks in the starting portfolio (150 stocks exist in the pool at $t = 120$); thus obtaining a similar number as the number of stocks held by funds in the real data, where the average number is 143 stocks. Next, for each time point $t > 120$, some new stocks start to exist and some old stocks disappear. For the new stocks in the pool, the manager will hold the stocks, and the weights are random numbers between 0% and 5%. For stocks in the pool that disappear, the fund will not hold these stocks.\(^{10}\) For stocks that exist in the previous period and do not disappear, the fund continues to hold them without changing the holdings (the weights are changing because prices fluctuate). The simulated weights thus contain only past price information. The number of stocks in each subsequent portfolio is between 150 and 200.

The Investment Company Act of 1940 requires that the weight of each stock cannot be larger than 5% without triggering reporting requirements. As the simulation evolves there are multiple stocks with weights that could exceed 5%, so we use the following procedure to adjust the weights. We first rank the stocks by their weights in the portfolio. Then we select the stock with the highest weight and denote this weight by $w_1$ (by assumption, $w_1 > 5\%$). Next, we decrease the weight of this stock to 5%, and let the weights of lower-ranked stocks increase by $(w_1 - 5\%)/(N - 1)$, where $N$ is the number of stocks in the portfolio. Next, we select the stock with the second highest weight (denote this weight by $w_2$). If $w_2 > 5\%$, we decrease the weight of this stock to 5% and let the weights of lower-ranked stocks increase by $(w_1 - 5\%)/(N - 2)$ (there are $N - 2$ stocks with lower rank). We repeat this process until all the stocks have weights less than or equal

\(^{10}\) The delisting returns are included in the return data to avoid selection bias.
to 5%. The only exception is the rare case when there are less than 20 stocks in the portfolio. In this case, we do not reduce the weights to 5%. To remove the effect of the initial conditions, we let the funds start 20 months earlier (from $t=100$). However, the initial conditions do not change the results significantly.

5.4.2 Momentum strategy weights

Similar to the buy-and-hold strategy, the momentum strategy starts off holding all the stocks in the pool that have non-missing returns at time $t=120$. After the first month, the fund manager rebalances her portfolio. She ranks the stocks by average returns from the preceding 2 to 12 months, and removes 6% of the stocks with the lowest past returns. This is to match the average monthly turnover in the data of 6% per month\(^\text{11}\). The stocks that have been removed are replaced by the stocks with the largest average past returns. The stocks added are weighted in proportion to their average past returns. On average, the portfolios also contain 150 to 200 stocks.

5.4.3 Random selection strategy weights

The random selection strategy is simulated as follows. The manager holds an equally weighted portfolio of stocks in the first period. During all the following periods, she rebalances 6% of the portfolio randomly. Specifically, the manager randomly selects some stocks from her portfolio, and sells all the shares of these stocks. The total weight of these stocks is around 6%\(^\text{12}\). Next, she randomly selects the same number of stocks from

\(^{11}\) We would expect that the momentum manager may have a higher turnover ratio than an average fund manager. Therefore, we also simulate the momentum strategy with a 10% or 20% turnover ratio.

\(^{12}\) We also simulate the random selection strategy with higher ratios higher than 6%.
the pool, and buys an equally weighted portfolio of these stocks. The total weight of these added stocks is the same as the total weight of the stocks she sells; thus, the total weight of the portfolio is always 100\%.

5.5 Simulation Results for Hypothetical Strategies

Using the bootstrapped returns and the simulated strategy weights we estimate the various performance measures and calculate the standard errors and T-statistics in each simulation trial. We use the simulated quarterly weights, and the quarterly returns from compounding the simulated monthly returns. By simulating the returns and strategies 1000 times, we have 1000 point estimates, standard errors, and T-ratios for each case. The mean of the estimated performance measure is calculated as the sample average of the 1000 estimated performance measures.

The averages of the performance measures across the 1,000 trials are shown in the first row of each panel of Table 3. The Diff IV and Haj2010 estimators are close to zero; less than four basis points per year. Thus, any performance captured by the classical estimators is driven by the average alpha effect (and a small average style effect for the GT measures). The expected values of the classical GT and DGTW measures are positive for both the buy-and-hold and momentum strategies. The values are economically large; more than 7½ percent per year for the DGTW measure. Thus, there are large positive average alpha effects in the classical measures under both the buy-and-hold and momentum strategies.

With 1000 T-ratios, we can study the distributional properties of the T-statistics for the performance measures. If the estimated T-statistics have standard normal distributions,
we should expect that the 5%, 10%, 90% and 95% critical values of the simulated distributions should be close to the corresponding critical values of the standard normal distribution. We rank the 1000 simulated T-ratios and select the 25th, 50th, 950th and 975th values as the two-tailed 5%, 10%, 90% and 95% critical values of the distribution.

The distributions of the T-ratios in Table 3 show that the classical GT and DGTW performance measures would find significant positive performance in a large fraction of the cases (nearly 100% for DGTW). This is driven by the positive average alpha effect in the classical measures. When the average alpha effect is removed using the Diff IV or Haj2010 estimator, the distributions of the T-ratios appear slightly more peaked than a normal, with slightly thinner tails. However, the critical values of T-statistics are close to the corresponding values for the standard normal distribution. Although the classical measures are strongly influenced by the average alpha effect, the distributions of the simulated T-statistics are affected mainly in their location. The distributions of T-statistics for the bias-adjusted measures are close to a standard normal. This is found for both the buy-and-hold strategy and the momentum strategy.

The results of the random selection strategy are presented in panel C of Table 3 (forthcoming). The average alpha effects for all the performance measures are smaller than those with the other two uninformed strategies, but the value of the classical DGTW measure is still much larger than zero. The average DGTW benchmark adjusted return, taken over all stocks and all times, is positive; and thus there is a positive average alpha

---

13 We test whether the distributions of 1000 T-ratios for each performance measure are normal following the Jarque-Bera (JB 1986) test and Lilliefors (1967) test. The JB test compares the skewness and kurtosis of the distributions with those of a normal distribution. The Lilliefors test examines the difference between the CDF of the sample distribution and that of a normal distribution. The JB and Lilliefors tests show that, for most of the measures after the average alpha effect is removed, the distributions of the T-statistics are not significantly different from those of a normal.
effect in the measure.

Overall, our simulations based on the hypothetical strategies reveal the following main results. First, the average alpha effect is a substantial contributor to the classical performance measures, under either the buy-and-hold or momentum strategy. Its effect is mainly to shift the location of the estimates, but not the shape of the distribution, which approximates a normal.

5.6 Simulations to Address Power (preliminary)

We simulate informed funds to appraise the power of the performance measures. When a manager is informed, the weights should predict future stock returns. There are various components of ability, such as factor timing, volatility timing and security selection. Our simulations focus on ability in security selection for simplicity. Thus, the simulated weights of an informed manager are correlated with the future benchmark-adjusted stock returns. The benchmark-adjusted stock returns are simulated from the following dynamics:

\[ r_{i+1}^i - b_i^i r_{i+1}^B = \alpha_i^i + \beta_i s_i^i + \epsilon_{i+1}^i, \]  

(22)

where \( b_i^i r_{i+1}^B \) is the benchmark return for stock \( i \) using the linear factor model, or as a special case, the DGTW benchmark. The symbol \( s_i^i \) denotes the informed portfolio weights and the value of \( \beta \) in (22) controls the informativeness of the weights of the informed manager. The value of \( \beta \) is determined by calibrating the R-squared of the predictive panel regression (22). Theoretically, \( \beta = R \frac{\sigma_{\epsilon_i}}{\sigma_i} \), where \( R \) controls the
correlation between the weights and future stock returns. The larger is the value of $R$, the more relevant is the information in the weights about the future stock returns. We present results for a range of $R$ values. Note that there are fixed effects in the data generating process, and we evaluate the power of estimators with fixed effects, so there are no average alpha effects considered in the power analyses.

