

# Generalized Portfolio Sorts for Factor Validation\*

**Daniel Hoechle<sup>#</sup>**

Institute for Finance, FHNW School of Business, CH-4002 Basel, Switzerland

**Markus Schmid**

Swiss Institute of Banking and Finance, University of St. Gallen, Swiss Finance Institute (SFI),  
and European Corporate Governance Institute (ECGI), CH-9000 St. Gallen, Switzerland

**Heinz Zimmermann**

University of Basel, Department of Finance, CH-4002 Basel, Switzerland

This version: March 2025

## Abstract

Portfolio sorts are widely used in empirical asset pricing to identify firm characteristics that predict stock returns. However, such tests can conflate genuine characteristic-based predictability with persistent, firm-level heterogeneity. To address this limitation, we propose a Generalized Portfolio Sorts (GPS) model, which can exactly replicate results from all variants of conventional portfolio sorts, but can also be specified so that it separates a firm characteristic's genuine predictive power from stable firm-level factors. We also derive a statistical test to detect whether return predictability arises from the sorting characteristic itself or from persistent, firm-level traits. Applied to a large set of proposed asset pricing predictors, we find that nearly half lose significance once persistent, firm-level heterogeneity is accounted for. The GPS-model thus strengthens factor validation, advances our understanding of the factor zoo, and provides a more robust foundation for empirical asset pricing tests.

**Keywords:** Portfolio sorts, Cross-section of expected returns, Tests of asset pricing models, Random effects assumption

**JEL classification :** C21, G14, D1

---

\* We are grateful to Yakov Amihud, Vitor Azevedo, Brad Barber, Trond Daskeland, John Driscoll, Can Gao, Frank Graef, Michael Halling, Bryan Kelly, Alexander Kempf, Sunyoung Kim, Tim Kroencke, Beni Lauterbach, Nora Laurinaityte, Juhani Linnainmaa, Ernst Maug, Hannes Mohrschladt, Honkanen Pekka, Andy Neuhierl, Stefan Ruenzi, Paul Söderlind, Martin Weber, Michael Weber, and conference and seminar participants at the 2021 American Finance Association (AFA) Annual Meeting in Chicago (virtual conference), 17th Colloquium on Financial Markets in Cologne, 21st Annual Conference of the Swiss Society for Financial Market Research (SGF) in Zurich, the 27th European Financial Management Association Annual Meetings in Milan, the 25th Annual Meeting of the German Finance Association (DGF) in Trier, the 5th SAFE Asset Pricing Workshop at Goethe University Frankfurt, the 16th Paris December Finance Meeting, 6th Paris Financial Management Conference, University of Basel, University of St. Gallen, WHU – Otto Beisheim School of Management, and University of Mannheim for helpful discussions and valuable comments. A Stata program implementing our methodology, a sample dataset, and a tutorial illustrating our approach can be downloaded using the following three Stata commands (1) `ssc install xtsc, all replace`, (2) `ssc install estout, all replace`, and (3) `ssc install xtgps, all replace`. The tutorial illustrating our approach is stored in Stata's active directory as "xtgps\_material.zip".

<sup>#</sup> Corresponding author: Tel.: +41-61-279-1773; Email address: daniel.hoechle@fhnw.ch.

Address: Institute for Finance, FHNW School of Business, Peter Merian-Strasse 86, CH-4002 Basel, Switzerland.

## 1. Introduction

The identification of firm characteristics that predict stock returns is a cornerstone of empirical asset pricing. Yet, amid the ever growing “factor zoo” of proposed return predictors, disagreement about which characteristics truly drive expected returns persists. Portfolio sorts, an intuitive two-step methodology introduced by Black et al. (1972) and popularized by Fama and French (1993), are widely used to assess whether portfolios formed on specific characteristics earn alpha relative to standard asset pricing models.<sup>1</sup>

In this paper, we challenge the conventional interpretation of portfolio sorts by demonstrating that their standard implementation can conflate genuine characteristic-based predictability with persistent, firm-level heterogeneity. This problem emanates from the implicit assumption of standard portfolio sorts that unobserved firm-specific risk is uncorrelated with the sorting characteristic(s), generally referred to as the “random effects assumption” (RE assumption) in econometrics. When the sorting characteristic coincidentally correlates with some long-lived firm traits (e.g., innovativeness or management quality), the random effects assumption fails and portfolio sorts are prone to overstating a characteristic’s true predictive power by blending genuine return predictability with persistent firm-level components. In such cases, it is unclear whether observed predictive power originates from the sorting characteristic itself or from persistent, unobserved firm-level attributes correlated with that characteristic.

Our paper contributes to the sparse literature analyzing the econometric properties of portfolio sorts.<sup>2</sup> We are the first to systematically study how persistent firm-level heterogeneity affects risk-adjusted returns from portfolio sorts. To this end, we propose

---

<sup>1</sup> Recent examples of empirical asset pricing studies using the portfolio sorts approach are Novy-Marx (2013), Frazzini and Pedersen (2014), Ball et al. (2015), Fama and French (2015), Bali et al. (2017), Gerakos and Linnainmaa (2018), Weber (2018), van Binsbergen and Opp (2019), Ehsani and Linnainmaa (2022), and Friewald et al. (2022), among many others. The portfolio sorts approach is also widely used in other areas of empirical finance research, such as derivative research (e.g., Heston et al., 2023), cryptocurrency research (e.g., Liu et al., 2022), ESG research (e.g., Hsu et al., 2023), household finance (e.g., Barber and Odean, 2000, 2001; Bali et al., 2023), research on insider trading (e.g., Hvide and Meisner Nielsen, 2025), and studies analyzing the performance of mutual funds and hedge funds (e.g., Kacperczyk et al., 2008; Fung et al., 2008; Gerakos et al., 2021).

<sup>2</sup> Research covering methodological and econometric aspects of the portfolio sorts approach includes Fama (1998), Lyon et al. (1999), Loughran and Ritter (2000), Kothari and Warner (2008), and Cochrane (2011). More recently, Cattaneo et al. (2020) analyze the statistical properties of portfolio sorts and apply a nonparametric estimator to obtain optimal choices for the number of portfolios.

a “Generalized Portfolio Sorts” (GPS) model that nests the standard approach as a special case. In our GPS-model, individual firms’ monthly excess returns are regressed on a set of market factors (e.g., Fama-French factors), one or more firm characteristics (e.g., gross profitability or volatility), and the interaction terms between these factors and characteristics – all within a single (weighted) pooled OLS regression.<sup>3</sup> We prove theoretically and confirm empirically that this firm-level regression can exactly replicate the alpha and factor exposure estimates (and their  $t$ -statistics) from all variants of conventional portfolio sorts. This result underscores a central insight: Since portfolio sorts can be replicated by a pooled OLS regression, which depends on the random effects assumption, a significant correlation between a sorting characteristic and some persistent firm-level attributes can inflate inferences about a characteristic’s predictive power.

Unobserved heterogeneity is often present in empirical finance research, and a natural solution for addressing it, is to include firm fixed effects (Petersen, 2009; Gormley and Matsa, 2014). Unfortunately, traditional portfolio sorts cannot accommodate firm fixed effects by construction as the predictive power of a firm characteristic is inferred from a *time-series* regression where portfolio returns (e.g., top-minus-bottom decile portfolio returns) are regressed on a set of factors (e.g., the five Fama-French (2015) factors). Our GPS-model remedies this drawback of standard portfolio sorts: By estimating a single firm-level regression, our approach can seamlessly integrate firm fixed effects. The GPS-model thus offers a straightforward solution to isolating the genuine predictive power (if any) of the sorting characteristic from persistent, firm-specific traits.

To evaluate the importance of accounting for persistent, firm-specific risk components in empirical asset pricing applications, we consider an established set of 94 asset pricing predictors from Gu et al. (2020, 2021). For each predictor, we estimate 12 specification variants – three versions of the characteristic (quintile dummies, decile dummies, cross-sectionally standardized) times four factor models (Fama-French three-

---

<sup>3</sup> For ease of exposition, but without loss of generality, we simply refer to “firms” throughout the paper. However, our methodology applies to all areas of empirical finance research, where portfolio sorts are applied. Furthermore, our technique can be used for evaluating the long-term performance of firms, funds, or investors.

factor, Fama-French five-factor, Carhart four-factor, and market model). Of these 1,128 specifications, 532 appear significant under a standard portfolio sorts setup (or pooled OLS estimation). However, when we control for persistent, firm-specific attributes by including firm fixed effects in the analysis, only 270 predictors remain significant, while 262 (49%) either lose their predictive power (215, or 40%) or flip sign while staying significant (47, or 9%). These findings indicate that many documented anomalies may simply reflect time-invariant firm attributes rather than genuine return predictability.

To systematically determine whether a newly proposed factor genuinely forecasts returns instead of reflecting unobserved firm-level attributes, we develop a Hausman (1978) type portfolio sorts specification test. If the test’s null hypothesis (that the random effects assumption holds) is rejected, this is likely *not* the case and the sorting characteristic’s apparent predictive power may in part be due to some stable firm-level attributes that coincidentally correlate with the characteristic. In that scenario, incorporating firm fixed effects into the GPS-model yields a more reliable measure of the characteristic’s genuine predictive power. This highlights that even a large  $t$ -statistic (Harvey et al., 2016) may not guarantee true return predictability; it could reflect p-hacking (Harvey, 2017; Chordia et al., 2020; Hou et al., 2020; Chen, 2021) or, as we demonstrate in this study, systematic cross-sectional correlations that persist across firms. Hence, our GPS-model coupled with the portfolio sorts specification test does not just provide a means to control for firm-level heterogeneity in portfolio-based asset pricing tests. Rather, our approach serves as a robust diagnostic tool: Any characteristic that continues to predict returns after firm-specific fixed effects are included in the analysis is likely to represent a genuine return predictor and not just a proxy for long-lived firm traits.

Our GPS-model relates to the extensive literature documenting hundreds of return predictors, sometimes called the “factor zoo” (Cochrane, 2011). Researchers have pursued various refinements – including machine learning techniques or trading-friction adjustments – to manage the proliferation of proposed anomalies, while noting that

biases related to micro-cap stocks or frequent rebalancing can inflate paper alphas.<sup>4</sup> We point out another concern: Stable, unobserved firm traits may masquerade as characteristic-driven predictability. This complements Dello-Preite et al. (2023), who build on Merton’s (1987) model to derive a stochastic discount factor (SDF) that can price both systematic and unsystematic risk, showing that unsystematic risk accounts for over 70% of their SDF’s variation. Integrating firm fixed effects into our GPS-model helps ensure that persistent, firm-specific attributes do not create an illusion of characteristic-based predictability.

Beyond addressing unobserved firm-level risk, the GPS-model also mitigates the “curse of dimensionality” inherent in standard portfolio sorts. When researchers attempt multiple sorts, the number of portfolios can grow rapidly, limiting empirical work to just a few characteristics (Cochrane, 2011). By contrast, the GPS-model accommodates any number of continuous or discrete characteristics in a single regression, while retaining the intuitive alpha focus that makes portfolio sorts appealing. This flexibility applies regardless of whether the “optimal” number of portfolios is five, ten, or more (Cattaneo et al., 2020), and avoids discarding potentially valuable information by focusing solely on top and bottom group portfolios (Patton and Timmermann, 2010).

Why is it important to revisit the conventional interpretation of portfolio sorts? Portfolio sorts are a key method in empirical asset pricing, as well as in studies of mutual fund performance, IPO underperformance, household finance, and other domains. Their intuitive appeal has made them a convenient benchmark for deciding which characteristics “explain” the cross-section of average stock returns. But if these characteristics largely capture long-lived firm traits, they may not constitute genuine drivers of expected returns. Researchers, practitioners, and policymakers rely on these empirical findings to guide investment strategies, evaluate managerial skill, and shape regulatory discussions. Ensuring that persistent firm-level components are not mistaken for true return predictability is thus of substantial practical and theoretical importance. By transferring well-known econometric tools (firm fixed effects and Hausman tests) into a

---

<sup>4</sup> For example, see Bryzgalova (2016), Pukthuanthong et al. (2018), Huang et al. (2019), Harvey and Liu (2021), and Novy-Marx and Velikov (2024). Examples for machine learning techniques used to tame the factor zoo include Feng et al. (2020), Freyberger et al. (2020), Kozak et al. (2020), and Bryzgalova et al. (2023).

portfolio sorts context, the GPS-model offers an easy to use, yet powerful and widely applicable framework for reliably screening and validating characteristic-based factors in asset pricing, and for investigating the performance of firms, funds, or investors in empirical finance research more broadly.

## 2. A regression-based approach to analyzing asset returns

In this section, we begin by describing the GPS-model setup, its economic interpretation, and the statistical properties of our single-step, firm-level regression approach for analyzing asset returns. Next, we illustrate the methodology's flexibility in handling multiple dimensions and continuous firm (or fund or investor) characteristics. We then propose a framework to distinguish cross-sectional relations from genuine return predictability, a crucial distinction closely related to Hausman's (1978) specification test. Specifically, we show that a firm characteristic that appears to predict the cross-section of stock returns well may not qualify as a genuine return predictor if it fails to exhibit robust explanatory power once persistent firm-level effects are accounted for.

### 2.1 Model setup

Our Generalized Portfolio Sorts (GPS) model for analyzing the cross-section of stock returns employs the following firm-level regression specification, with statistical inference based on Driscoll and Kraay (1998) standard errors that are robust to heteroskedasticity and general forms of cross-sectional and temporal dependence (Driscoll and Kraay, 1998; Hoechle, 2007):

$$r_{it} = (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + c_i + v_{it} \quad (i = 1, \dots, N, \quad t = 1, \dots, T) \quad (1)$$

The dependent variable  $r_{it}$  is the period  $t$  (excess) return of firm  $i$ . Vector  $\mathbf{z}_{it} = [1 \quad z_{2,it} \quad \dots \quad z_{M,it}]$  includes a constant and a set of firm-specific characteristics  $z_{m,it}$  ( $m = 2, \dots, M$ ) which may vary across both firms and time. Vector  $\mathbf{x}_t = [1 \quad x_{1t} \quad \dots \quad x_{Kt}]$  consists of a constant and a series of market-level factor variables  $x_{k,t}$  ( $k = 1, \dots, K$ ) that vary over time but not across firms. Common examples of variables in vector  $\mathbf{x}_t$  include the market excess return, the Fama and French (1993, 2015) size,

value, profitability, and investment factors, as well as the Carhart (1997) momentum factor. The operator  $\otimes$  denotes the Kronecker product. Hence, regression (1) incorporates three types of explanatory variables: individual firm characteristics, market-level factor variables, and all interactions between those firm characteristics and factors. This fully interacted regression model has  $M \times (K + 1)$  explanatory variables, whose coefficients are stored in vector  $\theta$ . Finally,  $c_i$  is an unobserved firm-specific effect with  $E(c_i) = 0$ , and  $v_{it}$  is the disturbance term, assumed to be strictly exogenous (i.e.,  $E[v_{it} | c_i, (z_{i1} \otimes x_1), \dots, (z_{iT} \otimes x_T)] = 0$  for all  $i$  and  $t$ ).

If persistent firm-level effects are either absent ( $c_i = 0 \ \forall i$ ) or uncorrelated with the regressors ( $E[c_i | (z_{it} \otimes x_t)] = 0$ ), regression (1) can be consistently estimated without firm fixed effects using (weighted) pooled OLS.<sup>5</sup> In this scenario, i.e. when  $c_i$  is omitted, regression (1) can *exactly* reproduce the results of all variants of standard portfolio sorts (see Section 2.2). We therefore refer to regression (1) as the “Generalized Portfolio Sorts” model, or the GPS-model.