The process for the simulated weights $s_i^t$ is described as:

$$s_{t+1}^i = \gamma^i + \rho s_t^i + \zeta_{t+1}^i.$$  \hspace{1cm} (23)

The monthly $\rho$ is set to be 0.95, which determines the persistence of the weight process and is close to the average value implied by Table 1. The value of $\gamma^i$ is calibrated by running the weight autoregression on actual data. First, an initial value $s_0^i$ is randomly chosen from the benchmark-adjusted returns at time $t=1$. At each time $t \geq 2$, we bootstrap $\zeta_t^i$ independently from the pool of demeaned benchmark-adjusted stock returns at time $t+1$. In this way, the innovations for the simulated weights at time $t$ are correlated with future idiosyncratic stock returns for time $t+1$. The simulated weights are built up recursively following the dynamics of equation (23). In order to match the variances, $\zeta_t^i$ is scaled by $\sqrt{1-\rho^2}$ after it is randomly selected from the pool. Given the AR(1) process for the informed weights,

$$\text{var}(s_{t+1}^i) = \rho^2 \text{var}(s_t^i) + (1-\rho^2)\text{var}(r_{t+1}^i - b_t^i r_{t+1}^B - \frac{1}{T} \sum_t (r_{t+1}^i - b_t^i r_{t+1}^B)).$$  \hspace{1cm} (24)

Since

$$\text{var}(s_0^i) = \text{var}(r_{t+1}^i - b_t^i r_{t+1}^B - \frac{1}{T} \sum_t (r_{t+1}^i - b_t^i r_{t+1}^B)).$$  \hspace{1cm} (25)
it follows that \( \text{var}(s_i^t) \)'s are the same for all \( t \).

For a given value of R, the values of \( \sigma_i \) and \( \sigma_s \) that determine \( \beta \) in (22) are calibrated to the sample volatilities of the actual benchmark-adjusted returns and the process \( \{s_i^t\} \). Finally, the residual \( \epsilon_{i, t+1} \) in Equation (22) is independently bootstrapped from the pool of demeaned benchmark adjusted returns and scaled by \( \sqrt{1 - \beta^2} \). After simulating the adjusted returns from Equation (22), we independently simulate the benchmark returns as in the previous subsection, and the simulated stock returns are the summation of the two.

The simulated weights for informed managers are based on the simulated process \( \{s_i^t\} \). In general, the funds do not short sell stocks. Even if they do, the reported weights do not contain short positions. Therefore, we assume that the weights cannot be negative. We assume that the manager only selects the stocks with \( s_i^t > 0 \), and we set the weight for the stock \( i \) at time \( t \) to be zero when \( s_i^t \leq 0 \). In addition, since the summation of weights should be equal to one, the weight for each stock is scaled by the factor \( \sum s_i^t \).

With the simulated stock returns and informed weights, we evaluate the power of the performance measures. Power is the probability of rejecting the null hypothesis when it is false. We rank all 1000 simulated T-statistics, count the number of values that are larger than the 95% critical value from the simulated distribution of T-statistics when \( R = 0 \), and divide this number by 1000. This rejection ratio should be larger when the power is larger\(^{14}\). We report results for various values of \( R \).

\(^{14}\) The reason we do not count the number of the values that are larger than 1.96 is because the null distribution (when \( R = 0 \)) of the T-statistics may not be a standard normal (due to the bias). Therefore, its 95% critical value can be different from 1.96.
As a point of comparison we also report the power of a simple returns-based approach using the alphas in the factor model regressions. In this approach our simulated hypothetical fund returns, $\Sigma_i r_{t+1}^i s_{t}^i$, from (23) are regressed over time on the simulated benchmark returns $r_{t+1}^B$, both measured net of a short term Treasury bill return, and the regression intercepts are the alphas. Standard heteroskedasticity-consistent standard errors from those time series regressions are used to compute the t-ratios for the alphas. The power of the linear factor model is then computed in a similar way.

Table 4 shows … (forthcoming).

6. Empirical Results for Mutual Funds

In Table 5 we present the classical performance measures estimated by pooled OLS as in Equation (6), alongside estimates from the fixed effect model that extracts the average alpha effects. These are estimated using the Hajmarlsson (2010) approach and denoted by the subscript, $H$. In the Internet Appendix Table 5A we use the differenced IV approach. The results are similar except as noted. Funds are sorted into five groups by their values of the classical measures in the first column.

Panel A of Table 5 shows that the average style exposure change component of the GT measure is small; less than two basis points per month, and this component is absent from the other measures, so we may consider the difference between an original measure and the fixed effects estimate as approximately equal to the average alpha effect as explained above. The GT and GT$_H$ estimates show similar patterns across the GT

---

$15$ To calculate the average style change effect in the GT measure, we estimate the betas for each stock using daily stock returns with the carhart four factors as the benchmark, and compute $(\Sigma_i \Delta w_i^j \beta_i) r_{t+1}^B$. 

quintiles, until we find the lowest performance group with a large, negative average alpha
effect. Thus, the poorest performance under GT is driven by the average alpha effect. When
the average alpha effect is removed in the $\text{GT}_H$ measure, the lower quintile funds’
performance changes from the previous -0.29% per month to only -0.08% per month. In
contrast, the performance of the top quintile group is reduced by only 0.09% per month by
the removal of the average alpha effect. As a result, the difference between the high and
low performance quintile changes from 0.64% per month in the original measure, to 0.34%
per month using the $\text{GT}_H$ measure.

The bottom rows of each panel of Table 5 show results for a randomly-selected
index fund, the Wilshire 5000 Index Portfolio Investment Class Shares (W5000). The
classical GT measure does not like the index fund during 1997-2012, recording a
performance of -0.07% per month. This reflects a negative average alpha effect of -0.06%,
a small, positive average style effect and a small performance, net of the average alpha
effect, of -0.03%. The negative average alpha effect contrasts with what we would expect
under pure buy-and-hold according to the simulations in Table 3. This suggests that the
index fund departs from a buy-and-hold strategy to obtain a negative average alpha effect
under the GT measure.

The DGTW measure is examined in Panel B of Table 5. The difference between
the original DGTW estimate and the $\text{DGTW}_H$, is the average alpha effect, which ranges
from -0.19% to +0.25% per month across the quintiles, and dominates the differences
across the quintiles. Removing the average alpha effect, the low quintile performance
changes from -0.30% per month to -0.12% per month (-0.01% under $\text{DiffIV}$). The high
performance quintile numbers change from 0.25% to 0.00%. The difference between the
high and low-performance quintiles is thus only 0.12% per month, compared with 0.55% in the original measure. The fact that the average alpha effect dominates the variation in the DGTW measure across the quintiles is striking. Sorting funds by their classical DGTW measures is a lot like sorting them on their average alpha effects.

In the bottom row of Table 5, Panel B, the W5000 displays overall performance near zero, with a positive average alpha effect of 0.11% per month under DGTW. The average alpha effect is again smaller than what would be expected under pure buy-and-hold, as the simulations in Table 3 suggest an effect of +0.78% per month for this case. If the Hajmarlsson (2010) estimates of DGTW capture selectivity as the measure is designed to do, then selectivity for the W5000 is small and negative. DGTW characterizes the W5000 fund as having poor selectivity performance (-0.10% to -0.12% per year) masked by a small but positive average alpha effect. The sum of the two parts is only 0.01% per month in the original DGTW measure. Compared with the first panel, this illustrates that the results for a particular fund can be highly sensitive to the performance measure (GT versus DGTW), a feature also observed in the previous generation of returns-based performance measures (e.g., see Lehmann and Modest, 1987).\textsuperscript{16}

Panel C of Table 5 presents results for the conditional weight measure, CWM. The classical measure is \((1/T)\sum_i \sum_t ((r_{i,t+1} - r_{i,C,t+1})(w_{i,t} - w_{libt}))\), with the conditional mean of the stock returns as benchmark return: \(r_{i,C,t+1} = \delta_i' Z_{i,t}\), which is estimated by a regression in the first step (the CWM results are preliminary).