Unfortunately, the random effects (RE) assumption  $E[c_i | (z_{it} \otimes x_t)] = 0$  cannot generally be assumed to hold. Given the wide-ranging firm-level heterogeneity in size, investments, capital structure, business culture, and management quality, it is plausible that persistent, firm-specific attributes correlate with the explanatory variables in regression (1). However, if  $E[c_i | (z_{it} \otimes x_t)] \neq 0$ , (weighted) pooled OLS is *not* reliable for estimating GPS-model (1). By contrast, the fixed effects (FE) estimator is robust to such correlations because it treats  $c_i$  as an unobserved, long-lived component that may be correlated with the regressors. Hence, by controlling for these stable firm-level influences, FE estimation consistently recovers GPS-model (1) and prevents the sorting characteristic from being confounded with long-lived attributes. As we will show below, this is a major advantage of the GPS-model over standard portfolio sorts.

## 2.2 GPS-model vs. portfolio sorts

Popularized by Fama and French’s (1993, 1996) influential research, the portfolio sorts methodology has become an important empirical tool in finance. The method offers an

---

<sup>5</sup> For details, see Cameron and Trivedi (2005, chapter 21).

intuitive economic interpretation and supports robust statistical inference even in the presence of cross-sectional and temporal dependence (Lyon et al., 1999). In this section, we show that GPS-model (1) can exactly reproduce the results of all variants of standard portfolio sorts. Hence, the GPS-model proposed in this paper has a sound theoretical and statistical foundation, and the coefficient estimates from regression (1) retain the straightforward economic interpretation associated with portfolio sorts.

### 2.2.1 Formal exposition of the portfolio sorts approach

The portfolio sorts procedure generally involves two steps. First, for each period  $t$ , the portfolio (excess) return  $r_{pt}$  is computed for a group of firms  $i$  as follows:

$$r_{pt} = \sum_{i=1}^{N_t} w_{it} r_{it} \quad (2)$$

where  $w_{it}$  denotes the beginning-of-period  $t$  portfolio weight of firm  $i$  ( $i = 1, \dots, N_t$ ), and  $r_{it}$  refers to the firm's stock (excess) return in period  $t$ . The second step assesses the (risk-adjusted) performance of portfolio  $p$  by estimating a linear  $K$ -factor time-series regression with  $r_{pt}$  from (2) as the dependent variable:

$$r_{pt} = \beta_0 + \beta_1 x_{1t} + \dots + \beta_K x_{Kt} + \varepsilon_t \quad (3)$$

In most applications, variables  $x_{kt}$  ( $k = 1, \dots, K$ ) are chosen so that (3) represents a Jensen (1968), Fama and French (1993, 2015), or Carhart (1997) type factor model. The intercept  $\beta_0$  (often called “alpha”) measures the risk-adjusted performance of portfolio  $p$ .

### 2.2.2 Evaluating the performance of a single portfolio with the GPS-model

By estimating the firm-level GPS-model (1) via (weighted) pooled OLS, we can replicate the results of time-series regression (3). To do so, we set  $\mathbf{z}_{it} = [1]$  and  $\mathbf{x}_t = [1 \ x_{1t} \ \dots \ x_{Kt}]$  to obtain<sup>6</sup>

$$r_{it} = (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + v_{it} = \theta_0 + \theta_1 x_{1t} + \dots + \theta_K x_{Kt} + v_{it} \quad (4)$$

---

<sup>6</sup> For consistency with the portfolio sorts approach, throughout Section 2.2 we assume  $c_i = 0$  (for all  $i$ ) and omit the firm-specific effects  $c_i$  from the analysis. This allows us to estimate GPS-model (1) with (weighted) pooled OLS.

For mathematical tractability (without loss of generality), we consider a balanced panel with  $N$  firms,  $T$  time periods, and equal weights  $w_{it} = 1/N$ . Under these conditions, the following result holds true:<sup>7</sup>

**Proposition 1 (Single Portfolio)**

- **Part A – Coefficient estimates.** *Estimating linear regression (4) with pooled OLS yields coefficient estimates identical to those from time-series regression (3) estimated with OLS, i.e.,  $\hat{\theta}_k \equiv \hat{\beta}_k$  ( $\forall k = 0, 1, \dots, K$ ).*
- **Part B – Standard errors.** *For a given lag length  $H$ , Driscoll and Kraay (1998) standard errors for  $\hat{\theta}_k$  ( $k = 0, 1, \dots, K$ ) in pooled OLS regression (4) coincide with Newey and West (1987) standard errors for  $\hat{\beta}_k$  in (3).*

Proof: See Internet Appendix A.1.

Part A of Proposition 1 follows from a known property of portfolio theory: the portfolio beta equals the weighted sum of individual asset betas. Part B of Proposition 1 is intuitive since the Driscoll and Kraay (1998, p. 552) “covariance matrix estimator is precisely the standard Newey and West (1987) heteroskedasticity and serial correlation consistent covariance matrix estimator, applied to the sequence of cross-sectional averages” of the moment conditions.

### 2.2.3 Using the GPS-model to obtain portfolio sorts

In empirical research, the portfolio sorts procedure often extends beyond a single group of firms, as researchers frequently form five or ten portfolios defined by specific characteristics. In many studies, these portfolios are built using a single firm characteristic (e.g., the book-to-market ratio), although double or higher-dimensional sorts based on multiple firm characteristics are also common. For each portfolio, the researcher then evaluates its risk-adjusted performance using a Jensen (1968), Fama and French (1993, 2015), or Carhart (1997) style time-series regression, as outlined in Section 2.2.1.

---

<sup>7</sup> In our empirical analysis, we also consider unbalanced panels (i.e., time-varying cross-sections) and value weighted portfolios. In Section 3, we demonstrate that our theoretical results also hold in this more general setup.

Formally, in this broader setup, the first step groups firms into characteristic-based portfolios  $p$  (for  $p = 1, \dots, P$ ) whose average month- $t$  excess return,  $r_{pt}$ , is:

$$r_{pt} = \sum_{i=1}^{N_t} w_{it}^{(p)} z_{it}^{(p)} r_{it} \quad (5)$$

where  $r_{it}$  is the period- $t$  excess return of firm  $i$ ,  $w_{it}^{(p)}$  is the beginning-of-period  $t$  weight of firm  $i$  in portfolio  $p$ , and  $z_{it}^{(p)}$  is a dummy variable equal to one if firm  $i$  belongs to portfolio  $p$ , and zero otherwise. For each portfolio  $p$ , the weights sum to  $\sum_{i=1}^{N_t} w_{it}^{(p)} z_{it}^{(p)} = 1$ , and the period  $t$  cross-section comprises  $N_t = \sum_{p=1}^P \sum_{i=1}^{N_t} z_{it}^{(p)}$  firms. The second step evaluates the (risk-adjusted) performance  $\beta_{p,0}$  via a  $K$ -factor time-series regression with  $r_{pt}$  from (5) as the dependent variable:

$$r_{pt} = \beta_{p,0} + \beta_{p,1}x_{1t} + \dots + \beta_{p,K}x_{Kt} + \varepsilon_{pt} \quad (6)$$

If the coefficient estimate  $\hat{\beta}_{p,0}$  is positive (negative) and statistically significantly different from zero, portfolio  $p$  has a positive (negative) alpha.

Crucially, GPS-model (1) can replicate the results from time-series regression (6) for every portfolio  $p$  ( $p = 1, \dots, P$ ) by estimating a single firm-level regression with (weighted) pooled OLS. To this end, we set  $\mathbf{x}_t = [1 \ x_{1t} \ \dots \ x_{Kt}]$  and  $\mathbf{z}_{it} = [z_{it}^{(1)} \ z_{it}^{(2)} \ \dots \ z_{it}^{(P)}]$  to obtain

$$\begin{aligned} r_{it} &= \left( \begin{bmatrix} z_{it}^{(1)} & z_{it}^{(2)} & \dots & z_{it}^{(P)} \end{bmatrix} \otimes [1 \ x_{1t} \ \dots \ x_{Kt}] \right) \boldsymbol{\theta} + v_{it} \\ &= \theta_{1,0} z_{it}^{(1)} + \theta_{1,1} x_{1t} z_{it}^{(1)} + \dots + \theta_{1,K} x_{Kt} z_{it}^{(1)} \\ &+ \dots \\ &+ \theta_{P,0} z_{it}^{(P)} + \theta_{P,1} x_{1t} z_{it}^{(P)} + \dots + \theta_{P,K} x_{Kt} z_{it}^{(P)} + v_{it} \end{aligned} \quad (7)$$

Under the assumptions of Proposition 1, and if portfolios  $p = 1, \dots, P$  remain constant over time (i.e.,  $z_{it}^{(p)} = z_i^{(p)}$  for all  $t$ ), we obtain:<sup>8</sup>

---

<sup>8</sup> For Proposition 2 to hold in the general case of an unbalanced panel with time-varying portfolios, GPS-model (7) needs to be estimated with weighted pooled OLS. Details on the weighting scheme reproducing the results of portfolio sorts with value-weighted portfolios are provided in Section 3.1.

### Proposition 2 (Portfolio sorts)

- **Part A – Coefficient estimates.** For each portfolio  $p \in \{1, \dots, P\}$ , pooled OLS estimates  $\hat{\theta}_{p,k}$  in (7) match  $\hat{\beta}_{p,k}$  from time-series regression (6), i.e.,  $\hat{\theta}_{p,k} \equiv \hat{\beta}_{p,k}$  ( $\forall k = 0, 1, \dots, K$ ).
- **Part B – Standard errors.** For a given lag length  $H$ , Driscoll and Kraay (1998) standard errors for  $\hat{\theta}_{p,k}$  ( $k = 0, 1, \dots, K$  and  $p = 1, \dots, P$ ) in (7) coincide with Newey and West (1987) standard errors for  $\hat{\beta}_{p,k}$  in (6).

Proof: See Internet Appendix A.2.

According to Proposition 2, the coefficient estimates in (7) have a clear economic interpretation:  $\hat{\theta}_{p,0}$  ( $p = 1, \dots, P$ ) measures the alpha of portfolio  $p$ , and  $\hat{\theta}_{p,k}$  (with  $k = 1, \dots, K$ ) represents portfolio  $p$ 's exposure to factor  $k$ . Note that GPS-model (7) replicates a set of  $P$  independent time-series regressions (6) through one linear regression at the individual firm level. Consequently, a standard Wald test can be used to verify joint significance of the  $P$  portfolios' alphas:

$$\begin{aligned} H_0: \theta_{1,0} = \theta_{2,0} = \dots = \theta_{P,0} = 0 \quad \text{vs.} \\ H_1: \theta_{p,0} \neq 0 \text{ for at least one } p \in \{1, \dots, P\} \end{aligned} \tag{8}$$

This multiple hypothesis test provides an alternative to the widely used Gibbons, Ross, and Shanken (1989) or “GRS” test. Estimating (7) with Driscoll and Kraay (1998) standard errors ensures valid statistical inference even if  $v_{it}$  is heteroskedastic, autocorrelated, and cross-sectionally dependent.

#### 2.2.4 Performance differences between two portfolios

The portfolio sorts approach is also widely used to test whether the alpha of a top group portfolio differs from that of a bottom group portfolio. For example, to analyze whether the decile-portfolio of the highest book-to-market firms (“value stocks”) outperforms the decile of the lowest book-to-market firms (“growth stocks”), the first step computes average month- $t$  excess returns for both top and bottom group portfolios:

$$r_{p,t} = \sum_{i=1}^{N_t} w_{it}^{(p)} z_{it}^{(p)} r_{it} \quad (9)$$

where  $r_{it}$  is the period  $t$  excess return of firm  $i$ ,  $z_{it}^{(p)}$  is a dummy variable equal to one if firm  $i$  belongs to group  $p$  ( $p = \text{high}, \text{low}$ ), and zero otherwise.<sup>9</sup> In both portfolios, weights sum to  $\sum_{i=1}^{N_t} w_{it}^{(p)} z_{it}^{(p)} = 1$ . The difference in returns between the two portfolios is:

$$\Delta r_{p,t} = r_{\text{high},t} - r_{\text{low},t} \quad (10)$$

The second step then evaluates the risk-adjusted performance of the zero-investment portfolio (10) using a  $K$ -factor time-series regression:

$$\Delta r_{p,t} = \beta_{\Delta 0} + \beta_{\Delta 1} x_{1t} + \dots + \beta_{\Delta K} x_{Kt} + \varepsilon_{\Delta t} \quad (11)$$

If  $\hat{\beta}_{\Delta 0}$  exceeds zero (and is significant), the “*high*” portfolio is considered to outperform the “*low*” portfolio.

Estimating GPS-model (1) with (weighted) pooled OLS, we can replicate the results of (11). For that purpose, we confine the sample to the top and bottom group portfolios and set  $\mathbf{z}_{it} = [1 \ z_{it}^{(\text{high})}]$  and  $\mathbf{x}_t = [1 \ x_{1t} \ \dots \ x_{Kt}]$  to obtain:

$$\begin{aligned} r_{it} &= (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + v_{it} \\ &= \theta_{\text{low},0} + \theta_{\text{low},1} x_{1t} + \dots + \theta_{\text{low},K} x_{Kt} \\ &\quad + \theta_{\Delta 0} z_{it}^{(\text{high})} + \theta_{\Delta 1} x_{1t} z_{it}^{(\text{high})} + \dots + \theta_{\Delta K} x_{Kt} z_{it}^{(\text{high})} + v_{it} \end{aligned} \quad (12)$$

Under the assumptions of Proposition 1, and assuming the “*high*” and “*low*” portfolios remain constant over time, the following holds:<sup>10</sup>

---

<sup>9</sup> For simplicity but without loss of generality, we label the portfolios as “high” and “low” here. However, subscript  $p$  could also refer to “IPO firms” and “mature firms”, “firms with a female CEO” and “firms with a male CEO”, or any other set of two portfolios that are meant to be compared.

<sup>10</sup> For Proposition 3 to hold in the general case of an unbalanced panel with time-varying portfolios, regression (12) has to be estimated with weighted pooled OLS. See Section 3.3 for details.

### Proposition 3 (Performance difference between two groups)

#### Part A – Coefficient estimates.

Pooled OLS estimates  $\hat{\theta}_{\Delta k}$  in GPS-model (12) match  $\hat{\beta}_{\Delta k}$  in time-series regression (11), i.e.,  $\hat{\theta}_{\Delta k} \equiv \hat{\beta}_{\Delta k} (\forall k = 0, 1, \dots, K)$ . Likewise,  $\hat{\theta}_{low,k}$  in (12) matches  $\hat{\beta}_{low,k}$  from (6) for portfolio  $p$ ="low".

#### Part B – Standard errors.

For a given lag length  $H$ , Driscoll and Kraay (1998) standard errors for  $\hat{\theta}_{\Delta k}$  ( $\forall k = 0, 1, \dots, K$ ) in (12) coincide with Newey and West (1987) standard errors for  $\hat{\beta}_{\Delta k}$  in (11). Similarly, Driscoll and Kraay (1998) standard errors for  $\hat{\theta}_{low,k}$  in (12) match the Newey and West (1987) standard errors for  $\hat{\beta}_{low,k}$  in time-series regression (6) for portfolio  $p$ ="low".

Proof: See Internet Appendix A.3.

While Proposition 3 shows how to use the GPS-model for analyzing performance differences between two portfolios, the GPS-model can be further generalized to compare a base portfolio (e.g., portfolio  $p = 1$ ) with any other (e.g., quintile or decile) portfolio. Assuming that portfolio  $p = 1$  is the base portfolio, we set  $\mathbf{z}_{it} = [1 \ z_{it}^{(2)} \ \dots \ z_{it}^{(P)}]$  and  $\mathbf{x}_t = [1 \ x_{1t} \ \dots \ x_{Kt}]$  to obtain:

$$\begin{aligned}
 r_{it} &= (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + v_{it} \\
 &= \theta_{1,0} + \theta_{1,1} x_{1t} + \dots + \theta_{1,K} x_{Kt} \\
 &+ \theta_{\Delta(2 \rightarrow 1),0} z_{it}^{(2)} + \theta_{\Delta(2 \rightarrow 1),1} x_{1t} z_{it}^{(2)} + \dots + \theta_{\Delta(2 \rightarrow 1),K} x_{Kt} z_{it}^{(2)} \\
 &+ \dots \\
 &+ \theta_{\Delta(P \rightarrow 1),0} z_{it}^{(P)} + \theta_{\Delta(P \rightarrow 1),1} x_{1t} z_{it}^{(P)} + \dots + \theta_{\Delta(P \rightarrow 1),K} x_{Kt} z_{it}^{(P)} + v_{it}
 \end{aligned} \tag{13}$$

A direct consequence of Propositions 2 and 3 is:

**Corollary 1 (Relative performance versus a base portfolio)**

$\hat{\theta}_{\Delta(p \rightarrow 1),k}$  from (13) equals  $\hat{\beta}_{\Delta k}$  in (11) ( $\forall k = 0, 1, \dots, K$ ). Furthermore, Driscoll and Kraay (1998) standard errors for  $\hat{\theta}_{\Delta(p \rightarrow 1),k}$  from (13) match the Newey and West (1987) standard errors for  $\hat{\beta}_{\Delta k}$  in (11).

■

### 2.3 Applications of the GPS-model beyond traditional portfolio sorts

Standard portfolio sorts face several drawbacks. First, they are generally limited to analyzing only a few firm characteristics (Cochrane, 2011), making it challenging to test competing hypotheses. Second, it is difficult to assess the functional relationship across multiple portfolios, so that researchers often focus on a comparison of top and bottom portfolios for simplicity (Patton and Timmermann, 2010). Third, the results in Section 2.2 imply that standard portfolio sorts implicitly assume the random effects (RE) assumption to hold. This follows from the fact that GPS-model (1) must be estimated with pooled OLS (which is only consistent under the RE assumption) in order to replicate the outcomes of portfolio sorts. Consequently, if firm-specific effects exist and correlate with the explanatory variables, apparent characteristic-based return predictability may reflect long-lived firm-level traits rather than genuine return predictability.

The GPS-model proposed in this paper does not suffer from these limitations and therefore supports research questions beyond the scope of standard portfolio sorts. We now consider a series of such applications. First, we show how to interpret results from GPS-models that incorporate multiple (or continuous) firm characteristics. Then, we discuss consistent estimation of the GPS-model in the presence of firm-specific effects, and we propose a Hausman (1978)-type specification test to detect whether persistent firm-level heterogeneity is conflated with true characteristic-based predictability. This test clarifies when fixed effects are indispensable and thereby ensures that stable firm attributes do not overshadow genuine return predictors.

### 2.3.1 Multivariate and continuous firm characteristics

GPS-model (1) can be specified so that  $\mathbf{z}_{it}$  contains multiple binary or continuous firm characteristics. When specified this way, the GPS-model addresses the “multidimensional challenge” of standard portfolio sorts (Cochrane, 2011) and provides a straightforward framework for formal tests of competing hypotheses and for robustness checks.

Despite this flexibility, the GPS-model retains a clear economic interpretation even when multiple firm characteristics are included. To illustrate this, we consider a CAPM factor structure and two firm characteristics.<sup>11</sup> Let  $r_{mt}$  be the period- $t$  market return in excess of the risk-free rate, and specify  $\mathbf{x}_t = [1 \ r_{mt}]$  and  $\mathbf{z}_{it} = [1 \ z_{1,it} \ z_{2,it}]$ . Assuming no firm-specific effects ( $c_i = 0 \ \forall i$ ), we obtain:

$$\begin{aligned} r_{it} &= (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + v_{it} = ([1 \ z_{1,it} \ z_{2,it}] \otimes [1 \ r_{mt}]) \boldsymbol{\theta} + v_{it} \\ &= (\theta_{\alpha 0} + \theta_{\alpha 1} z_{1,it} + \theta_{\alpha 2} z_{2,it}) + (\theta_{\beta 0} + \theta_{\beta 1} z_{1,it} + \theta_{\beta 2} z_{2,it}) \times r_{mt} + v_{it} \quad (14) \\ &\equiv \alpha_{it} + \beta_{it} \times r_{mt} + v_{it} \end{aligned}$$

where  $\alpha_{it} = \theta_{\alpha 0} + \theta_{\alpha 1} z_{1,it} + \theta_{\alpha 2} z_{2,it}$  and  $\beta_{it} = \theta_{\beta 0} + \theta_{\beta 1} z_{1,it} + \theta_{\beta 2} z_{2,it}$ . Thus, the GPS-model decomposes the risk-adjusted performance  $\alpha_{it}$  and the factor exposure  $\beta_{it}$  according to the firm characteristics in  $\mathbf{z}_{it}$ . Hence,  $\alpha_{it}$  and  $\beta_{it}$  in (14) represent *conditional* measures. For instance, suppose an estimation of (14) yields:

$$\begin{aligned} r_{it} &= \hat{\alpha}_{it} + \hat{\beta}_{it} \times r_{mt} + v_{it} \\ &= (0.2 + 0.5z_1 - 0.8z_2) + (0.8 - 0.3z_1 + 0.1z_2) \times r_{mt} + v_{it} \end{aligned}$$

In this case, the alpha loads positively on firm characteristic  $z_1$  and negatively on  $z_2$ , so that firms with higher  $z_1$  and lower  $z_2$  have higher  $\hat{\alpha}_{it}$ . The beta, on the other hand, loads negatively on firm characteristic  $z_1$  and positively on  $z_2$ .

When persistent firm-specific effects are present and correlate with the firm characteristics in  $\mathbf{z}_{it}$ , (weighted) pooled OLS no longer yields consistent estimates. In this case, GPS-model (14) must be estimated using the fixed effects (FE) estimator to ensure

---

<sup>11</sup> Note that the specification of vectors  $\mathbf{x}_t$  and  $\mathbf{z}_{it}$  can easily be extended to comprise multiple factor variables or firm characteristics without changing the logic of how to interpret the results from estimating a GPS-model.

valid statistical inference. Importantly, this does not affect the interpretation of the GPS-model results. The key coefficients remain  $\hat{\theta}_{\alpha,1}$  and  $\hat{\theta}_{\alpha,2}$ , which measure by how much the Jensen alpha changes if  $z_{1,it}$  or  $z_{2,it}$  shifts by one unit. By contrast,  $\hat{\theta}_{\alpha,0}$  depends on the sample means of  $z_{1,it}$  and  $z_{2,it}$ , making it less informative. Under FE estimation, the intercept  $\hat{\theta}_{\alpha,0}$  drops out via the within-transformation, while the slope coefficients of interest remain consistently estimated. Hence, the GPS-model's economic interpretation does not change, regardless of whether it is estimated with or without firm fixed effects.

### 2.3.2 Portfolio sorts specification test as a diagnostic tool for factor validation

In this section, we derive a Hausman (1978) type “portfolio sorts specification test” to analyze whether the RE assumption holds in a conventional portfolio sorts setup. For this purpose, we start with GPS-model (12), thereby confining the sample to the top and bottom group portfolios, and denoting a dummy variable that is equal to one if firm  $i$  belongs to the top group portfolio (e.g., the top quintile or top decile portfolio) by  $z_{it}^{(high)}$ :

$$\begin{aligned} r_{it} &= (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + c_i + v_{it} \\ &= \left( \begin{bmatrix} 1 & z_{it}^{(high)} \end{bmatrix} \otimes \begin{bmatrix} 1 & x_{1t} & \dots & x_{Kt} \end{bmatrix} \right) \boldsymbol{\theta} + c_i + v_{it} \end{aligned}$$

Using Mundlak's (1978) correlated RE assumption, we then model  $c_i$  as

$$c_i = \left( \sum_{k=0}^K \xi_k \bar{q}_{ki} \right) + u_i \equiv \bar{\mathbf{q}}_i \boldsymbol{\xi} + u_i \quad (15)$$

where  $u_i$  is a mean-zero firm-specific effect uncorrelated with  $(\mathbf{z}_{it} \otimes \mathbf{x}_t)$ , and  $\bar{q}_{ki}$  is firm  $i$ 's time-series average of the interaction between dummy variable  $z_{it}^{(high)}$  and

factor variable (or constant)  $x_{kt}$ . We collect all such time-series averages in row vector  $\bar{q}_i$  and store the  $\xi_k$  coefficients in column vector  $\xi$ .<sup>12</sup>

Replacing  $c_i$  in GPS-model (12) with the expression in (15) yields:

$$r_{it} = (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + \bar{q}_i \boldsymbol{\xi} + v_{it} \quad (16)$$

where  $v_{it} = u_i + v_{it}$  is a mean zero disturbance that is uncorrelated with  $(\mathbf{z}_{it} \otimes \mathbf{x}_t)$ . As shown in Wooldridge (2010, Chapter 10.7.3), regression (16) forms a regression-based variant of the Hausman (1978) specification test. The pooled OLS estimates for  $\boldsymbol{\theta}$  represent the fixed effects (or “within”) coefficient estimates, while the pooled OLS estimates for  $\boldsymbol{\xi}$  capture by how much the between estimator differs from the fixed effects estimator. To ensure robust inference, we estimate (16) with Driscoll and Kraay (1998) standard errors that are robust to cross-sectional dependence, autocorrelation, and heteroskedasticity. A test of

$$H_0: \boldsymbol{\xi} = 0 \quad \text{vs.} \quad H_1: \boldsymbol{\xi} \neq 0 \quad (17)$$

in (16) then represents our “portfolio sorts specification test”. If the null hypothesis of  $\boldsymbol{\xi} = 0$  is not rejected, the RE assumption can be accepted, and standard portfolio sorts are likely valid. However, if the null hypothesis is rejected, the analysis suggests that part of the alpha stems from stable, firm-level traits rather than from the sorting characteristic itself, indicating that a GPS-model with firm fixed effects should be used instead. Our portfolio sorts specification test thus provides a novel diagnostic tool for addressing the factor zoo (Cochrane, 2011). By clarifying whether the random effects (RE) assumption holds, the test helps distinguishing factors that genuinely predict stock returns from those whose apparent predictability arises primarily from correlation with persistent firm-level attributes.

### 3. Empirical Analysis using the GPS-model

In this section, we begin by empirically validating Propositions 1 through 3 from Section 2. For illustrative purposes and ease of exposition, we initially focus on a single

---

<sup>12</sup> In an unbalanced panel, the time-series averages of the factor variables in vector  $\mathbf{x}_t$  vary across firms. Therefore, when working with an unbalanced panel, they have to be included in vector  $\bar{q}_i$  as well.

firm characteristic – operating profitability (Ball et al., 2015) – to demonstrate how the GPS-model replicates standard portfolio sorts and highlights the role of persistent firm-level heterogeneity. We then assess the broader implications of unobserved heterogeneity in portfolio-based asset pricing tests by estimating multiple GPS-model specifications (with and without firm fixed effects) for the 94 firm-level asset pricing predictors in Gu et al. (2020, 2021). All our empirical analyses use the merged CRSP-Compustat database (stock returns and market capitalization), combined with the asset pricing predictors from Dacheng Xiu’s website.<sup>13</sup> and the market-level factor returns from Kenneth French’s data library.<sup>14</sup> The sample period extends from July 1963 to December 2019.

### 3.1 The performance of a single portfolio

To empirically validate Proposition 1, we begin with the standard portfolio sorts approach. Each month, we sort stocks into quintiles based on NYSE breakpoints and hold the resulting portfolios for the subsequent month. For each portfolio  $p$  ( $p = 1, \dots, 5$ ), we compute monthly value-weighted excess returns ( $r_{pt}$ ) as

$$r_{pt} = \sum_{i=1}^{N_t} w_{it}^{(p)} \text{operprof}_{it}^{(p)} r_{it} \quad (18)$$

where  $w_{it}^{(p)}$  is the beginning-of-month  $t$  weight of stock  $i$  in quintile portfolio  $p$ ,  $r_{it}$  is stock  $i$ ’s month- $t$  excess return,  $\text{operprof}_{it}^{(p)}$  is a dummy variable equal to one if stock  $i$  belongs to operating profitability portfolio  $p$  in month  $t$ , and  $N_t$  is the total number of stocks that month.

Using  $r_{pt}$  from (18) as the dependent variable, we estimate the Fama and French (1993) three-factor model:<sup>15</sup>

<sup>13</sup> Data on the 94 stock-level predictor variables can be retrieved at: <https://dachxiu.chicagobooth.edu/>

<sup>14</sup> Data on the Fama and French factors are from [https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

<sup>15</sup> For ease of exposition, we restrict the analysis in Sections 3.1 to 3.5 to the Fama and French (1993) three factor model. When demonstrating the importance of firm fixed effects for characteristics-based asset pricing tests more broadly in Section 3.6, we consider a wide range of factor models including the market model, the Fama and French (1993, 2015) three- and five-factor models, and the Carhart (1997) model.

$$r_{pt} = \alpha + \beta_{RMRF} RMRF_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \varepsilon_t \quad (19)$$

where  $RMRF_t$  is the market excess return,  $SMB_t$  is the return of a zero-investment size portfolio, and  $HML_t$  is the return of a zero-investment book-to-market portfolio.

According to Proposition 1, the results from estimating time-series regression (19) for portfolio  $p$  can be replicated by estimating the following GPS-model (1):

$$r_{it} = \theta_\alpha + \theta_{RMRF} RMRF_t + \theta_{SMB} SMB_t + \theta_{HML} HML_t + v_{it} \quad (20)$$

As we consider a single portfolio  $p$ , GPS-model (20) only includes observations for which  $operprof_{it}^{(p)} = 1$  (i.e., the stocks in portfolio  $p$ ). To match the portfolio sorts methodology, we must account for the following. First, the portfolio returns in (18) are value-weighted returns (not equal-weighted returns). Second, the CRSP-Compustat database forms an unbalanced panel with time-varying cross-sections. Hence, we estimate GPS-model (20) by weighted pooled OLS, aligning the observation weights with the (implicit) weighting scheme in (18). Specifically, we set:

$$w_{it} = \frac{operprof_{it}^{(p)} ME_{it}}{\sum_{i=1}^{N_t} (operprof_{it}^{(p)} ME_{it})} \quad (21)$$

where  $ME_{it}$  is the beginning-of-month  $t$  market value of equity for stock  $i$ .

Table 1 reports the results for the top-quintile portfolio sorted on operating profitability. The first row (“Portfolio sorts approach”) shows the results from OLS estimation of time-series regression (19), with Newey and West (1987) standard errors (lag-length of three months). In line with Ball et al. (2015), this top-quintile portfolio has a statistically significant Fama and French (1993) alpha of +0.21% per month, or +2.47% per year. The second row (“GPS-model (weighted pooled OLS)”) presents estimates of GPS-model (20) under weighted pooled OLS, using Driscoll and Kraay (1998) standard errors (also with a lag-length of three months). The estimates match those from time-series regression (19), empirically validating Proposition 1 and demonstrating that GPS-model (20) precisely replicates the portfolio sorts approach when the RE assumption holds.