Like the GT measure, the CWM is not impressed by the performance of the W5000, assigning performance of -0.10% per month. Ferson and Khang (2002) also found negative

\textsuperscript{16} We examine the correlations of the various measures across individual funds in the Internet Appendix, Table A.1 (forthcoming).
CWMs for passive simulated strategies. The result is consistent with the DGTW characterization of the W5000 as showing poor selectivity offset by a positive average alpha effect. In the CWM there is no offsetting average alpha effect, so the negative performance shines through.

Panel D of Table 5 presents results for the FM measure (preliminary). The parameters in \( m_{t+1} = a-b\, r_{B,t+1} \) are estimated separately using the benchmark returns and the fitted values of \( m_{t+1} \) are plugged in as described above. These preliminary results show very large values of the FM measures at the high and low quintiles, and also large average alpha effects that vary between -1.44% and +.80% per month across the quintiles. Once again, the average alpha effect largely drives the variation in the performance across the quintiles. Removing the average alpha effects, all of the quintiles indicate negative performance, with a difference between the high and low quintile of 0.08% per month.

Like the other measures, the FM measure dislikes the W5000. The alpha is -0.38% per month. This reflects a small negative average alpha effect and otherwise negative performance, with \( FM_H \) equal to -0.29% per month (this is -0.27% under DiffIV; breaking down as negative selectivity (-0.26%), negative level (-0.35%) and small positive volatility timing, 0.13%). The negative selectivity estimates from FM and DGTW are similar.\(^{17}\)

### 6.1 Robustness

Table 5 uses monthly stock returns and monthly weights constructed from the quarterly holdings data as described above. These features raise potential concerns about

---

\(^{17}\) Poor timing as well as selectivity in passive portfolios is not surprising for unconditional performance measures (Ferson and Schadt, 1996), although the poor conditional performance according to the CWM may come as a surprise, but these results are preliminary.
the robustness of the results.

The first concern is errors or microstructure effects in the closing prices used to compute the weights and the returns. At the beginning of each month the same price appears in the weight and in the denominator of the future return calculation, so an error in the price might induce a spurious negative relation between the lagged weight and the future return. In momentum studies (e.g. Grinblatt and Titman, 1993) it is common to skip a day between the formation period and the future return calculation interval to handle such concerns. In a similar spirit we replicate Table 5 using 29-day returns on the left hand side of the regressions, skipping a day relative to the price data in the weights.

Table 5B in the Internet Appendix presents results using the 29-day returns. The holdings-based measures are not larger than in the original Table 5, in constrast to what would be expected if errors in prices caused a negative bias in the original results. The DGTW measures are actually smaller using the 29-day returns, but otherwise the results are very similar.

A second potential concern about the fund level results is the implicit buy-and-hold assumption that is applied between the quarterly reporting dates. We saw in Table 3 that a simulated buy-and-hold strategy can generate a positive average alpha effect. The concern here is that the large average alpha effects of the actual funds are partly an artifact of the buy-and-hold assumption.

Table 5C in the Internet Appendix replicates Table 5 using quarterly data for the returns and portfolio weights, this avoiding the buy-and-hold assumption. The results for the GT measure show slightly more extreme average alpha effects in both the high and low quintiles, as would be expected given the higher estimation error of the quarterly measures.
However, the differences with the Table 5 results are only .01% or .02% per month. The average alpha effect under the GT measure for the W5000 fund is more negative in the quarterly data. It changes from -0.06% to -0.13% per month under GT. The W5000 average alpha effect is also slightly smaller at +0.06% per month (versus 0.11%) under DGTW in the quarterly data. However, the DGTW measures display slightly higher average alpha effects across the quintiles in the quarterly data (about 0.01% to 0.03% higher), in contrast to what would be expected if there was a positive bias in the monthly results from the buy-and-hold assumption between quarterly reporting dates.

6.2 Performance in relation to fund characteristics

Previous studies find evidence that fund performance measures are higher for funds with more active management, as measured by higher return gaps (Kacperczyk, Sialm and Zheng, 2008) lower factor model regression R-squares (Amihud and Goyenko, 2013), larger active weights (Doshi, Elkamhi and Simutin, 2015) and lower fund return volatility (Jordan and Riley, 2016). In this section we re-examine these results in view of the average alpha effect.

In Table 6 we sort the funds by different fund characteristics and calculate the various original and Hajmarlsson (2010) measures associated with each of the groups. (The Internet Appendix Table 6A shows the similar results using DiffIV.) The sample period for return-gap and active weight is from 1980 to 2012. For R-squared and fund return volatility we start the analyses in 1999 when daily fund return data become available. The estimated measures and their T-statistics are shown in the table. To compute the T-statistics we cluster by time and apply a Newey-West (1987) estimator for panels allowing for nonzero
autocovariance up to 30 months.

Sorting on return gap in Panel A of Table 6 shows that the differences between the original measures and the Hajmarlsson (2010) estimates, approximating the average alpha effects, are fairly similar across the return gap quintiles. The average alpha effects are about the same for the extreme quintiles under GT and DGTW, and differ by relatively small amounts under the other two measures. As a result, the Low-High quintile spreads at the bottom of the panel are similar for the original estimators and for the estimators that extract the average alpha effects. This suggests that the average alpha effects to not vary much with return gap in the cross-section of mutual funds.

The remaining three panels of Table 6 record similar results for the other measures of active management. There is relatively little variation in the average alpha effects across the quintiles of funds, sorted by R-square, active weight or return volatility. The average alpha effects are large; on the same orders of magnitude as the average spread in the extreme quintile returns, but as they vary little across the quintiles, so it does not change the quintile spreads very much when the average alpha effects are removed. There is a tendency for the measures and spreads to be closer to zero when the average alpha effects are removed, and the number of quintile spread returns with t-ratios in excess of two is smaller. For example, with active weight the extreme quintile spreads are reduced to 40% of their previous magnitudes or less. We find seven cases of large t-ratios in the original measures; only two in the measures that extract the average alpha effects, out of 60 cases reported. (Note: this does not include the CWM results. We are still trouble-shooting the results for the CWM.)

7. Conclusions
In this paper, we introduce a panel predictive regression framework for holdings-based measures of portfolio performance. Previous measures from the literature appear as special cases of a panel regression, which facilitates analysis and provides new insights about the measures.

We find that there is an average alpha effect in the classical measures that is isolated using stock fixed effects. The average alpha effect is a relation across the stocks held by a fund, between the time-averaged fund portfolio weights and the average stock alphas, and it is the dominant component of the performance measures in many instances. It arises because the average alphas of the stocks are not zero, which is an unavoidable situation for any risk model in practice. Its interpretation is open to debate, and our contribution is to isolate it. The average alpha effects are economically large for the classical measures when confronted with simple strategies such as buy-and-hold, and also appear large for a passive index fund. As the average alpha effect depends on the stock-level alphas, the large effects found here underscore the importance of benchmark choice in performance measures, an issue addressed by Roll (1978), Lehman and Modest (1989) and others, and more recently by Cremers, Petajisto and Zitzewitz (2012).

While the different models we examine use different benchmarking methods, a full examination of the likely high sensitivity to the choice among literally hundreds of potential factor models is beyond the scope of our analysis.