### 3.2 Analyzing portfolio sorts with the GPS-model

According to Proposition 2, GPS-model (1) can be specified so that it exactly reproduces the results of multiple sorted portfolios through a single firm-level regression. To validate Proposition 2 empirically, we apply the portfolio sorts procedure from Section 3.1 to each of the five operating profitability quintile portfolios. Panel A of Table 2 shows the results. Consistent with Ball et al. (2015), the alphas increase monotonically with operating profitability. While the lowest-profitability ( $p = 1$ ) portfolio has a statistically significant negative alpha of -0.32% per month, the top-quintile ( $p = 5$ ) portfolio exhibits a statistically significant positive alpha of +0.21% per month.

To replicate these results from portfolio sorts using the GPS-model, we estimate regression (7) via weighted pooled OLS. We set the weight of observation  $it$  equal to stock  $i$ 's value-weight in the respective operating profitability portfolio (see equation 21). Panel B of Table 2 presents the results in a two-dimensional matrix. The portfolio dummy variables in  $\mathbf{z}_{it} = \left[ \text{operprof}_{it}^{(1)} \quad \text{operprof}_{it}^{(2)} \quad \dots \quad \text{operprof}_{it}^{(5)} \right]$  define the columns, and the factor variables in  $\mathbf{x}_t = [1 \quad \text{RMRF}_t \quad \text{SMB}_t \quad \text{HML}_t]$  define the rows. Each element of this matrix represents a coefficient estimate (and its  $t$ -statistic) for the interaction between  $\text{operprof}_{it}^{(p)}$  and factor variable (or constant)  $x_{kt}$ .

A comparison of Panels A and B in Table 2 shows that both the estimates and  $t$ -statistics (based on Driscoll-Kraay standard errors) from the GPS-model match those from the portfolio sorts approach. This finding supports Proposition 2, demonstrating that, when the random effects (RE) assumption holds, a single firm-level regression can fully replicate the alpha and factor exposures obtained from multiple sorted portfolios.

### 3.3 Top versus bottom portfolio performance

To empirically confirm Proposition 3, we first apply the standard portfolio sorts approach to measure the performance difference between the top ( $p = 5$ ) and bottom ( $p = 1$ ) quintile portfolios sorted on operating profitability. To this end, we first compute the month  $t$  excess returns ( $r_{pt}$ ) for the top and bottom quintile portfolios as outlined in expression (18) above. We then estimate time-series regression (19) using  $\Delta r_{pt} = r_{5,t} - r_{1,t}$  as the dependent variable. Panel A of Table 3 reports the results. On

a risk-adjusted basis, the portfolio of stocks with high operating profitability outperforms the low-profitability portfolio by a significant +0.52% per month (or +6.29% per year). Despite a slightly different sample period, the coefficient estimates and  $t$ -values are very similar to those reported in Table 6 of Ball et al. (2015).

To replicate these portfolio sorts results via a regression on the individual firm-level, we estimate GPS-model (12) with weighted pooled OLS. Here, we restrict the sample to stocks with top ( $p = 5$ ) or bottom ( $p = 1$ ) quintile operating profitability and set the observation weights equal to the stocks' value-weights in their operating profitability portfolio. Panel B of Table 3 reports the results in a two-dimensional matrix. The elements of vector  $\mathbf{z}_{it} = [1 \text{ } operprof_{it}^{(5)}]$  define the columns while the factor variables in  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]$  identify the rows. All elements in the results matrix thus represent coefficient estimates (and  $t$ -statistics based on Driscoll-Kraay standard errors) for interactions between the elements in  $\mathbf{z}_{it}$  and those in  $\mathbf{x}_t$ . The GPS-model estimates in Panel B match the portfolio sorts results in Panel A, thereby empirically validating Proposition 3.

### 3.4 Testing the validity of standard portfolio sorts

As demonstrated above, GPS-model (1) estimated with (weighted) pooled OLS can replicate the results of standard portfolio sorts. This implies, however, that portfolio sorts remain valid only if the random effects (RE) assumption holds. We now examine whether any firm-specific effects (if present) are indeed uncorrelated with the characteristic(s) used for portfolio formation. To test the validity of standard portfolio sorts, we employ the “portfolio sorts specification test” from Section 2.3.2.

To this end, we confine the sample to stocks with top ( $p=5$ ) or bottom ( $p=1$ ) quintile operating profitability and estimate the correlated RE model (16) by specifying  $\mathbf{z}_{it} = [1 \text{ } operprof_{it}^{(5)}]$  and  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]$ , and by adding the firm-specific time-series averages (collected in  $\bar{\mathbf{q}}_i$ ) of all variables in  $(\mathbf{z}_{it} \otimes \mathbf{x}_t)$  to the regression. Table 4 reports the results of estimating the resulting regression  $r_{it} = (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + \bar{\mathbf{q}}_i \boldsymbol{\xi} + \varepsilon_{it}$  with pooled OLS, weighting observations by each stock's beginning-of-month  $t$  value-weight in the operating profitability quintile portfolios.

The results are reported in matrix form. The first block of results, labeled “Coefficient estimates on explanatory variables  $\mathbf{z}_{it} \otimes \mathbf{x}_t$ ”, provides the estimation results for vector  $\boldsymbol{\theta}$  in similar format as in Panel B of Table 3. The second block of Table 4, labeled “Coefficient estimates on firm-specific time-series averages”, reports the results for vector  $\boldsymbol{\xi}$ . As discussed in Section 2.3.2, standard portfolio sorts remain valid if the coefficients in  $\boldsymbol{\xi}$  are jointly equal to zero. However, if the null hypothesis of  $\boldsymbol{\xi} = 0$  is rejected, the random effects (RE) assumption fails, likely invalidating standard portfolio sorts. Applying Wald test (17) yields an F-statistic of 3.04, rejecting  $\boldsymbol{\xi} = 0$  at all conventional levels of statistical significance. Thus, for portfolio sorts on operating profitability, the RE assumption does not hold, necessitating a GPS-model (1) that includes firm fixed effects. Taken together, Tables 3 and 4 illustrate that firm fixed effects (and thus unobservable heterogeneity across firms) can significantly affect the results of empirical asset pricing tests, underscoring the need to control for persistent firm-level components.

Figure 1 confirms this graphically. Panel A plots the GPS-model-estimated firm fixed effects against operating profitability, while Panel B does the same for an investment factor (defined as capital expenditures plus inventory), shown elsewhere to explain the cross-section of returns (e.g., Fama and French, 2015; Hou et al., 2015). Operating profitability is strongly correlated with the firm fixed effects ( $\rho = 0.34$ ), whereas the investment factor shows only a weak correlation ( $\rho = 0.02$ ). Accordingly, we expect the investment factor to remain a more robust predictor than operating profitability once firm fixed effects enter the analysis. Indeed, re-estimating GPS-models under the FE estimator confirms that while operating profitability loses significance, the investment factor persists as a genuine predictor of the cross-section of average returns (see subsequent Section 3.5).

### *3.5 Firm fixed effects and continuous and multivariate firm characteristics*

In this section, we illustrate how the GPS-model provides a versatile framework for investigating stock return predictability and assessing the role of firm fixed effects. To demonstrate the handling of multiple and continuous firm characteristics, we augment our model with two additional firm characteristics:

$$\begin{aligned}
r_{it} &= (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + e_{it} \\
&= ([1 \text{ } operprof_{it} \text{ } invest_{it} \text{ } mom12m_{it}] \\
&\quad \otimes [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]) \boldsymbol{\theta} + e_{it}
\end{aligned} \tag{22}$$

where  $operprof_{it}$  is operating profitability,  $invest_{it}$  measures capital expenditures and inventory, and  $mom12m_{it}$  is a 12-month momentum factor. The firm-characteristics in vector  $\mathbf{z}_{it}$  relate to factors included in popular factor models other than the Fama and French (1993) factor model, for instance in the Fama and French (2015) five-factor model and the Carhart (1997) four-factor model.<sup>16</sup> The variables in  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]$  consist of a constant and the three Fama and French (1993) factors.

Panel A of Table 5 presents the results from estimating GPS-model (22) with weighted pooled OLS, where observation weights are set equal to the firms' month  $t$  value weights. For brevity, we focus on the coefficients that decompose the Fama-French three-factor model alpha.<sup>17</sup> In Column (1), we estimate regression (22) with  $\mathbf{z}_{it} = [1 \text{ } operprof_{it}]$ . The results show that alpha increases with operating profitability. In line with Ball et al. (2015), the coefficient estimate is positive (+1.74) and highly statistically significant. In Column (2), we estimate GPS-model (22) with  $\mathbf{z}_{it} = [1 \text{ } invest_{it}]$ . Here the coefficient estimate is negative and significant, consistent with Chen et al. (2011) and Fama and French (2015). In Column (3), we estimate GPS-model (22) with  $\mathbf{z}_{it} = [1 \text{ } mom12m_{it}]$  and find alpha to increase in the momentum factor, a finding consistent with the empirical literature on momentum strategies and factors (e.g., Jegadeesh and Titman, 1993). Finally, Column (4) takes advantage of the GPS-model's capability to handle multiple firm characteristics. Estimating regression (22) with  $\mathbf{z}_{it} = [1 \text{ } operprof_{it} \text{ } invest_{it} \text{ } mom12m_{it}]$ , yields similar coefficients and  $t$ -statistics as when estimated individually in Columns (1) to (3).

---

<sup>16</sup> While the selection of firm characteristics clearly is arbitrary, this analysis only serves illustrative purposes. In Section 3.6, we empirically test the robustness of all 94 firm characteristics in the Gu et al. (2020, 2021) database.

<sup>17</sup> Note that conditional on firm  $i$ 's characteristics in period  $t$ , the Fama-French three-factor model alpha is obtained as  $\hat{\alpha}_{it} = \hat{\theta}_0 + \hat{\theta}_1 \times operprof_{it} + \hat{\theta}_2 \times invest_{it} + \hat{\theta}_3 \times mom12m_{it}$ . For details, see Section 2.3.1.

However, the analysis in Panel A of Table 5 is only valid if the random effects (RE) assumption holds. In Panel B, we re-run the estimations from Panel A using the fixed effects (FE) estimator, which accommodates firm-specific effects that may correlate with the explanatory variables in  $(\mathbf{z}_{it} \otimes \mathbf{x}_t)$ . Consistent with the results in Table 3, Column (5) shows that operating profitability no longer significantly predicts the cross-section of average stock returns once firm fixed effects are included. By contrast, both the investment and momentum factors remain robust predictors even under FE estimation (Columns (6) and (7)). In Column (8), we estimate the GPS-model with multiple characteristics via the FE estimator, confirming the single-characteristic results in Columns (5) to (7).

### *3.6 The relevance of firm fixed effects in empirical asset pricing tests*

In this section, we assess the broader importance of controlling for firm fixed effects – and thus long-lived firm-level attributes – when predicting the cross-section of average stock returns. Specifically, we ask how many of the 94 firm-level asset pricing predictors from Gu et al. (2020, 2021), that significantly predict the cross-section of stock returns in a standard portfolio sorts analysis, lose their apparent predictive power (or change sign) once we account for firm fixed effects. In the latter scenario, the characteristic likely proxies rather for persistent firm traits than for genuine cross-sectional return predictability.

To investigate this, we exploit the GPS-model’s ability to replicate standard portfolio sorts. Specifically, we estimate individual GPS models with pooled WLS for each of the 94 characteristics in Gu et al. (2020, 2021). In these tests, we first restrict the sample to top- or bottom-quintiles (or deciles) of the relevant characteristic and specify vector  $\mathbf{z}_{it}$  to include a constant and a dummy variable equal to one for observations in the top group (either quintile or decile). As a third variant, we include all observations and use cross-sectionally standardized versions of the continuous characteristics. Moreover, to ensure that our results do not hinge on a specific factor model, we run each test series under the Fama and French (1993) three-factor model, the Fama and French (2015) five-factor model, the Carhart (1997) four-factor model, and the market model. Hence, each of the 94 characteristics is examined 12 times (four alternative factor

models times three variants of each firm characteristic). We then reestimate all models using the fixed effects (FE) estimator. This allows us to assess how many of the 94 firm characteristics, which significantly predict the cross-section of stock returns in a standard portfolio sorts setting, continue to predict stock returns when unobservable heterogeneity across firms is accounted for.

Detailed results are depicted in Figures IA.1 to IA.12 in the Internet Appendix, comparing each characteristic's alpha from standard portfolio sorts to the alpha from a GPS-model with firm fixed effects. For simplicity, Table 6 summarizes the key outcomes in a single table. Each row corresponds to one of the 12 alternative specifications: Rows (1) to (3) use the Fama and French (1993) three-factor returns in  $\mathbf{x}_t$ . Rows (4) to (6) report the results for the Fama and French (2015) five-factor model. Rows (7) to (9) apply the Carhart (1997) four-factor model, and rows (10) to (12) show the findings for the market model. Rows (1), (4), (7), and (10) limit the sample to top- or bottom-quintiles. Rows (2), (5), (8), and (11) use deciles instead of quintiles; and Rows (3), (6), (9), and (12) report the results from GPS models that include cross-sectionally standardized versions of the continuous firm characteristics in  $\mathbf{z}_{it}$ .

Results in Row (1) show that only 22 of 48 firm characteristics, which significantly predict returns in standard quintile portfolios sorts, remain significant once we include firm fixed effects. Hence, 26 firm characteristics (54%) either lose their predictive power (20, or 41%) or remain significant but switch signs (six, or 13%). When applying decile instead of quintile sorts in Row (2), the percentage of characteristics that lose significance or switch sign when firm fixed effects are included increases to 56%. In Row (3), when using cross-sectionally standardized variables, the percentage of characteristics that lose significance or switch sign when firm fixed effects are included amounts to 49%. Rows (4) to (6), in which the Fama and French (1993) three-factor model is replaced by the Fama and French (2015) five-factor model, and Rows (7) to (9), using the Carhart (1997) four-factor model, show qualitatively similar results. When using the market model in Rows (10) to (12), the proportion of characteristics that lose or reverse predictive power is slightly lower (34%–37%). The bottom row of Table 6 aggregates across all 12 rows: Of the 532 specifications that are significant under a standard portfolio sorts setup, only 270 remain significant once firm fixed

effects are included. The remaining 262 (49%) either lose their predictive power (215, or 40%) or flip sign while staying significant (47, or 9%). Nearly half of the 94 factors thus appear to proxy for persistent firm traits rather than representing genuine drivers of expected returns.

#### **4. Conclusion**

This paper shows that standard portfolio sorts can conflate genuine characteristic-driven predictability with persistent, unobserved firm-level heterogeneity. We introduce the GPS-model, a firm-level regression framework that exactly replicates results from all variants of standard portfolio sorts, yet also accommodates more complex analyses with multiple characteristics and continuous controls. Most importantly, when estimated with firm fixed effects, the GPS-model disentangles cross-sectional predictability truly stemming from firm characteristics from that arising due to persistent, firm-specific traits.

Our theoretical results show that the alpha and factor exposures obtained from standard portfolio sorts match pooled OLS estimates only under a random effects (RE) assumption – an assumption that may not always hold in practice. To detect potential violations of this assumption, we develop a Hausman-type specification test for portfolio sorts. This test helps researchers discern whether observed return predictability from a firm-characteristic genuinely reflects a new factor or rather arises from persistent, firm-level heterogeneity.

Empirically, controlling for unobserved heterogeneity can markedly alter inferences about a factor’s predictive power. For instance, while operating profitability appears to predict average returns under standard portfolio sorts, its predictive power vanishes once firm fixed effects are included. Extending this analysis to 94 predictors from Gu et al. (2020, 2021) points to broader implications: Nearly half lose their significance after accounting for firm-level heterogeneity. Hence, a substantial portion of the factor zoo may reflect stable, firm-specific influences rather than true characteristic-based predictability.