Framing holdings based measures as panel regression estimators, we observe that the classical estimates of holdings-based performance suffer from a lagged stochastic regressor (a.k.a. “Stambaugh” 1999) bias. We provide bias-adjusted and GLS measures to improve efficiency and power, we simulate informed strategies and we study the
power of various holdings-based performance measures to detect informed strategies (in process).

The measures are applied to real mutual fund data. We find that when the classical measures are sorted from high to low, it is often the average alpha effect that dominates the differences in the performance across funds. We also examine the relation of the average alpha effect to active fund management, proxied by well-known fund characteristics such as active share, factor model R-square, active weight and fund volatility. While the average alpha effect is large relative to the performance differences associated with these variables, it bears little relation to the measures of active management.

In addition to mutual fund performance, the panel regression approach introduced in this paper can be used to study hedge fund managers or individual investors’ skills or for other purposes, as more holdings data become available.
Appendix

A.1 Difference Estimation

If we plug \( r_{t+1}^i = \alpha^i + \beta w_t^i + \epsilon_{t+1}^i \) into the numerator of the difference estimator, Equation (19) is equal to \( \beta \) plus a term whose numerator is:

\[
\sum_i \sum_t ((w_t^i - w_{t-\tau}^i)(\epsilon_{t+1}^i - \epsilon_{t+1-\tau}^i)). \tag{A.1}
\]

Using the AR(1) assumption for the weights,

\[
w_t^i = \gamma^i (1 + \sum_{j=1}^{r-1} \rho^i) + w_{t-\tau}^i \rho^\tau + \sum_{j=1}^\tau \rho^j v_{t-j}^i, \tag{A.2}
\]

and equation (A.1) becomes:

\[
(w_{t-\tau}^i (1 - \rho^\tau) - \gamma^i (1 + \sum_{j=1}^{r-1} \rho^i) - \sum_{j=1}^\tau \rho^j v_{t-j}^i)(\epsilon_{t+1}^i - \epsilon_{t+1-\tau}^i). \tag{A.3}
\]

Under the null hypothesis of no ability, the past weights and future regression residuals are uncorrelated. However, if the innovations in the weights \( v_{t+1-\tau}^i \) and the return innovations \( \epsilon_{t+1-\tau}^i \) are contemporaneously correlated, as seems highly likely, and if \( \rho \) is nonzero, the expected value of \(- \sum_{j=1}^\tau \rho^j v_{t-j}^i)(\epsilon_{t+1}^i - \epsilon_{t+1-\tau}^i)\) is not zero when \( \tau \geq 2 \). Thus, there is a Stambaugh bias in the classical difference estimator when \( \tau \geq 2 \).

We now show that the differenced IV estimator of Equation (20) is consistent. Plug in \( r_{t+1}^i = \alpha^i + \beta w_t^i + \epsilon_{t+1}^i \), the numerator of Equation (20) can be written as

\[
\sum_i \sum_t (w_t^i - w_{t-\tau}^i)(\beta (w_t^i - w_{t+\tau}^i) + (\epsilon_{t+1}^i - \epsilon_{t+1+\tau}^i)). \tag{A.4}
\]

By assumption, \( \epsilon_{t+1}^i - \epsilon_{t+1+\tau}^i \) is not correlated with \( w_t^i - w_{t+\tau}^i \) for \( \tau \geq 1 \). Therefore, equation (A.4) converges, for large \( T \), to \( \sum \beta E((w_t^i - w_{t-\tau}^i)(w_t^i - w_{t+\tau}^i)) \). Using the AR(1)
assumption for the weights we find:

\[ E((w_i' - w_{i-\tau})^T(w_i' - w_{i-\tau})) = (1 - \rho^*)^2 \text{Var}(w_i') \]  \hspace{1cm} (A.5)

The last equation depends on the stationarity assumption: \( \text{cov}(w_i', w_j') = \text{cov}(w_{i-\tau}', w_{j-\tau}') \).

With this result, A.4 converges to \( \sum_i \beta (1 - \rho^*)^2 \text{Var}(w_i') \). By similar logic, the denominator of equation (18) converges to \( \sum_i (1 - \rho^*)^2 \text{Var}(w_i') \). Therefore, when number of stocks is large, the differenced IV estimator converges to \( \beta \).

A.2 Weighted Least Squares Estimation

In general, WLS divides both the left and right-hand side variables by the volatility of the regression residuals to obtain observations with similar variances, and thus improves the efficiency of OLS on the transformed variables. The WLS measure here is defined for the differenced IV case as:

\[
\frac{1}{T} \sum_{t=1}^{T} ((w_t' - w_{t-\tau})' \Sigma^{-1} (r_{t+1} - r_{t+1-\tau})) \cdot \frac{1}{T} \sum_{t=1}^{T} (w_t' - w_{t-\tau})' (w_t' - w_{t+1-\tau}) \cdot \frac{1}{T} \sum_{t=1}^{T} (w_t' - w_{t-\tau})' \Sigma^{-1} (w_t' - w_{t+1-\tau})
\]  \hspace{1cm} (A.6)

Here, \( \Sigma = \text{diag}(\text{var}_1, \cdots, \text{var}_N) \), where \( \text{diag}(\cdot) \) represents the diagonal matrix, and the diagonal elements are the variances of \( N \) stock return residuals. We construct the WLS measures of the DGTW measure by replacing the stock returns with the DGTW benchmark adjusted returns, and replacing the variance of the stock return residuals with the variance of the DGTW benchmark adjusted returns. Measures with other benchmarks are treated
similarly. The term 
\[
\frac{1}{T} \sum_{t=1}^{T} (w_t - w_{t-\tau})' (w_t - w_{t+1:tau})
\]
is a normalization. If the number of stocks and number of periods are large,

\[
\frac{1}{T} \sum_{t=1}^{T} (w_t - w_{t-\tau})' \Sigma^{-1} (w_t - w_{t+1:tau})
\]

\[
- \frac{1}{T} \sum_{t=1}^{T} (r_{t+1} - r_{t+1:tau})' \rightarrow 0,
\]
since the two terms above separated by the minus sign are consistent estimators of the numerator of the slope coefficient. Hence, the WLS measure should converge to 

\[
E((w_t - w_{t-\tau})' (r_{t+1} - r_{t+1:tau})).
\]

In other words, if the performance measures are unbiased, the WLS measures should converge to the measures that do not adjust for heterogeneity in the variances when the number of stocks is large enough.

A.3 Feasible GLS Estimation

Our GLS estimation allows for correlations among the error terms through the factor model regression 
\[r_{it+1} = \alpha_i + \beta_i r_{Bt+1} + u_{it+1}.\]
For the DGTW measure, 
\[u_{it+1}\]
is the DGTW benchmark adjusted return. We assume that the factor model error terms are uncorrelated across stocks and over time. (We find that the covariances between different stocks in their DGTW adjusted returns are relatively small.) Like in the WLS case, instead of using the covariance of the error term \(\varepsilon_i'\), in the regression of the returns on the weights, we use the covariance of the stock return residuals in the factor model or the DGTW benchmark adjusted returns. Stack up the \(\varepsilon_i'\) into a TN x 1 vector, \(\varepsilon\), where the first
N elements are the observations for t=1 and the last N are for t=T. We model \( \text{Cov}(\varepsilon) = I_T \otimes [BV_B B' + \Sigma] \), where B is the N x K matrix of factor model betas, \( \beta \), \( V_B \) is the K x K covariance matrix of the benchmark factors \( r_B \), and \( \Sigma \) is the N x N diagonal matrix of the factor model regression residual variances. All of the inputs are estimated from the factor model regression (4), run for each stock as an OLS seemingly-unrelated regression.

Write the panel regression with fixed effects as \( r = X \gamma + \varepsilon \), where \( r \) is the TN x 1 vector of excess stock returns, \( X \) is a TN x (N+1) data matrix, where the first N columns are T rows of N x N identity matrices, \( I_N \), and the last column contains the TN x 1 vector of weights, \( w \). The (N+1) parameter vector to be estimated, \( \gamma \), includes the N fixed effects, \( a_i \), followed by the panel slope coefficient, \( \beta \). The model is estimated by OLS on the transformed variables, where \( \text{Cov}(\varepsilon)^{1/2} r \) is regressed on \( \text{Cov}(\varepsilon)^{1/2} X \).

### A.4. Auxiliary Parameter Estimation Error

The first step is to estimate the K+1 parameters \((a,b)\) in the SDF model. This is accomplished through an exactly identified GMM (Hansen, 1982) system:

\[
g_1 = (1/T) \Sigma_t [a - b'r_B]r_Bt \quad (A.7)
\]
\[
g_2 = (1/T) \Sigma_t [a - b'r_B]R_{ft} - 1. \quad (A.8)
\]

The sample moment condition \( g=(g_1',g_2')' \) is set equal to zero, and \( R_{ft} \) is the gross (one plus the rate of) return on the short term Treasury bill. The optimal covariance matrix from Hansen (1982), in which we assume zero moving average terms, delivers the covariance of the parameter estimates, \( V(a,b) \).
Following the previous subsection, the OLS panel regression estimator given the true values of the parameters \((a, b)\) can be written as 
\[
\hat{\beta} = (X'X)^{-1}X'y,
\]
where \(y\) is the TN x 1 vector of the \(r_{it+1}(a-b'r_{Bt+1})\). Letting \(\hat{y}\) be the TN x 1 vector plugging in the first step parameter estimates, 
\[
y_{it+1} = r_{it+1}(\hat{a} - \hat{b}'r_{Bt+1})
\]
and let 
\[
\hat{\beta} = (X'X)^{-1}X'\hat{y}.
\]
We can write:
\[
\hat{\beta} - \beta = \hat{\beta} - \beta + (X'X)^{-1}X' \left[ (\hat{a} - a) - (\hat{b} - b)'r_B \right] r,
\]
(A.9)
where \(r\) is the TN x 1 vector of the \(r_{it}\). Applying the delta method, \(\hat{\beta} - \beta\) is a function \(f(\varphi)\) of the \(K+2\) parameter estimates \(\varphi = (\hat{\beta}, \hat{a}, \hat{b})\). We approximate the sample variance
\[
V(\hat{\beta}) \approx \partial f / \partial(\varphi) \text{Var}(\varphi) \partial f / \partial(\varphi)',
\]
where the gradients follow from (A.9). We approximate the \((K+2)\) square matrix, \(\text{Var}(\varphi)\), by putting the panel estimator of \(\text{Var}(\hat{\beta})\) in its upper left position, the \((K+1)\) square matrix \(V(a,b)\) in its lower right block, and we fill in out the matrix with two \(K\)-vectors of zeros. This assumes that the covariances between the panel slope estimate and the estimates of the parameters \((a,b)\) in the first step is negligible.
References


Ferreira, Kwswanit, Miguel and Ramos, 2013,


Hansen, Lars P., 1982, Large sample properties of generalized method of moments


Jordan, Bradford and Timothy Riley, 2016, Volatility and mutual fund manager skill, working paper, University of Kentucky.


Roll, Richard, 1978, Ambiguity when performance is measured by the security market


Table 1: Summary Statistics

In panel A, for each fund in the sample we compute the time-series average of its total net assets (TNA) in millions of dollars, the average number of stocks held, the return gap and the active weight. We compute the sample standard deviations of the reported fund returns, σ, and the R-squares of factor model regressions in the Carhart (1997) four-factor model. We also compute the sample autocorrelations of the funds’ portfolio weights, ρ_j, averaged across the holdings, at various lags j, j=1,…5. The means, std, max and min are taken across the funds in the sample. Error Corr is the correlation between the errors of the portfolio weight autocorrelation regression and the panel regression of future stock returns on the weights, averaged across the holdings. The sample period is from 1984 to 2012, and the number of funds is 3596. In Panel B, we compute the autoregression for the weights for each stock and average across the funds that hold the stock. The descriptive statistics are calculated at the stock level. In Panel C, we compute the autocorrelation of the first differences in the weights.

Panel A: Descriptive statistics at the fund level

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNA ($million)</td>
<td>684</td>
<td>2102</td>
<td>44496</td>
<td>1.03</td>
</tr>
<tr>
<td>Number of stocks</td>
<td>114</td>
<td>198</td>
<td>3317</td>
<td>10</td>
</tr>
<tr>
<td>Return Gap (%)</td>
<td>0.03</td>
<td>0.23</td>
<td>1.85</td>
<td>-2.6</td>
</tr>
<tr>
<td>Active Weight</td>
<td>0.40</td>
<td>0.10</td>
<td>0.90</td>
<td>0.01</td>
</tr>
<tr>
<td>Return σ</td>
<td>0.01</td>
<td>0.01</td>
<td>0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>R-squares</td>
<td>0.87</td>
<td>0.14</td>
<td>0.99</td>
<td>0.06</td>
</tr>
<tr>
<td>ρ1</td>
<td>0.9424</td>
<td>0.0398</td>
<td>0.9981</td>
<td>0.1957</td>
</tr>
<tr>
<td>ρ2</td>
<td>0.8854</td>
<td>0.0730</td>
<td>0.9966</td>
<td>-0.0306</td>
</tr>
<tr>
<td>ρ3</td>
<td>0.8282</td>
<td>0.1092</td>
<td>0.9962</td>
<td>-0.5256</td>
</tr>
<tr>
<td>ρ4</td>
<td>0.7823</td>
<td>0.1277</td>
<td>0.9962</td>
<td>-0.5187</td>
</tr>
<tr>
<td>ρ5</td>
<td>0.7366</td>
<td>0.1478</td>
<td>0.9962</td>
<td>-0.5187</td>
</tr>
<tr>
<td>Error Corr</td>
<td>-0.0043</td>
<td>0.0154</td>
<td>0.2430</td>
<td>-0.1868</td>
</tr>
</tbody>
</table>

Panel B: AR coefficients of the weights at the stock level

<table>
<thead>
<tr>
<th>ρ1</th>
<th>Mean</th>
<th>Std</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9027</td>
<td>0.0380</td>
<td>0.9971</td>
<td>0.2638</td>
<td></td>
</tr>
<tr>
<td>0.8132</td>
<td>0.0656</td>
<td>0.9921</td>
<td>0.1978</td>
<td></td>
</tr>
<tr>
<td>0.7289</td>
<td>0.0901</td>
<td>0.9868</td>
<td>-0.5539</td>
<td></td>
</tr>
<tr>
<td>0.6553</td>
<td>0.1040</td>
<td>0.9853</td>
<td>-0.4970</td>
<td></td>
</tr>
<tr>
<td>0.5856</td>
<td>0.1172</td>
<td>0.9852</td>
<td>-0.3851</td>
<td></td>
</tr>
</tbody>
</table>
Panel C: Average Autoregressive coefficients for weight differences

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_1$</td>
<td>-0.0091</td>
<td>0.0180</td>
<td>0.0627</td>
<td>-0.1223</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.0090</td>
<td>0.0185</td>
<td>0.0634</td>
<td>-0.1833</td>
</tr>
<tr>
<td>$\rho_3$</td>
<td>-0.0271</td>
<td>0.0422</td>
<td>0.1211</td>
<td>-0.4513</td>
</tr>
<tr>
<td>$\rho_4$</td>
<td>-0.0091</td>
<td>0.0184</td>
<td>0.0650</td>
<td>-0.1897</td>
</tr>
<tr>
<td>$\rho_5$</td>
<td>-0.0086</td>
<td>0.0177</td>
<td>0.0638</td>
<td>-0.0910</td>
</tr>
</tbody>
</table>
Table 2: Simulations to Address Stochastic Regressor Bias