In sum, the GPS-model provides a robust and flexible framework for evaluating new factor candidates and interpreting asset pricing anomalies. By isolating genuine characteristic-based factors from persistent firm-level traits, it enhances the rigor of empirical asset pricing tests and strengthens the credibility of inferred market anomalies.

## References

- Bali, Turan G., Stephen J. Brown, and Yi Tang, 2017, Is economic uncertainty priced in the cross-section of stock returns?, *Journal of Financial Economics* 126, 471-489.
- Bali, Turan G., A. Doruk Gunaydin, Thomas Jansson, and Yigitcan Karabulut, 2023, Do the rich gamble in the stock market? Low risk anomalies and wealthy households, *Journal of Financial Economics* 150, 103715.
- Ball, Ray, Joseph Gerakos, Juhani T. Linnainmaa, and Valeri V. Nikolaev, 2015, Deflating profitability, *Journal of Financial Economics* 117, 225-248.
- Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773-806.
- Barber, Brad M., and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics* 116, 261-292.
- van Binsbergen, Jules H., and Christian C. Opp, 2019, Real anomalies, *Journal of Finance* 74, 1659-1706.
- Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The Capital Asset Pricing Model: Some empirical findings, in: Jensen, M., Ed., *Studies in the Theory of Capital Markets*, Praeger Publishers, New York, 79-124.
- Bryzgalova, Svetlana, 2016, Spurious factors in linear asset pricing models, Working Paper, Stanford University.
- Bryzgalova, Svetlana, Markus Pelger, and Jason Zhu, 2023, Forest through the trees: Building cross-sections of stock returns, *Journal of Finance*, forthcoming.
- Cameron, A. Colin, and Pravin K. Trivedi, 2005, *Microeconometrics: methods and applications*, Cambridge University Press.
- Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.
- Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Ernst Schaumburg, 2020, Characteristic-sorted portfolios: Estimation and inference, *Review of Economics and Statistics* 102, 531-551.
- Chen, Andrew Y., 2021, The limits of p-hacking: Some thought experiments, *Journal of Finance* 76, 2447-2480.

Chen, Long, Novy-Marx, Robert, and Lu Zhang, 2011, An alternative three-factor model, Working Paper, Ohio State University.

Chordia, Tarun, Amit Goyal, and Alessio Saretto, 2020, Anomalies and false rejections, *Review of Financial Studies* 33, 2134-2179.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047-1108.

Dello-Preite, Massimo, Raman Uppal, Paolo Zaffaroni, and Irina Zviadadze, 2023, What is missing in asset-pricing factor models? Working Paper, Imperial College London.

Driscoll, John C., and Aart C. Kraay, 1998, Consistent covariance matrix estimation with spatially dependent panel data, *Review of Economics and Statistics* 80, 549-560.

Ehsani, Sina, and Juhani T. Linnainmaa, 2022, Factor momentum and the momentum factor, *Journal of Finance* 77, 1877-1919.

Fama, Eugene F., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283-306.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns of stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2020, Taming the factor zoo: A test of new factors, *Journal of Finance* 75, 1327-1370.

Frazzini, Andrea, and Lasse H. Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1-25.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2020, Dissecting characteristics nonparametrically, *Review of Financial Studies* 33, 2326-2377.

Friewald, Nils, Florian Nagler, and Christian Wagner, 2022, Debt refinancing and equity returns, *Journal of Finance* 77, 2287-2329.

Fung, William, David A. Hsieh, Narayan Y. Naik, and Tarun Ramadorai, 2008, Hedge funds: Performance, risk, and capital formation, *Journal of Finance* 63, 1777-1803.

Gerakos, Joseph, and Juhani T. Linnainmaa, 2018, Decomposing value, *Review of Financial Studies* 31, 1825-1854.

Gerakos, Joseph, Juhani T. Linnainmaa, and Adair Morse, 2021, Asset managers: Institutional performance and factor exposures, *Journal of Finance* 76, 2035-2075.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Gormley, Todd A., and David A. Matsa, 2014, Common errors: How to (and not to) control for unobserved heterogeneity, *Review of Financial Studies* 27, 617-661.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *Review of Financial Studies* 33, 2223-2273.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2021, Autoencoder asset pricing models, *Journal of Econometrics* 222, 429-450.

Harvey, Campbell R., 2017, Presidential Address: The Scientific Outlook in Financial Economics, *Journal of Finance* 72, 1399-1440.

Harvey, Campbell R., and Yan Liu, 2021, Lucky factors, *Journal of Financial Economics* 141, 413-435.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, ...and the cross-section of expected returns. *Review of Financial Studies* 29, 5-68.

Hausman, Jerry A., 1978, Specification tests in econometrics. *Econometrica* 46, 1251-1271.

Heston, Steven L., Christopher S. Jones, Mehdi Khorram, Shuaiqi Li, and Haitao Mo, 2023, Option momentum, *Journal of Finance* 78, 3141-3192.

Hoechle, Daniel, 2007, Robust standard errors for panel regressions with cross-sectional dependence. *Stata Journal* 7, 281-312.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650-705.

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 2019-2133.

- Huang, Dashan, Jiaen Li, and Guofu Zhou, 2023, Shrinking factor dimension: A reduced-rank approach, *Management Science* 69, 5501-5522.
- Hsu, Po-Hsuan, Kai Li, and Chi-Yang Tsou, 2023, The pollution premium, *Journal of Finance* 78, 1343-1392.
- Hvide, H.K., and K. Meisner Nielsen, 2025, Flying below the radar: Insider trading by executives below the top, *Journal of Financial Economics*, forthcoming.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.
- Jensen, Michael C., 1968, The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 389-416.
- Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379-2416.
- Kothari, S.P., and Jerold B. Warner, 2008, Econometrics of event studies, in: Eckbo, B. Espen (ed.), *Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1*, Elsevier/North-Holland, 3-36.
- Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the Cross Section, *Journal of Financial Economics* 135, 271-292.
- Liu, Yukun, Aleh Tsyvinski, and Xi Wu, 2022, Common risk factors in cryptocurrency, *Journal of Finance* 77, 1133-1177.
- Loughran, Tim, and Jay R. Ritter, 2000, Uniformly least powerful tests of market efficiency, *Journal of Financial Economics* 55, 361-389.
- Lyon, John D., Brad M. Barber, and Chih-Ling Tsai, 1999, Improved methods for tests of long-run abnormal stock returns, *Journal of Finance* 54, 165-201.
- Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.
- Mundlak, Yair, 1978, On the pooling of time series and cross section data. *Econometrica* 46, 69-85.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Novy-Marx, Robert, 2013, The other side of value: the gross profitability premium, *Journal of Financial Economics* 108, 1-28.

Patton, Andrew J., and Allan Timmermann, 2010, Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts, *Journal of Financial Economics* 98, 605-625.

Petersen, Mitchell A., 2009, Estimating standard errors in financial panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435-480.

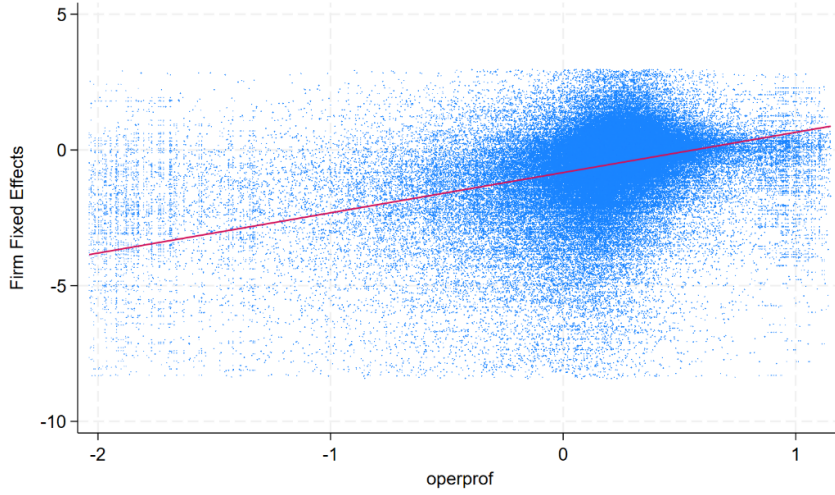
Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam, 2018, A protocol for factor identification, *Review of Financial Studies* 32, 1573-1607.

Weber, Michael, 2018, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* 128, 486-503.

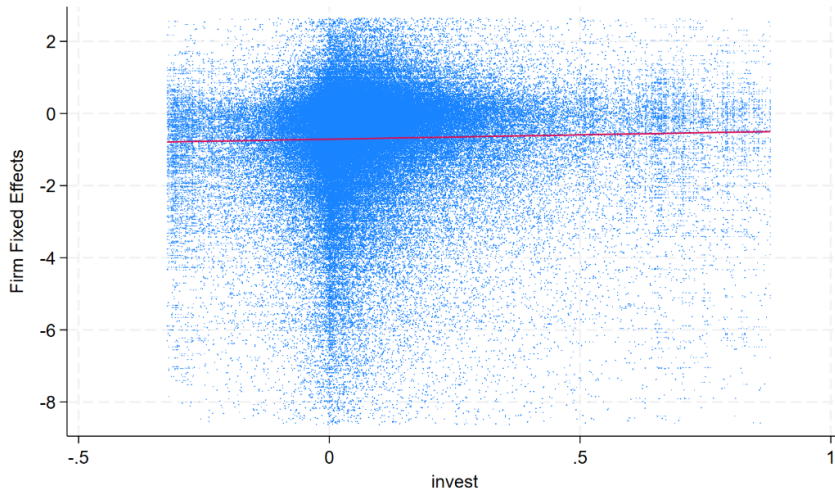
Wooldridge, Jeffrey M., 2010, *Econometric analysis of cross section and time series data*, MIT Press.

**Figure 1: Relationship between firm-specific fixed effects and firm characteristics**

Panel A: Firm fixed effects versus operating profitability (*operprof*)



Panel B: Firm fixed effects versus capital expenditures and inventory (*invest*)



This figure plots firm fixed effects versus operating profitability (Panel A) and capital expenditures and inventory (Panel B). The sample period is July 1963 to December 2019. Firm fixed effects are derived from estimating a firm-level panel regression as shown in equation (1) with the fixed effects (FE) estimator. The market-level variables in vector  $\mathbf{x}_t$  comprise a constant and the three Fama and French (1993) factors, i.e., RMRF (market excess return), SMB (small minus big), and HML (high minus low). In Panel A, regression (1) is estimated with vector  $\mathbf{z}_{it}$  containing five quintile dummy variables based on NYSE breakpoints for operating profitability (variable *operprof* of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). The correlation between the estimated firm fixed effects and *operprof* is 0.34. In Panel B, vector  $\mathbf{z}_{it}$  contains five quintile dummy variables based on NYSE breakpoints for capital expenditures and inventory (variable *invest* of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). The correlation between the estimated firm fixed effects and *invest* is 0.02. The charts do not display the top and bottom 1% of observations for the firm fixed effects, *operprof*, and *invest*, respectively.

**Table 1: Single portfolio analysis**

|                                     | Operating profitability ( <i>operprof</i> ): Top-quintile portfolio |                         |                        |                        |                               |               |                  |
|-------------------------------------|---------------------------------------------------------------------|-------------------------|------------------------|------------------------|-------------------------------|---------------|------------------|
|                                     | Three-factor model                                                  |                         |                        |                        | Three-factor model statistics |               |                  |
|                                     | <i>a</i>                                                            | <i>b<sub>RMRF</sub></i> | <i>b<sub>SMB</sub></i> | <i>b<sub>HML</sub></i> | <i>R-squared</i>              | <i>N Obs.</i> | <i>N Stocks.</i> |
| (1) Portfolio sorts approach        | 0.206***<br>(4.30)                                                  | 0.943***<br>(63.29)     | -0.128***<br>(-5.09)   | -0.146***<br>(-4.11)   | 0.919                         | 678           |                  |
| (2) GPS-model (weighted pooled OLS) | 0.206***<br>(4.30)                                                  | 0.943***<br>(63.29)     | -0.128***<br>(-5.09)   | -0.146***<br>(-4.11)   | 0.256                         | 300,778       | 6,188            |

This table reports the average return as well as the 3-factor model alpha along with RMRF (market), SMB (small minus big), and HML (high minus low) factor loadings for portfolios sorted by operating profitability (variable *operprof* of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). The portfolio sort is based on NYSE breakpoints and the sample period is July 1963 to December 2019. All results are for quintile 5 (high) portfolio, comprising stocks with the highest operating profitability. Row (1) reports results from a standard portfolio sort where the portfolio's excess return is regressed on the three Fama and French (1993) factors. Row (2) presents results from estimating GPS-model (22) with weighted pooled OLS, where observation weights are set equal to the beginning-of-time *t* value-weights of the stocks. *t*-statistics from testing for significance against a value of zero are presented in parentheses. Statistical inference for the portfolio sorts approach in Row (1) is based on Newey and West (1987) standard errors with a lag-length of three. The GPS-model in Row (2) is estimated with Driscoll and Kraay (1998) standard errors with a lag-length of three. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels (two-tailed).

**Table 2: Analysis of quintile portfolios sorted on operating profitability**

| Panel A: Standard portfolio sorts |                       |                       |                       |                       |                       |
|-----------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                                   | Quintile Portfolio    |                       |                       |                       |                       |
|                                   | 1 (low)               | 2                     | 3                     | 4                     | 5 (high)              |
| $\alpha$                          | -0.318***<br>(-3.83)  | -0.0596<br>(-1.00)    | 0.0242<br>(0.48)      | 0.0511<br>(1.26)      | 0.206***<br>(4.30)    |
| $b_{RMRF}$                        | 1.118***<br>(42.55)   | 0.945***<br>(52.81)   | 0.948***<br>(61.16)   | 0.981***<br>(62.87)   | 0.943***<br>(63.29)   |
| $b_{SMB}$                         | 0.382***<br>(11.35)   | 0.0647*<br>(1.92)     | 0.0243<br>(0.67)      | -0.101***<br>(-4.18)  | -0.128***<br>(-5.09)  |
| $b_{HML}$                         | 0.0172<br>(0.37)      | 0.0899***<br>(2.81)   | -0.0135<br>(-0.51)    | -0.0896***<br>(-4.82) | -0.146***<br>(-4.11)  |
| $R\text{-squared}$                | 0.886                 | 0.885                 | 0.911                 | 0.939                 | 0.919                 |
| $N\text{ Obs.}$                   | 678                   | 678                   | 678                   | 678                   | 678                   |
| Panel B: GPS-model                |                       |                       |                       |                       |                       |
| Vector $\mathbf{z}_t \rightarrow$ | $operprof_{it}^{(1)}$ | $operprof_{it}^{(2)}$ | $operprof_{it}^{(3)}$ | $operprof_{it}^{(4)}$ | $operprof_{it}^{(5)}$ |
| Vector $\mathbf{x}_t \downarrow$  |                       |                       |                       |                       |                       |
| 1 (Intercept)                     | -0.318***<br>(-3.83)  | -0.0596<br>(-1.00)    | 0.0242<br>(0.48)      | 0.0511<br>(1.26)      | 0.206***<br>(4.30)    |
| $RMRF_t$                          | 1.118***<br>(42.55)   | 0.945***<br>(52.81)   | 0.948***<br>(61.16)   | 0.981***<br>(62.87)   | 0.943***<br>(63.29)   |
| $SMB_t$                           | 0.382***<br>(11.35)   | 0.0647*<br>(1.92)     | 0.0243<br>(0.67)      | -0.101***<br>(-4.18)  | -0.128***<br>(-5.09)  |
| $HML_t$                           | 0.0172<br>(0.37)      | 0.0899***<br>(2.81)   | -0.0135<br>(-0.51)    | -0.0896***<br>(-4.82) | -0.146***<br>(-4.11)  |
| $R\text{-squared}$                | 0.228                 |                       |                       |                       |                       |
| $N\text{ Obs.}$                   | 1,951,536             |                       |                       |                       |                       |
| $N\text{ Stocks}$                 | 15,772                |                       |                       |                       |                       |

This table reports value-weighted 3-factor model alphas along with RMRF (market excess return), SMB (small minus big), and HML (high minus low) factor loadings for portfolios sorted by operating profitability (variable  $operprof$  of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). Panel A reports results from standard portfolio sorts, where a portfolio's excess return is regressed on the three Fama and French (1993) factors. Panel B presents results from estimating a single GPS-model with weighted pooled OLS, where observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Coefficient estimates and  $t$ -statistics (in parentheses) in Panel B (GPS-model) are for the interaction of the (market-level) variables and a constant contained in vector  $\mathbf{x}_t$  (displayed on the vertical axis) and the quintile dummy variables  $operprof_{it}^{(q)}$  ( $q = 1, \dots, 5$ ) for  $operprof$  in vector  $\mathbf{z}_{it}$  (displayed on the horizontal axis). The quintile portfolios (Panel A) and dummy variables (Panel B) are formed based on NYSE breakpoints. The sample period is July 1963 to December 2019. Statistical inference on the portfolio sorts (Panel A) is based on Newey and West (1987) standard errors with a lag-length of three. The GPS-model in Panel B is estimated with Driscoll and Kraay (1998) standard errors with a lag-length of three. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels (two-tailed).

**Table 3: Comparison of high vs. low quintile portfolios sorted on operating profitability**

| Panel A: Standard portfolio sorts |                      |                       | Panel B: GPS-model (weighted pooled OLS) |                                      |                         |                       |
|-----------------------------------|----------------------|-----------------------|------------------------------------------|--------------------------------------|-------------------------|-----------------------|
|                                   | Q1 (low)             | Q5 - Q1               |                                          | Vector $\mathbf{z}_{it} \rightarrow$ | 1 $operprof_{it}^{(5)}$ |                       |
| $a$                               | -0.318***<br>(-3.83) | 0.524***<br>(4.64)    | Vector $\mathbf{x}_t$<br>$\leftarrow$    | 1 (Intercept)                        | -0.318***<br>(-3.83)    | 0.524***<br>(4.64)    |
| $b_{RMRF}$                        | 1.118***<br>(42.55)  | -0.176***<br>(-4.85)  |                                          | $RMRF_t$                             | 1.118***<br>(42.55)     | -0.176***<br>(-4.85)  |
| $b_{SMB}$                         | 0.382***<br>(11.35)  | -0.510***<br>(-10.00) |                                          | $SMB_t$                              | 0.382***<br>(11.35)     | -0.510***<br>(-10.00) |
| $b_{HML}$                         | 0.0172<br>(0.37)     | -0.164**<br>(-2.15)   |                                          | $HML_t$                              | 0.0172<br>(0.37)        | -0.164**<br>(-2.15)   |
| $R\text{-squared}$                | 0.886                | 0.326                 |                                          | $R\text{-squared}$                   | 0.212                   |                       |
| $N\text{ Obs.}$                   | 678                  | 678                   |                                          | $N\text{ Obs.}$                      | 1,035,812               |                       |
|                                   |                      |                       |                                          | $N\text{ Stocks}$                    | 14,500                  |                       |

This table reports value-weighted 3-factor model alphas along with RMRF (market excess return), SMB (small minus big), and HML (high minus low) factor loadings for portfolios sorted on operating profitability (variable  $operprof$  of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). Panel A reports results from standard portfolio sorts, where the bottom quintile portfolio's excess return (first column) or the return difference between the top quintile and the bottom quintile profitability portfolio (second column) is regressed on the three Fama and French (1993) factors. Statistical inference for the standard portfolio sorts approach is based on Newey and West (1987) standard errors with a lag-length of three. Panel B reproduces the results from standard portfolio sorts by using the GPS-model in Equation (12). The GPS-model is estimated with weighted pooled OLS, where observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks in the bottom- and the top-quintile profitability portfolio, respectively. Coefficient estimates and  $t$ -statistics (in parentheses) for the GPS-model in Panel B are for the product of the market-level factor variables (plus a constant) contained in vector  $\mathbf{x}_t$  (displayed on the vertical axis) and firm-characteristics vector  $\mathbf{z}_{it}$ . Vector  $\mathbf{z}_{it}$  comprises a constant and dummy variable  $operprof_{it}^{(5)}$ , which is one for stocks with  $operprof$  in the top quintile (based on NYSE breakpoints), and zero otherwise. The GPS-model in Panel B only includes firms, which belong to the top- or bottom-quintile groups of firms sorted on operating profitability. Statistical inference on the GPS-model is based on Driscoll and Kraay (1998) standard errors with a lag-length of three. The sample period is July 1963 to December 2019. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels (two-tailed).

**Table 4: Portfolio sorts specification test**

|                                           | Coefficient estimates on<br>explanatory variables $\mathbf{z}_{it} \otimes \mathbf{x}_t$<br>(coefficient vector $\boldsymbol{\theta}$ ) |                       | Coefficient estimates on<br>firm-specific time-series averages<br>(coefficient vector $\boldsymbol{\xi}$ ) |                       |
|-------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|-----------------------|------------------------------------------------------------------------------------------------------------|-----------------------|
|                                           | 1                                                                                                                                       | $operprof_{it}^{(5)}$ | 1                                                                                                          | $operprof_{it}^{(5)}$ |
| 1 (Intercept)                             | -0.189*<br>(-1.83)                                                                                                                      | 0.264<br>(1.55)       | ---<br>---                                                                                                 | 0.109<br>(0.41)       |
| RMRF                                      | 1.117***<br>(43.05)                                                                                                                     | -0.177***<br>(-4.94)  | 0.0378<br>(0.30)                                                                                           | 0.405*<br>(1.94)      |
| SMB                                       | 0.383***<br>(11.29)                                                                                                                     | -0.509***<br>(-9.79)  | 0.0212<br>(0.13)                                                                                           | 0.167<br>(0.67)       |
| HML                                       | 0.0229<br>(0.51)                                                                                                                        | -0.165**<br>(-2.20)   | -0.192<br>(-0.90)                                                                                          | -0.0546<br>(-0.21)    |
| <i>R-squared</i>                          | 0.212                                                                                                                                   |                       |                                                                                                            |                       |
| <i>N Obs.</i>                             | 1,035,812                                                                                                                               |                       |                                                                                                            |                       |
| <u>Portfolio sorts specification test</u> |                                                                                                                                         |                       |                                                                                                            |                       |
| <i>F</i> (7, 677)                         | 3.04                                                                                                                                    |                       |                                                                                                            |                       |
| <i>p-value</i>                            | 0.004                                                                                                                                   |                       |                                                                                                            |                       |

This table reports the results from estimating regression  $r_{it} = (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + \bar{\mathbf{q}}_i \boldsymbol{\xi} + \varepsilon_{it}$  for portfolio sorts on operating profitability (variable  $operprof$  of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). Vector  $\mathbf{z}_{it}$  is specified as  $\mathbf{z}_{it} = [1 \quad operprof_{it}^{(5)}]$ , where  $operprof_{it}^{(5)}$  is a dummy variable equal to one if firm  $i$  belongs to the top-quintile portfolio sorted on operating profitability. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t \quad SMB_t \quad HML_t]$  with RMRF (market excess return), SMB (small minus big), and HML (high minus low) representing monthly Fama and French (1993) factor returns. Vector  $\bar{\mathbf{q}}_i$  comprises the (cross-sectionally centered) firm-specific time-series averages from all variables in  $(\mathbf{z}_{it} \otimes \mathbf{x}_t)$ . Estimation results for coefficient vectors  $\boldsymbol{\theta}$  and  $\boldsymbol{\xi}$  are reported in matrix form. Each coefficient estimate and  $t$ -statistic (in parentheses) is for the product of the market-level factor variable in vector  $\mathbf{x}_t$  (displayed on the vertical axis) and the portfolio dummy variable (or constant) in vector  $\mathbf{z}_{it}$  (left panel), or for the time-series average of the interaction between the factor variable in  $\mathbf{x}_t$  and the portfolio dummy variable in  $\mathbf{z}_{it}$  (right panel). The regressions are estimated with weighted pooled OLS, where observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks in the quintile portfolios sorted on operating profitability. The portfolio sorts specification test displayed in the lower left part is a Wald-test on  $H_0: \boldsymbol{\xi} = 0$ . Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels (two-tailed).

**Table 5: Continuous and multivariate firm characteristics**

|                 | Panel A: Weighted pooled OLS |                      |                    |                      | Panel B: Weighted FE estimation |                      |                    |                      |
|-----------------|------------------------------|----------------------|--------------------|----------------------|---------------------------------|----------------------|--------------------|----------------------|
|                 | (1)                          | (2)                  | (3)                | (4)                  | (5)                             | (6)                  | (7)                | (8)                  |
| Constant        | -0.038<br>(-1.19)            | 0.053***<br>(3.00)   | -0.036<br>(-1.38)  | -0.101***<br>(-3.08) | -0.001<br>(-0.02)               | 0.052***<br>(2.79)   | -0.023<br>(-0.94)  | -0.050<br>(-1.33)    |
| <i>operprof</i> | 0.174***<br>(2.96)           |                      |                    | 0.157***<br>(2.87)   | 0.088<br>(1.36)                 |                      |                    | 0.073<br>(1.18)      |
| <i>invest</i>   |                              | -0.129***<br>(-3.59) |                    | -0.132***<br>(-3.70) |                                 | -0.125***<br>(-3.53) |                    | -0.112***<br>(-3.21) |
| <i>mom12m</i>   |                              |                      | 0.414***<br>(4.38) | 0.401***<br>(4.12)   |                                 |                      | 0.335***<br>(3.50) | 0.317***<br>(3.21)   |
| (within) $R^2$  | 0.239                        | 0.238                | 0.235              | 0.246                | 0.239                           | 0.238                | 0.235              | 0.246                |
| <i>N Obs.</i>   | 1,951,536                    | 2,010,654            | 2,392,018          | 1,913,216            | 1,951,536                       | 2,010,654            | 2,392,018          | 1,913,216            |
| <i>N Stocks</i> | 15,772                       | 16,047               | 19,616             | 15,641               | 15,772                          | 16,047               | 19,616             | 15,641               |

This table reports the coefficient estimates and  $t$ -statistics (in parentheses) from GPS-models with multivariate and continuous firm characteristics in vector  $\mathbf{z}_{it}$ . The GPS-model is specified as

$$r_{it} = (\mathbf{z}_{it} \otimes \mathbf{x}_t) \boldsymbol{\theta} + v_{it} \\ = ([1 \text{ } operprof_{it} \text{ } invest_{it} \text{ } mom12m_{it}] \otimes [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]) \boldsymbol{\theta} + v_{it}$$

$operprof_{it}$  is operating profitability (variable *operprof* of the firm characteristics dataset of Gu, Kelly, and Xiu, 2018, 2019).  $invest_{it}$  is capital expenditures and inventory (variable *invest* of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021).  $mom12m_{it}$  denotes twelve-month momentum (variable *mom12m* of the firm characteristics dataset of Gu, Kelly, and Xiu, 2020, 2021). Variables  $operprof_{it}$ ,  $invest_{it}$ , and  $mom12m_{it}$  are cross-sectionally standardized, such that each month their cross-sectional mean is zero and the cross-sectional standard deviation is one. Note that  $operprof_{it}$ ,  $invest_{it}$ , and  $mom12m_{it}$  are lagged as in Gu, Kelly, and Xiu (2020, 2021). The table only displays results for coefficient estimates of the variables in vector  $\mathbf{z}_{it}$ . The Fama-French three-factor model alpha (as a measure of the risk-adjusted performance) conditional on firm  $i$ 's cross-sectionally standardized characteristics in period  $t$  is obtained as  $\hat{\alpha}_{it} = \hat{\theta}_0 + \hat{\theta}_1 \times operprof_{it} + \hat{\theta}_2 \times invest_{it} + \hat{\theta}_3 \times mom12m_{it}$ . The results for the interactions of vector  $\mathbf{z}_{it}$  firm characteristics with the market ( $RMRF_t$ ),  $SMB_t$  (small minus big), and  $HML_t$  (high minus low) factors are not shown in the table. The sample period is July 1963 to December 2019. The GPS-models in Panel A are estimated with weighted pooled OLS, where observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Panel B reports the results from estimating the GPS-models with the fixed effects (FE) estimator, where observation weights are also set equal to the beginning-of-time  $t$  value-weights of the stocks.  $t$ -statistics test for significance against a value of zero. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels (two-tailed).

**Table 6: Comparison of pooled OLS and FE estimators for Gu et al. (2020, 2021) predictors**

|       |         |                       | OLS<br>significant | FE<br>significant | Loses<br>significance | Switches<br>sign | FE different |
|-------|---------|-----------------------|--------------------|-------------------|-----------------------|------------------|--------------|
| (1)   | FF3     | Hi-Low, 5 Portfolios  | 48                 | 22                | 20 (42%)              | 6 (13%)          | 26 (54%)     |
| (2)   |         | Hi-Low, 10 Portfolios | 48                 | 21                | 22 (46%)              | 5 (10%)          | 27 (56%)     |
| (3)   |         | Standardized          | 49                 | 25                | 22 (45%)              | 2 (4%)           | 24 (49%)     |
| (4)   | FF5     | Hi-Low, 5 Portfolios  | 46                 | 20                | 18 (39%)              | 8 (17%)          | 26 (57%)     |
| (5)   |         | Hi-Low, 10 Portfolios | 42                 | 18                | 15 (36%)              | 9 (21%)          | 24 (57%)     |
| (6)   |         | Standardized          | 44                 | 26                | 14 (32%)              | 4 (9%)           | 18 (41%)     |
| (7)   | Carhart | Hi-Low, 5 Portfolios  | 39                 | 16                | 20 (51%)              | 3 (8%)           | 23 (59%)     |
| (8)   |         | Hi-Low, 10 Portfolios | 42                 | 14                | 24 (57%)              | 4 (10%)          | 28 (67%)     |
| (9)   |         | Standardized          | 34                 | 18                | 15 (44%)              | 1 (3%)           | 16 (47%)     |
| (10)  | MM      | Hi-Low, 5 Portfolios  | 44                 | 29                | 13 (30%)              | 2 (5%)           | 15 (34%)     |
| (11)  |         | Hi-Low, 10 Portfolios | 47                 | 30                | 14 (30%)              | 3 (6%)           | 17 (36%)     |
| (12)  |         | Standardized          | 49                 | 31                | 18 (37%)              | 0 (0%)           | 18 (37%)     |
| Total |         |                       | 532                | 270               | 215 (40%)             | 47 (9%)          | 262 (49%)    |