The table reports the average across 1000 simulation trials, of estimated holdings based performance measures. The units are percent excess return per quarter. True denotes the actual values of the measures used to calibrate the simulations, bootstrapping from versions of Equations (8) and (14). These are found at various fractiles in the cross-section of actual fund data. The “No alpha” OLS estimates the slope coefficient in the baseline panel regression of (6), Diff IV is the differenced IV estimator, Diff no IV is the classical difference estimator of the regression (8) with stock dummies and Within is the classical within-group (least squares with dummy variables) estimator. OLS with alpha is based on the model in footnote 7. Haj2010 is the bias adjusted estimator in (18) and Haj2008 is the bias adjusted estimator in (16). GT is the weight change measure of Grinblatt and Titman (1993), DGTW is the measure of Grinblatt, Titman and Wermers (1997). CWM is the conditional weight measure of Ferson and Khang (2002). FM is the stochastic discount factor measure of Ferson and Mo (2016).

Panel A: Alternative Holdings-based Performance Estimators

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>No alpha OLS</th>
<th>OLS with alpha</th>
<th>Haj2008</th>
<th>Within</th>
<th>Diff</th>
<th>Diff IV</th>
<th>Haj2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.29</td>
<td>0.13</td>
<td>-0.16</td>
<td>-0.25</td>
<td>-0.37</td>
<td>-0.44</td>
<td>-0.30</td>
<td>-0.30</td>
</tr>
<tr>
<td></td>
<td>-0.19</td>
<td>0.17</td>
<td>-0.01</td>
<td>-0.14</td>
<td>-0.32</td>
<td>-0.22</td>
<td>-0.22</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
<td>0.17</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>0.48</td>
<td>0.39</td>
<td>0.42</td>
<td>0.37</td>
<td>0.35</td>
<td>0.40</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>0.58</td>
<td>0.92</td>
<td>0.58</td>
<td>0.60</td>
<td>0.46</td>
<td>0.40</td>
<td>0.60</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Panel B: Classical Measures when there are no Fixed Effects

<table>
<thead>
<tr>
<th></th>
<th>True (GT)</th>
<th>GT</th>
<th>True (DGTW)</th>
<th>DGTW</th>
<th>True (CWM)</th>
<th>Diff IV CWM</th>
<th>True (FM)</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.29</td>
<td>-0.29</td>
<td>-0.26</td>
<td>-0.29</td>
<td>-0.38</td>
<td>-0.37</td>
<td>-0.54</td>
<td>-0.54</td>
</tr>
<tr>
<td></td>
<td>-0.19</td>
<td>-0.17</td>
<td>-0.18</td>
<td>-0.18</td>
<td>-0.25</td>
<td>-0.24</td>
<td>-0.37</td>
<td>-0.38</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.06</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>0.41</td>
<td>0.22</td>
<td>0.23</td>
<td>0.16</td>
<td>0.14</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>0.58</td>
<td>0.59</td>
<td>0.35</td>
<td>0.32</td>
<td>0.24</td>
<td>0.21</td>
<td>0.38</td>
<td>0.38</td>
</tr>
</tbody>
</table>
Table 3: Simulations of Hypothetical Strategies

Two strategies are simulated: buy-and-hold and momentum strategies. The returns are assumed to follow the DGTW factor model, i.e. the return is a summation of the DGTW benchmark adjusted return plus the DGTW benchmark return. The weights are simulated following section 5.4. The average value of the measures (annualized percentage) are shown in the first row of each panel. The distributions of the T-statistics are summarized by reporting the values at the 2.5%, 5%, 50%, 95%, and 97.5% tails of the 1000 simulations. The percentage of simulated T-statistics larger than 1.96 are presented in the bottom row of each panel. GT and DGTW are the original estimates, while DiffIV uses the differenced IV method and Haj10 uses the Hjalmarsson (2010) estimator for the models with fixed effects.

Panel A: Simulated Buy-and-Hold

<table>
<thead>
<tr>
<th></th>
<th>GT</th>
<th>DGTW</th>
<th>DiffIV GT</th>
<th>DiffIV DGTW</th>
<th>Haj10 GT</th>
<th>Haj10 DGTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Measure</td>
<td>1.04</td>
<td>9.38</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.03</td>
</tr>
<tr>
<td>2.5%</td>
<td>-0.96</td>
<td>3.00</td>
<td>-2.37</td>
<td>-2.10</td>
<td>-2.24</td>
<td>-2.33</td>
</tr>
<tr>
<td>5.0%</td>
<td>-0.74</td>
<td>3.51</td>
<td>-1.97</td>
<td>-1.76</td>
<td>-1.98</td>
<td>-2.04</td>
</tr>
<tr>
<td>50%</td>
<td>1.18</td>
<td>7.50</td>
<td>-0.09</td>
<td>-0.01</td>
<td>-0.15</td>
<td>-0.34</td>
</tr>
<tr>
<td>95%</td>
<td>2.50</td>
<td>10.08</td>
<td>1.78</td>
<td>1.69</td>
<td>1.52</td>
<td>1.26</td>
</tr>
<tr>
<td>97.5%</td>
<td>2.77</td>
<td>10.48</td>
<td>2.10</td>
<td>1.86</td>
<td>1.74</td>
<td>1.52</td>
</tr>
<tr>
<td>&gt;1.96</td>
<td>16.80%</td>
<td>99.80%</td>
<td>3.10%</td>
<td>1.90%</td>
<td>1.40%</td>
<td>0.50%</td>
</tr>
</tbody>
</table>

Panel B: Simulated Momentum Trading

<table>
<thead>
<tr>
<th></th>
<th>GT</th>
<th>DGTW</th>
<th>DiffIV GT</th>
<th>DiffIV DGTW</th>
<th>Haj10 GT</th>
<th>Haj10 DGTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Measure</td>
<td>0.42</td>
<td>7.57</td>
<td>-0.06</td>
<td>-0.10</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>2.5%</td>
<td>-1.14</td>
<td>9.61</td>
<td>-2.31</td>
<td>-2.45</td>
<td>-2.20</td>
<td>-2.37</td>
</tr>
<tr>
<td>5.0%</td>
<td>-0.85</td>
<td>9.87</td>
<td>-2.00</td>
<td>-2.16</td>
<td>-1.83</td>
<td>-1.96</td>
</tr>
<tr>
<td>50%</td>
<td>0.84</td>
<td>11.39</td>
<td>-0.09</td>
<td>-0.38</td>
<td>0.10</td>
<td>-0.09</td>
</tr>
<tr>
<td>95%</td>
<td>2.70</td>
<td>13.38</td>
<td>1.90</td>
<td>1.48</td>
<td>1.87</td>
<td>1.73</td>
</tr>
<tr>
<td>97.5%</td>
<td>3.15</td>
<td>13.79</td>
<td>2.33</td>
<td>1.87</td>
<td>2.25</td>
<td>2.09</td>
</tr>
<tr>
<td>&gt;1.96</td>
<td>16.70%</td>
<td>100.00%</td>
<td>4.80%</td>
<td>1.90%</td>
<td>4.20%</td>
<td>3.10%</td>
</tr>
</tbody>
</table>
Table 5: Holdings-based Performance Measures

The various performance measures are estimated for each fund in the sample using panel regressions. GT is the portfolio change measure, DGTW is the DGTW characteristic selectivity measure, FM is the Ferson and Mo SDF-based measure and CWM is the conditional weight based measure. Funds are sorted into five groups on the basis of the original measures, estimated using Equation (6) and shown in the first column. The second column, subscripted with H, indicates an estimate that includes stock fixed effects in the panel regression, using the Hjalmarsson (2010) method. The third column is the average alpha effect extracted by the fixed effects. The average style component of the GT measure is described in the text. The units of the performance measures are percent per month. The sample period is from 1980 to 2012, and the number of funds is 3596. Data for the Wilshire 5000 index fund cover July, 1999 through June, 2012.