This table provides a summary of results obtained from estimating GPS models either with pooled OLS or a fixed effects estimator using one of three alternative factor models in vector  $\mathbf{x}_t$  and each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021) in vector  $\mathbf{z}_{it}$ . Models summarized in Rows “FF3” use the Fama and French (1993) three-factor model with vector  $\mathbf{x}_t$  being set to  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]$  with RMRF (market excess return), SMB (small minus big), and HML (high minus low) representing monthly Fama and French (1993) factor returns. Rows “FF5” use the Fama and French (2015) five-factor model with vector  $\mathbf{x}_t$  being set to  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t \text{ } RMW_t \text{ } CMA_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), RMW (robust minus weak), and CMA (conservative minus aggressive) representing monthly Fama and French (2015) factor returns. Rows “Carhart” use the Carhart (1997) four-factor model with vector  $\mathbf{x}_t$  set to  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t \text{ } MOM_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), and MOM (winners minus losers) representing monthly Carhart (1997) factor returns. Models summarized in Rows “MM” use a market model with vector  $\mathbf{x}_t$  being set to  $\mathbf{x}_t = [1 \text{ } RMRF_t]$  with RMRF representing the market excess return. Model estimations summarized in Rows “Hi-Low, 5 Portfolios” only include observations, which belong to either the top- or bottom-quintiles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy variable which is one for observations in the top quintile of the respective firm characteristic. Model estimations summarized in Rows “Hi-Low, 10 Portfolios” only include observations, which belong to either the top- or bottom-deciles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy variable which is one for observations in the top decile of the respective firm characteristic. Models summarized in Rows “Standardized” use firm characteristics of Gu, Kelly, and Xiu (2020, 2021) that are cross-sectionally standardized each month. In this case, vector  $\mathbf{z}_{it}$  includes a constant and the standardized firm characteristic. Column “OLS sign.” counts the number of model estimations with the coefficient on the standardized firm characteristic or top tercile, quintile, or decile dummy estimated significant at the 5% level or better. Column “FE sign.” counts the number of estimations with significant OLS coefficient estimates which remain statistically significant (and no change in sign) when the model is estimated with a fixed effects estimator. Column “Loses sign.” reports the number (and percentage) of models with significant estimate on the firm characteristic that turn insignificant when estimated with a fixed effects estimator. Column “Switch sign” reports the number (and percentage) of models with significant estimate on the firm characteristic that switch sign when estimated with a fixed effects estimator. Column “FE different” reports the number (and percentage) of models estimated with OLS that produce a significant coefficient estimate on the firm characteristic in vector  $\mathbf{z}_{it}$  that either turns insignificant or changes sign when estimated with a fixed effects estimator. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

# Internet Appendix to: Generalized Portfolio Sorts for Factor Validation

Daniel Hoechle, Markus Schmid, Heinz Zimmermann

## Internet Appendix A: Proof of Propositions 1 to 3

In this appendix, we first set the ground by rewriting both the GPS-model and the portfolio sorts approach in matrix notation. We then perform a series of basic transformations that apply to all three propositions stated in Section 2. Finally, we proof Propositions 1 to 3 mathematically. For ease of mathematical tractability and as outlined in Section 2, we thereby restrict our formal analysis to the case of a balanced panel ( $N$  firms with  $T$  regularly spaced observations), time-constant firm characteristics (i.e.,  $\mathbf{z}_{it} \equiv \mathbf{z}_i$ ), and equally weighted portfolios (i.e.,  $w_{it} = 1/N$ ). Under these simplifying assumptions, the GPS-model can reproduce the results of the portfolio sorts approach by aid of standard pooled OLS where all observations are equally weighted.

### A.1 Matrix notation and proof of Proposition 1

#### A.1.1 Coefficient estimates and standard errors for the GPS-model

Applying matrix notation, we can write the GPS-model in (1) as:<sup>1</sup>

$$\begin{bmatrix} r_{11} \\ r_{12} \\ \vdots \\ r_{1T} \\ r_{21} \\ \vdots \\ r_{NT} \end{bmatrix} = \left( \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_N \end{bmatrix} \otimes \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_T \end{bmatrix} \right) \boldsymbol{\theta} + \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1T} \\ v_{21} \\ \vdots \\ v_{NT} \end{bmatrix} \quad (\text{A1})$$

or, more briefly:

---

<sup>1</sup> Throughout Appendix A, we assume  $c_i = 0$  (for all  $i$ ) and omit the firm-specific effects  $c_i$  from the analysis.

$$vec(\mathbf{R}) = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = (\mathbf{Z} \otimes \mathbf{X}) \boldsymbol{\theta} + \mathbf{v} \quad (\text{A2})$$

where  $vec(\mathbf{R})$  represents a  $(NT \times 1)$ -vector of the firms' period  $t$  (excess) returns,  $\mathbf{z}_i$  is a  $(1 \times M)$ -vector of firm characteristics (which are assumed to remain constant over time),  $\mathbf{x}_t = [1 \ x_{1t} \ \dots \ x_{Kt}]$  refers to a  $(1 \times (K + 1))$ -vector of market-level variables (which apart from the constant change over time but do not vary in the cross-section), and  $(\mathbf{Z} \otimes \mathbf{X})$  denotes the Kronecker product of  $(N \times M)$ -dimensional matrix  $\mathbf{Z} = [\mathbf{z}'_1 \ \dots \ \mathbf{z}'_N]'$  with  $(T \times (K + 1))$ -dimensional matrix  $\mathbf{X} = [\mathbf{x}'_1 \ \dots \ \mathbf{x}'_T]'$ . Estimating regression model (A2) with pooled OLS, and applying the calculus rules for the Kronecker product, yields the following coefficient estimates for  $\boldsymbol{\theta}$ :

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= ((\mathbf{Z} \otimes \mathbf{X})'(\mathbf{Z} \otimes \mathbf{X}))^{-1}(\mathbf{Z} \otimes \mathbf{X})'vec(\mathbf{R}) \\ &= (\mathbf{Z}'\mathbf{Z} \otimes \mathbf{X}'\mathbf{X})^{-1}(\mathbf{Z}' \otimes \mathbf{X}')vec(\mathbf{R}) \\ &= ((\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' \otimes (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')vec(\mathbf{R}) \end{aligned} \quad (\text{A3})$$

Next, we use the following Lemma from linear algebra (e.g., see Sydsaeter, Strom, and Berck, 2000, p. 146):

**Lemma 1.** For any three matrices  $\mathbf{A} \in \mathbb{R}^{r,r}$ ,  $\mathbf{B} \in \mathbb{R}^{r,s}$ , and  $\mathbf{C} \in \mathbb{R}^{s,s}$  it holds true that  $vec(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})vec(\mathbf{B})$ .

Applying Lemma 1 to expression (A3) above yields

$$\tilde{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \quad (\text{A4})$$

Here,  $\tilde{\boldsymbol{\theta}}$  refers to the  $((K + 1) \times M)$ -dimensional matrix of coefficient estimates  $\hat{\theta}_{k,m}$  for the interaction of firm characteristic  $m$  (with  $m = 1, \dots, M$ ) from vector  $\mathbf{z}_i$  and market-level factor variable  $k$  (with  $k = 0, \dots, K$ ) in vector  $\mathbf{x}_t$ .

We now turn to the Driscoll and Kraay (1998) covariance matrix estimator for the pooled OLS regression model in (A1). For  $H$  lags, it has the following structure:

$$\begin{aligned} \tilde{\mathbf{V}}(\hat{\boldsymbol{\theta}}) &= ((\mathbf{Z} \otimes \mathbf{X})'(\mathbf{Z} \otimes \mathbf{X}))^{-1} \tilde{\mathbf{S}}_T ((\mathbf{Z} \otimes \mathbf{X})'(\mathbf{Z} \otimes \mathbf{X}))^{-1} \\ &= ((\mathbf{Z}'\mathbf{Z})^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}) \tilde{\mathbf{S}}_T ((\mathbf{Z}'\mathbf{Z})^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}) \end{aligned} \quad (\text{A5})$$

with  $\tilde{\mathbf{S}}_T = \hat{\mathbf{\Omega}}_0 + \sum_{j=1}^H \omega_{j,H} (\hat{\mathbf{\Omega}}_j + \hat{\mathbf{\Omega}}_j') , \quad \hat{\mathbf{\Omega}}_j =$   
 $\sum_{\tau=j+1}^T \mathbf{h}_\tau(\hat{\boldsymbol{\theta}}) \mathbf{h}_{\tau-j}'(\hat{\boldsymbol{\theta}}) ,$   
and  $\mathbf{h}_\tau(\hat{\boldsymbol{\theta}}) = (\mathbf{Z} \otimes \mathbf{x}_\tau)' \hat{\mathbf{v}}$

The modified Bartlett weights  $\omega_{j,H} = 1 - j/(H + 1)$  ensure positive semi-definiteness of  $\tilde{\mathbf{S}}_T$  and smooth the sample autocovariance function such that higher order lags receive less weight.

#### A.1.2 Coefficient estimates and standard errors for the portfolio sorts approach

The portfolio sorts approach comprises two steps. The first step involves computing the month  $t$  average return for portfolio  $p$  as outlined in Equation (5). In our case of a balanced panel, time-constant firm characteristics, and equally weighted portfolios (i.e.,  $w_{it} = N_p^{-1} = \left( \sum_{i=1}^N z_i^{(p)} \right)^{-1}$ ), we can rewrite Equation (5) as follows:

$$r_{pt} = \frac{1}{N_p} \sum_{i=1}^N z_i^{(p)} r_{it} = \mathbf{r}_t' \mathbf{d}_p (\mathbf{d}_p' \mathbf{d}_p)^{-1} \quad (\text{A6})$$

with  $\mathbf{r}_t' = [r_{1t} \ r_{2t} \ \dots \ r_{Nt}]$  and  $\mathbf{d}_p' = [z_1^{(p)} \ z_2^{(p)} \ \dots \ z_N^{(p)}]$ .

Here,  $z_i^{(p)}$  is a dummy variable with value one if firm  $i$  belongs to portfolio  $p$ , and zero otherwise. In the second step of the procedure,  $r_{pt}$  from (A6) is regressed on a constant and the  $K$  factor variables as outlined in Equation (6). This yields OLS coefficient estimates as follows:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_p &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \begin{bmatrix} r_{p1} \\ \vdots \\ r_{pT} \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \begin{bmatrix} \mathbf{r}_1' \mathbf{d}_p (\mathbf{d}_p' \mathbf{d}_p)^{-1} \\ \vdots \\ \mathbf{r}_T' \mathbf{d}_p (\mathbf{d}_p' \mathbf{d}_p)^{-1} \end{bmatrix} \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \begin{bmatrix} \mathbf{r}_1' \\ \vdots \\ \mathbf{r}_T' \end{bmatrix} \mathbf{d}_p (\mathbf{d}_p' \mathbf{d}_p)^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' [\boldsymbol{\gamma}_1 \ \boldsymbol{\gamma}_2 \ \dots \ \boldsymbol{\gamma}_N] \mathbf{d}_p (\mathbf{d}_p' \mathbf{d}_p)^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{R} \mathbf{d}_p (\mathbf{d}_p' \mathbf{d}_p)^{-1} \end{aligned} \quad (\text{A7})$$

The formula for computing the Newey and West (1987) covariance matrix with lag length  $H$  for the regression model in Equation (6) has the following structure:

$$\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}}_p) = (\mathbf{X}'\mathbf{X})^{-1}\widehat{\mathbf{S}}_T(\mathbf{X}'\mathbf{X})^{-1}$$

with

$$\widehat{\mathbf{S}}_T = \sum_{t=1}^T \widehat{\varepsilon}_{pt}^2 \mathbf{x}_t' \mathbf{x}_t + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( \widehat{\varepsilon}_{p\tau} \widehat{\varepsilon}_{p\tau-j} (\mathbf{x}_\tau' \mathbf{x}_{\tau-j} + \mathbf{x}_{\tau-j}' \mathbf{x}_\tau) \right) \quad (\text{A8})$$

Restricting the sample to two portfolios (or groups of firms), arbitrarily denoted as  $p = \text{"high"}$  and  $p = \text{"low"}$ , we can use matrix notation to rewrite Equation (10) computing the month  $t$  return difference as follows:

$$\Delta r_{p,t} = r_{high,t} - r_{low,t} = \mathbf{r}_t' \mathbf{Z} (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{e}_2 = \begin{bmatrix} r_{low,t} & \Delta r_{p,t} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (\text{A9})$$

where matrix  $\mathbf{Z}$  is specified as  $\mathbf{Z} = [\boldsymbol{\iota} \quad \mathbf{d}_{high}]$  and  $\boldsymbol{\iota}$  is a  $(N \times 1)$ -dimensional vector of ones. When regressing  $\Delta r_{p,t}$  from (A9) on a constant and the  $K$  factor variables according to Equation (11), one obtains the following OLS coefficient estimates:

$$\widehat{\boldsymbol{\beta}}_\Delta = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \begin{bmatrix} \Delta r_{p,1} \\ \vdots \\ \Delta r_{p,T} \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{R} \mathbf{Z} (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{e}_2 \quad (\text{A10})$$

The Newey and West (1987) covariance matrix estimator for the coefficient estimates in (A10) has the same structure as the one displayed in Equation (A8), with  $\widehat{\varepsilon}_{\Delta t}$  replacing  $\widehat{\varepsilon}_t$  in the formula.

### A.1.3 Proof of Proposition 1

Proposition 1 states that the GPS-model can reproduce the results of the portfolio sorts approach for the case of a single portfolio if vector  $\mathbf{z}_i \equiv \mathbf{z}_{it}$  is specified as  $\mathbf{z}_i = [1]$ . In this case, matrix  $\mathbf{Z}$  in Equation (A2) is given as  $\mathbf{Z} = \boldsymbol{\iota}$ . As a result, the coefficient estimates of the GPS-model in this case are

$$\widetilde{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{R} \boldsymbol{\iota} (\boldsymbol{\iota}' \boldsymbol{\iota})^{-1} \quad (\text{A11})$$

When there is only a single subject group, then  $z_i^{(p)}$  for all firms  $i$  is equal to 1, i.e.,  $\mathbf{d}_p = \iota$ . Consequently, the coefficient estimates for the portfolio sorts approach in Equation (A7) are equal to

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}\iota(\iota'\iota)^{-1} \quad (\text{A12})$$

As stated in Part A of Proposition 1, we thus have  $\tilde{\boldsymbol{\theta}} \equiv \hat{\boldsymbol{\beta}}$ . This completes the first part of the proof. ■

We next turn to the standard errors for the coefficient estimates. With  $r_{pt} = N^{-1} \sum_{i=1}^N r_{it}$  and  $\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\beta}}$ , the (estimated) residual  $\hat{\varepsilon}_t$  in Equation (3) is equal to

$$\hat{\varepsilon}_t = N^{-1} \sum_{i=1}^N \hat{v}_{it} \equiv N^{-1} V_t \quad (\text{A13})$$

where  $\hat{v}_{it}$  is the (estimated) residual from pooled OLS regression (4). Replacing  $\hat{\varepsilon}_t$  in (A8) by the corresponding term from (A13) yields

$$N^2 \hat{\mathbf{S}}_T = \sum_{t=1}^T V_t^2 \mathbf{x}_t' \mathbf{x}_t + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T (V_\tau V_{\tau-j} (\mathbf{x}_\tau' \mathbf{x}_{\tau-j} + \mathbf{x}_{\tau-j}' \mathbf{x}_\tau)) \quad (\text{A14})$$

In case of the GPS-model, we plug in  $\mathbf{Z} = \iota$  in Equation (A5). This gives

$$\begin{aligned} \tilde{\mathbf{V}}(\hat{\boldsymbol{\theta}}) &= ((\iota'\iota)^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}) \tilde{\mathbf{S}}_T ((\iota'\iota)^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}) \\ &= (\mathbf{X}'\mathbf{X})^{-1} \frac{\tilde{\mathbf{S}}_T}{N^2} (\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \quad (\text{A15})$$

Comparing  $\tilde{\mathbf{V}}(\hat{\boldsymbol{\theta}})$  in (A15) with  $\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}})$  from (A8) in case of a single subject group, we hence have to show that  $N^{-2} \tilde{\mathbf{S}}_T = \hat{\mathbf{S}}_T$ . Rewriting  $\mathbf{h}_\tau(\hat{\boldsymbol{\theta}})$  in (A5), we obtain:

$$\mathbf{h}_\tau(\hat{\boldsymbol{\theta}}) = (\iota \otimes \mathbf{x}_\tau)' \hat{\mathbf{v}} = [\mathbf{x}_\tau' \quad \dots \quad \mathbf{x}_\tau'] \hat{\mathbf{v}} = \begin{bmatrix} \sum_{i=1}^N \hat{v}_{i\tau} \\ x_{1\tau} \sum_{i=1}^N \hat{v}_{i\tau} \\ \vdots \\ x_{K\tau} \sum_{i=1}^N \hat{v}_{i\tau} \end{bmatrix} = \mathbf{x}_\tau' V_\tau \quad (\text{A16})$$