----

Panel A: Grinblatt Titman Portfolio Change Measure

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>GT</th>
<th>$GT_H$</th>
<th>Avg. Alpha</th>
<th>Avg. Style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-0.29</td>
<td>-0.08</td>
<td>-0.22</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>0.02</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>0.06</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>High</td>
<td>0.35</td>
<td>0.26</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Mean</td>
<td>0.02</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Wilshire 5000</td>
<td>-0.07</td>
<td>-0.03</td>
<td>-0.06</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Table 5, page 2

Panel B: DGTW Characteristic Selectivity Measure

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>DGTW</th>
<th>DGTW_H</th>
<th>Avg. Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-0.30</td>
<td>-0.12</td>
<td>-0.19</td>
</tr>
<tr>
<td>2</td>
<td>-0.07</td>
<td>-0.09</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>-0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>4</td>
<td>0.07</td>
<td>-0.05</td>
<td>0.12</td>
</tr>
<tr>
<td>High</td>
<td>0.25</td>
<td>0.00</td>
<td>0.25</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.01</td>
<td>-0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Wilshire 5000</td>
<td>0.01</td>
<td>-0.10</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Panel C: Conditional Weight-based Measure

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>CWM_H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-0.20</td>
</tr>
<tr>
<td>2</td>
<td>-0.09</td>
</tr>
<tr>
<td>3</td>
<td>-0.06</td>
</tr>
<tr>
<td>4</td>
<td>-0.04</td>
</tr>
<tr>
<td>High</td>
<td>0.01</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.08</td>
</tr>
<tr>
<td>Wilshire 5000</td>
<td>-0.10</td>
</tr>
</tbody>
</table>

Panel D: Ferson and Mo Stochastic Discount Factor Measure

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>FM</th>
<th>FM_H</th>
<th>Avg. Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-1.61</td>
<td>-0.18</td>
<td>-1.44</td>
</tr>
<tr>
<td>2</td>
<td>-0.44</td>
<td>-0.21</td>
<td>-0.23</td>
</tr>
<tr>
<td>3</td>
<td>-0.11</td>
<td>-0.14</td>
<td>0.03</td>
</tr>
<tr>
<td>4</td>
<td>0.13</td>
<td>-0.11</td>
<td>0.24</td>
</tr>
<tr>
<td>High</td>
<td>0.70</td>
<td>-0.10</td>
<td>0.80</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.27</td>
<td>-0.15</td>
<td>-0.12</td>
</tr>
<tr>
<td>Wilshire 5000</td>
<td>-0.38</td>
<td>-0.29</td>
<td>-0.09</td>
</tr>
</tbody>
</table>
Table 6: Estimates of Holdings-based Performance Measures sorted by fund-Characteristics

The funds are sorted by return-gap, R-square, active weight and volatility of returns in quantiles. Eight weight-based measures are calculated, and T-statistics are estimated. GT is the portfolio change measure, DGTW is the DGTW characteristic selectivity measure, and FM is the Ferson and Mo SDF-based total alpha measure and CWM is the conditional weight based measure. For each measure the second column, subscripted with $H$, indicates an estimate that includes stock fixed effects in the panel regression, estimated according to the method of Hajmarlsson (2010).


<table>
<thead>
<tr>
<th>Fund Quintile</th>
<th>GT</th>
<th>DGTW</th>
<th>CWM</th>
<th>FM</th>
<th>GT$^H$</th>
<th>DGTW$^H$</th>
<th>CWM$^H$</th>
<th>FM$^H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.20</td>
<td>0.04</td>
<td>-0.51</td>
<td>0.14</td>
<td>0.17</td>
<td>-0.00</td>
<td>-0.01</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(2.74)</td>
<td>(1.22)</td>
<td>(-5.67)</td>
<td>(0.53)</td>
<td>(2.58)</td>
<td>(-0.06)</td>
<td>(-0.17)</td>
<td>(-0.57)</td>
</tr>
<tr>
<td>2</td>
<td>0.13</td>
<td>0.04</td>
<td>-0.45</td>
<td>0.14</td>
<td>0.09</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(2.60)</td>
<td>(2.05)</td>
<td>(-5.62)</td>
<td>(0.53)</td>
<td>(2.20)</td>
<td>(-0.41)</td>
<td>(-0.19)</td>
<td>(-0.49)</td>
</tr>
<tr>
<td>3</td>
<td>0.09</td>
<td>0.02</td>
<td>-0.44</td>
<td>0.12</td>
<td>0.06</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(2.05)</td>
<td>(1.48)</td>
<td>(-5.87)</td>
<td>(0.45)</td>
<td>(1.44)</td>
<td>(-1.31)</td>
<td>(-0.57)</td>
<td>(-0.96)</td>
</tr>
<tr>
<td>4</td>
<td>0.08</td>
<td>0.04</td>
<td>-0.48</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(1.94)</td>
<td>(1.65)</td>
<td>(-6.59)</td>
<td>(0.40)</td>
<td>(1.89)</td>
<td>(-1.12)</td>
<td>(-0.70)</td>
<td>(-0.93)</td>
</tr>
<tr>
<td>Low</td>
<td>0.11</td>
<td>0.01</td>
<td>-0.52</td>
<td>0.07</td>
<td>0.08</td>
<td>-0.03</td>
<td>-0.06</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(2.45)</td>
<td>(0.85)</td>
<td>(-6.46)</td>
<td>(0.27)</td>
<td>(2.04)</td>
<td>(-0.75)</td>
<td>(-0.96)</td>
<td>(-0.81)</td>
</tr>
<tr>
<td>Low-High</td>
<td>-0.08</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.06</td>
<td>-0.08</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(-2.45)</td>
<td>(-1.36)</td>
<td>(-0.34)</td>
<td>(-1.13)</td>
<td>(-2.65)</td>
<td>(-1.70)</td>
<td>(-1.89)</td>
<td>(-0.59)</td>
</tr>
</tbody>
</table>
Panel B: R-square (1999-2012)

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>GT</th>
<th>DGTW</th>
<th>CWM</th>
<th>FM</th>
<th>GT_H</th>
<th>DGTW_H</th>
<th>CWM_H</th>
<th>FM_H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.05</td>
<td>0.08</td>
<td>-0.42</td>
<td>-0.05</td>
<td>0.03</td>
<td>-0.07</td>
<td>-0.08</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>(0.92)</td>
<td>(1.88)</td>
<td>(-2.48)</td>
<td>(-0.11)</td>
<td>(0.66)</td>
<td>(-1.88)</td>
<td>(-1.81)</td>
<td>(-1.53)</td>
</tr>
<tr>
<td>2</td>
<td>-0.04</td>
<td>0.07</td>
<td>-0.55</td>
<td>-0.13</td>
<td>-0.04</td>
<td>-0.13</td>
<td>-0.14</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>(-0.68)</td>
<td>(1.41)</td>
<td>(-3.81)</td>
<td>(-0.26)</td>
<td>(-0.68)</td>
<td>(-2.45)</td>
<td>(-2.29)</td>
<td>(-2.08)</td>
</tr>
<tr>
<td>3</td>
<td>-0.06</td>
<td>0.02</td>
<td>-0.61</td>
<td>-0.23</td>
<td>0.00</td>
<td>-0.12</td>
<td>-0.19</td>
<td>-0.36</td>
</tr>
<tr>
<td></td>
<td>(-1.01)</td>
<td>(0.46)</td>
<td>(-5.96)</td>
<td>(-0.48)</td>
<td>(0.07)</td>
<td>(-2.20)</td>
<td>(-1.99)</td>
<td>(-1.94)</td>
</tr>
<tr>
<td>4</td>
<td>-0.09</td>
<td>0.00</td>
<td>-0.59</td>
<td>-0.29</td>
<td>-0.01</td>
<td>-0.13</td>
<td>-0.13</td>
<td>-0.37</td>
</tr>
<tr>
<td></td>
<td>(-1.37)</td>
<td>(0.03)</td>
<td>(-7.26)</td>
<td>(-0.62)</td>
<td>(-0.10)</td>
<td>(-2.57)</td>
<td>(-2.06)</td>
<td>(-1.90)</td>
</tr>
<tr>
<td>High</td>
<td>-0.07</td>
<td>0.00</td>
<td>-0.59</td>
<td>-0.35</td>
<td>-0.02</td>
<td>-0.12</td>
<td>-0.09</td>
<td>-0.30</td>
</tr>
<tr>
<td></td>
<td>(-1.43)</td>
<td>(0.04)</td>
<td>(-7.25)</td>
<td>(-0.74)</td>
<td>(-0.53)</td>
<td>(-2.85)</td>
<td>(-2.14)</td>
<td>(-2.02)</td>
</tr>
<tr>
<td>High-low</td>
<td>-0.13</td>
<td>-0.08</td>
<td>-0.17</td>
<td>-0.30</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.01</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(-2.47)</td>
<td>(-1.84)</td>
<td>(-1.50)</td>
<td>(-3.28)</td>
<td>(-1.26)</td>
<td>(-1.46)</td>
<td>(-0.41)</td>
<td>(-1.12)</td>
</tr>
</tbody>
</table>

Panel C: Active weight (1980-2012)

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>GT</th>
<th>DGTW</th>
<th>CWM</th>
<th>FM</th>
<th>GT_H</th>
<th>DGTW_H</th>
<th>CWM_H</th>
<th>FM_H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.08</td>
<td>-0.01</td>
<td>-0.50</td>
<td>0.02</td>
<td>0.07</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(1.51)</td>
<td>(-0.73)</td>
<td>(-6.73)</td>
<td>(0.08)</td>
<td>(1.28)</td>
<td>(-1.27)</td>
<td>(-0.53)</td>
<td>(-0.77)</td>
</tr>
<tr>
<td>2</td>
<td>0.06</td>
<td>0.02</td>
<td>-0.46</td>
<td>0.06</td>
<td>0.05</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>(1.57)</td>
<td>(1.09)</td>
<td>(-5.15)</td>
<td>(0.22)</td>
<td>(1.07)</td>
<td>(-1.46)</td>
<td>(-0.76)</td>
<td>(-0.93)</td>
</tr>
<tr>
<td>3</td>
<td>0.10</td>
<td>0.02</td>
<td>-0.44</td>
<td>0.08</td>
<td>0.07</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(2.49)</td>
<td>(1.09)</td>
<td>(-5.38)</td>
<td>(0.30)</td>
<td>(1.42)</td>
<td>(-1.24)</td>
<td>(-0.45)</td>
<td>(-0.72)</td>
</tr>
<tr>
<td>4</td>
<td>0.09</td>
<td>0.03</td>
<td>-0.43</td>
<td>0.10</td>
<td>0.07</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(1.43)</td>
<td>(-5.80)</td>
<td>(0.38)</td>
<td>(1.44)</td>
<td>(-1.50)</td>
<td>(-0.96)</td>
<td>(-0.75)</td>
</tr>
<tr>
<td>High</td>
<td>0.14</td>
<td>0.06</td>
<td>-0.41</td>
<td>0.17</td>
<td>0.09</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(3.23)</td>
<td>(2.00)</td>
<td>(-4.88)</td>
<td>(0.63)</td>
<td>(2.17)</td>
<td>(0.01)</td>
<td>(-0.53)</td>
<td>(-0.10)</td>
</tr>
<tr>
<td>High-Low</td>
<td>0.06</td>
<td>0.07</td>
<td>0.09</td>
<td>0.15</td>
<td>0.02</td>
<td>0.04</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(2.95)</td>
<td>(1.76)</td>
<td>(2.49)</td>
<td>(0.66)</td>
<td>(2.27)</td>
<td>(0.31)</td>
<td>(1.65)</td>
</tr>
</tbody>
</table>

Panel D: Return volatility (1999-2012)

<table>
<thead>
<tr>
<th>Fund Quintile:</th>
<th>GT</th>
<th>DGTW</th>
<th>CWM</th>
<th>FM</th>
<th>GT_H</th>
<th>DGTW_H</th>
<th>CWM_H</th>
<th>FM_H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.29</td>
<td>-0.13</td>
<td>-0.00</td>
<td>-0.09</td>
<td>-0.07</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(0.88)</td>
<td>(-1.73)</td>
<td>(-0.28)</td>
<td>(-0.06)</td>
<td>(-2.56)</td>
<td>(-2.04)</td>
<td>(-2.05)</td>
</tr>
<tr>
<td>2</td>
<td>-0.04</td>
<td>0.03</td>
<td>-0.42</td>
<td>-0.19</td>
<td>-0.02</td>
<td>-0.11</td>
<td>-0.09</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>(-0.86)</td>
<td>(0.97)</td>
<td>(-2.51)</td>
<td>(-0.43)</td>
<td>(-0.58)</td>
<td>(-2.62)</td>
<td>(-2.14)</td>
<td>(-2.14)</td>
</tr>
<tr>
<td>3</td>
<td>-0.04</td>
<td>0.02</td>
<td>-0.58</td>
<td>-0.26</td>
<td>0.01</td>
<td>-0.11</td>
<td>-0.09</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>(-0.80)</td>
<td>(0.78)</td>
<td>(-4.68)</td>
<td>(-0.55)</td>
<td>(0.20)</td>
<td>(-2.39)</td>
<td>(-1.78)</td>
<td>(-1.77)</td>
</tr>
<tr>
<td>4</td>
<td>-0.07</td>
<td>0.05</td>
<td>-0.64</td>
<td>-0.25</td>
<td>-0.00</td>
<td>-0.12</td>
<td>-0.12</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(-0.94)</td>
<td>(0.86)</td>
<td>(-6.55)</td>
<td>(-0.50)</td>
<td>(-0.02)</td>
<td>(-2.45)</td>
<td>(-1.95)</td>
<td>(-1.95)</td>
</tr>
<tr>
<td>High</td>
<td>-0.10</td>
<td>0.03</td>
<td>-0.81</td>
<td>-0.22</td>
<td>-0.01</td>
<td>-0.14</td>
<td>-0.19</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>(-1.09)</td>
<td>(0.26)</td>
<td>(-6.71)</td>
<td>(-0.42)</td>
<td>(-0.21)</td>
<td>(-2.12)</td>
<td>(-2.18)</td>
<td>(-2.16)</td>
</tr>
<tr>
<td>High-Low</td>
<td>-0.12</td>
<td>-0.01</td>
<td>-0.52</td>
<td>-0.10</td>
<td>-0.01</td>
<td>-0.05</td>
<td>-0.12</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(-1.60)</td>
<td>(-0.04)</td>
<td>(-2.18)</td>
<td>(-0.45)</td>
<td>(-0.25)</td>
<td>(-0.94)</td>
<td>(-1.66)</td>
<td>(-1.66)</td>
</tr>
</tbody>
</table>