From (A16) it follows for  $\hat{\boldsymbol{\Omega}}_j$  in (A5) that  $\hat{\boldsymbol{\Omega}}_j = \sum_{\tau=j+1}^T \mathbf{h}_\tau(\hat{\boldsymbol{\theta}}) \mathbf{h}_{\tau-j}'(\hat{\boldsymbol{\theta}}) = \sum_{\tau=j+1}^T V_\tau V_{\tau-j} \mathbf{x}_\tau' \mathbf{x}_{\tau-j}$ , and consequently

$$\tilde{\mathbf{S}}_T = \sum_{t=1}^T V_t^2 \mathbf{x}'_t \mathbf{x}_t + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( V_\tau V_{\tau-j} (\mathbf{x}'_\tau \mathbf{x}_{\tau-j} + \mathbf{x}'_{\tau-j} \mathbf{x}_\tau) \right) \equiv N^2 \hat{\mathbf{S}}_T \quad (\text{A17})$$

This completes the proof. ■

### A.2 Proof of Proposition 2

Proposition 2 states that the GPS-model can reproduce the results of the portfolio sorts approach for multiple sorted portfolios by estimating a single pooled OLS regression on the individual firm level. To this end, we specify vector  $\mathbf{z}_i$  as  $\mathbf{z}_i = [z_i^{(1)} \ z_i^{(2)} \ \dots \ z_i^{(P)}]$ . Using the definition of  $\mathbf{d}_p$  in (A6),  $(N \times P)$ -matrix  $\mathbf{Z}$  in equation (A2) is given as follows:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_N \end{bmatrix} = \begin{bmatrix} z_1^{(1)} & \dots & z_1^{(P)} \\ \vdots & \ddots & \vdots \\ z_N^{(1)} & \dots & z_N^{(P)} \end{bmatrix} = [\mathbf{d}_1 \ \dots \ \mathbf{d}_P] \quad (\text{A18})$$

The coefficient estimates for the GPS-model are derived in (A4). Specifying matrix  $\mathbf{Z}$  according to (A18) thus results in a  $((K+1) \times P)$ -dimensional matrix of coefficient estimates  $\tilde{\boldsymbol{\theta}}$ . The  $p$ -th column of results matrix  $\tilde{\boldsymbol{\theta}}$  can be obtained as follows:

$$\tilde{\boldsymbol{\theta}}_p = \tilde{\boldsymbol{\theta}} \mathbf{e}_p = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{R} \mathbf{Z} (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{e}_p \quad (\text{A19})$$

where  $\mathbf{e}_p$  is a  $P$ -dimensional vector of zeroes with a one on position  $p$ . Proposition 2 claims that  $\tilde{\boldsymbol{\theta}}_p \equiv \hat{\boldsymbol{\beta}}_p$  where  $\hat{\boldsymbol{\beta}}_p$  refers to the portfolio  $p$  coefficient estimates from estimating the second-step time-series regression of the portfolio sorts approach. The respective coefficient estimates have been derived in Equation (A7). We thus have to proof that  $\tilde{\boldsymbol{\theta}}_p \equiv \hat{\boldsymbol{\beta}}_p = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{R} \mathbf{d}_p (\mathbf{d}'_p \mathbf{d}_p)^{-1}$  or, equivalently, that

$$\mathbf{Z} (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{e}_p \equiv \mathbf{d}_p (\mathbf{d}'_p \mathbf{d}_p)^{-1} \quad (\text{A20})$$

To show that (A20) indeed is an identity, we note that  $\mathbf{d}_p = \mathbf{Z} \mathbf{e}_p$  and that  $(\mathbf{Z}'\mathbf{Z})^{-1}$  is a diagonal matrix with element  $(\mathbf{Z}'\mathbf{Z})_{p,p}^{-1} = (\mathbf{d}'_p \mathbf{d}_p)^{-1}$ . With  $(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{e}_p = (\mathbf{d}'_p \mathbf{d}_p)^{-1} \mathbf{e}_p = \mathbf{e}_p (\mathbf{d}'_p \mathbf{d}_p)^{-1}$  it therefore holds true that

$$\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{e}_p = \underbrace{\mathbf{Z}\mathbf{e}_p}_{\mathbf{d}_p}(\mathbf{d}_p'\mathbf{d}_p)^{-1} = \mathbf{d}_p(\mathbf{d}_p'\mathbf{d}_p)^{-1} \quad (\text{A21})$$

This shows that  $\tilde{\boldsymbol{\theta}}_p \equiv \hat{\boldsymbol{\beta}}_p$  and, hence, completes the first part of the proof of Proposition 2. ■

Next, we turn to the standard errors for the coefficient estimates. With  $r_{pt} = N_p^{-1} \sum_{i=1}^N z_i^p r_{it}$  and  $\tilde{\boldsymbol{\theta}}_p = \hat{\boldsymbol{\beta}}_p$ , the (estimated) residual  $\hat{\varepsilon}_{pt}$  for the portfolio sorts approach in equation (6) is equal to

$$\hat{\varepsilon}_{pt} = N_p^{-1} \sum_{i=1}^N z_i^{(p)} \hat{v}_{it} \equiv N_p^{-1} V_{pt} \quad (\text{A22})$$

where  $\hat{v}_{it}$  is the (estimated) residual from pooled OLS regression (7). Replacing  $\hat{\varepsilon}_{pt}$  in (A8) by the corresponding term from (A22) yields

$$N_p^2 \hat{\mathbf{S}}_T = \sum_{t=1}^T V_{pt}^2 \mathbf{x}_t' \mathbf{x}_t + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( V_{p\tau} V_{p,\tau-j} (\mathbf{x}_\tau' \mathbf{x}_{\tau-j} + \mathbf{x}_{\tau-j}' \mathbf{x}_\tau) \right) \equiv \hat{\mathbf{S}}_T^{(p)} \quad (\text{A23})$$

As a consequence, we finally obtain the Newey and West (1987) standard errors in case of the portfolio sorts approach as follows:

$$\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_p) = N_p^{-2} (\mathbf{X}'\mathbf{X})^{-1} \hat{\mathbf{S}}_T^{(p)} (\mathbf{X}'\mathbf{X})^{-1} \quad (\text{A24})$$

We now consider the GPS-model with matrix  $\mathbf{Z}$  defined as in Equation (A18). The  $(P \times (K + 1))$ -dimensional column vector  $\mathbf{h}_\tau(\hat{\boldsymbol{\theta}})$  from (A5) in this case is equal to

$$\mathbf{h}_\tau(\hat{\boldsymbol{\theta}}) = (\mathbf{Z}' \otimes \mathbf{x}_\tau') \hat{\mathbf{v}} = \begin{bmatrix} \sum_{i=1}^N z_i^{(1)} \hat{v}_{i\tau} \\ x_{1\tau} \sum_{i=1}^N z_i^{(1)} \hat{v}_{i\tau} \\ \vdots \\ x_{K\tau} \sum_{i=1}^N z_i^{(1)} \hat{v}_{i\tau} \\ \sum_{i=1}^N z_i^{(2)} \hat{v}_{i\tau} \\ x_{1\tau} \sum_{i=1}^N z_i^{(2)} \hat{v}_{i\tau} \\ \vdots \\ x_{K\tau} \sum_{i=1}^N z_i^{(2)} \hat{v}_{i\tau} \\ \vdots \\ \sum_{i=1}^N z_i^{(P)} \hat{v}_{i\tau} \\ x_{1\tau} \sum_{i=1}^N z_i^{(P)} \hat{v}_{i\tau} \\ \vdots \\ x_{K\tau} \sum_{i=1}^N z_i^{(P)} \hat{v}_{i\tau} \end{bmatrix} = \begin{bmatrix} V_{1\tau} \\ x_{1\tau} V_{1\tau} \\ \vdots \\ x_{K\tau} V_{1\tau} \\ V_{2\tau} \\ x_{1\tau} V_{2\tau} \\ \vdots \\ x_{K\tau} V_{2\tau} \\ \vdots \\ V_{P\tau} \\ x_{1\tau} V_{P\tau} \\ \vdots \\ x_{K\tau} V_{P\tau} \end{bmatrix} = (\mathbf{I}_P \otimes \mathbf{x}_\tau') \begin{bmatrix} V_{1\tau} \\ \vdots \\ V_{P\tau} \end{bmatrix} \equiv (\mathbf{I}_P \otimes \mathbf{x}_\tau') \mathbf{V}_\tau \quad (\text{A25})$$

where  $\mathbf{I}_P$  is the  $P$ -dimensional identity matrix. With (A25) it follows for  $\widehat{\mathbf{\Omega}}_j$  in Equation (A5) that

$$\widehat{\mathbf{\Omega}}_j = \sum_{\tau=j+1}^T (\mathbf{I}_P \otimes \mathbf{x}'_{\tau}) \mathbf{V}_{\tau} \mathbf{V}'_{\tau-j} (\mathbf{I}_P \otimes \mathbf{x}_{\tau-j}) = \sum_{\tau=j+1}^T (\mathbf{V}_{\tau} \mathbf{V}'_{\tau-j}) \otimes (\mathbf{x}'_{\tau} \mathbf{x}_{\tau-j}) \quad (\text{A26})$$

As a result, matrix  $\tilde{\mathbf{S}}_T$  in (A5) can be written as follows:

$$\begin{aligned} \tilde{\mathbf{S}}_T &= \sum_{t=1}^T (\mathbf{V}_t \mathbf{V}'_t) \otimes (\mathbf{x}'_t \mathbf{x}_t) \\ &\quad + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( (\mathbf{V}_{\tau} \mathbf{V}'_{\tau-j}) \otimes (\mathbf{x}'_{\tau} \mathbf{x}_{\tau-j}) \right. \\ &\quad \left. + (\mathbf{V}_{\tau-j} \mathbf{V}'_{\tau}) \otimes (\mathbf{x}'_{\tau-j} \mathbf{x}_{\tau}) \right) \end{aligned} \quad (\text{A27})$$

We next define  $\tilde{\mathbf{S}}_T^{(p,q)}$  as follows

$$\begin{aligned} \tilde{\mathbf{S}}_T^{(p,q)} &\equiv \sum_{t=1}^T V_{pt} V_{qt} (\mathbf{x}'_t \mathbf{x}_t) \\ &\quad + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( V_{p\tau} V_{q,\tau-j} (\mathbf{x}'_{\tau} \mathbf{x}_{\tau-j}) + V_{p,\tau-j} V_{q\tau} (\mathbf{x}'_{\tau-j} \mathbf{x}_{\tau}) \right) \end{aligned} \quad (\text{A28})$$

where  $V_{pt}$  is a scalar as in (A22) above. Consequently, matrix  $\tilde{\mathbf{S}}_T$  is equal to

$$\tilde{\mathbf{S}}_T = \begin{bmatrix} \tilde{\mathbf{S}}_T^{(1,1)} & \dots & \tilde{\mathbf{S}}_T^{(1,P)} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{S}}_T^{(P,1)} & \dots & \tilde{\mathbf{S}}_T^{(P,P)} \end{bmatrix} \quad (\text{A29})$$

Moreover, with matrix  $\mathbf{Z}$  being defined according to (A18),  $\mathbf{Z}'\mathbf{Z}$  now is a  $(P \times P)$ -dimensional diagonal matrix with element  $(p, p)$  equal to  $N_p$  and all off-diagonal elements equal to zero. With  $\tilde{\mathbf{S}}_T$  structured according to Expression (A29), we can thus rewrite  $\tilde{\mathbf{V}}(\hat{\boldsymbol{\theta}})$  in (A5) as follows:

$$\begin{aligned} &\tilde{\mathbf{V}}(\hat{\boldsymbol{\theta}}) \\ &= \begin{bmatrix} N_1^{-2} (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{S}}_T^{(1,1)} (\mathbf{X}'\mathbf{X})^{-1} & \dots & N_1^{-1} N_P^{-1} (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{S}}_T^{(1,P)} (\mathbf{X}'\mathbf{X})^{-1} \\ \vdots & \ddots & \vdots \\ N_1^{-1} N_P^{-1} (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{S}}_T^{(P,1)} (\mathbf{X}'\mathbf{X})^{-1} & \dots & N_P^{-2} (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{S}}_T^{(P,P)} (\mathbf{X}'\mathbf{X})^{-1} \end{bmatrix} \end{aligned} \quad (\text{A30})$$

The second part of Proposition 2 claims that  $\text{SE}(\hat{\theta}_{p,k}) = \text{SE}(\hat{\beta}_{p,k})$  for  $k = 0, 1, \dots, K$  and  $p = 1, \dots, P$ . To proof that this holds true, it is sufficient to show that

$N_p^{-2}(\mathbf{X}'\mathbf{X})^{-1}\tilde{\mathbf{S}}_T^{(p,p)}(\mathbf{X}'\mathbf{X})^{-1}$  in Expression (A30) is identical to  $\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_p)$  in Equation (A24) for every  $p = 1, \dots, P$ . This in turn is equivalent to demonstrating that  $\hat{\mathbf{S}}_T^{(p)}$  from (A23) coincides with  $\tilde{\mathbf{S}}_T^{(p,p)}$  in (A28). By comparing the two expressions we see that this holds true, which completes the proof. ■

### A.3 Proof of Proposition 3

Proposition 3 states that the GPS-model can reproduce the results of the portfolio sorts approach for the case of performance differences between two portfolios. In order to compare the performance of firms in group “high” with that of firms belonging to group “low”, we have to specify vector  $\mathbf{z}_i$  as  $\mathbf{z}_i = [1 \ z_i^{(high)}]$  such that the  $(N \times 2)$ -dimensional matrix  $\mathbf{Z}$  comprising the characteristics of all  $N$  firms is equal to  $\mathbf{Z} = [\mathbf{1} \ \mathbf{d}_{high}]$ . This matches the definition of matrix  $\mathbf{Z}$  in Expression (A9) which is used for deriving the results of performance differences in case of the portfolio sorts approach. Estimating the GPS-model in (A2) with pooled OLS yields the coefficient estimates, structured as a  $((K + 1) \times 2)$ -dimensional matrix, in Equation (A4).

The second column of  $\tilde{\boldsymbol{\theta}}$  in (A4) contains the coefficient estimates for the interaction terms between dummy variable  $z_i^{(high)}$  and the factor variables in vector  $\mathbf{x}_t$ . In Equation (12), those coefficient estimates are named  $\hat{\theta}_{\Delta k}$ . To extract the coefficient estimates for  $\hat{\theta}_{\Delta k}$  from  $\tilde{\boldsymbol{\theta}}$ , we post-multiply Expression (A4) with  $\mathbf{e}_2$ . The resulting term coincides with the one in Equation (A10) for the portfolio sorts approach. This shows that  $\hat{\beta}_{\Delta k} \equiv \hat{\theta}_{\Delta k} \ (\forall k = 0, 1, \dots, K)$ .

The first column of  $\tilde{\boldsymbol{\theta}}$  in (A4) comprises the coefficient estimates for subject group “low”. In Equation (12), the respective coefficient estimates are labeled as  $\hat{\theta}_{low,k}$  (with  $k = 0, 1, \dots, K$ ). The coefficient estimates for  $\theta_{low,k}$  are retrieved by post-multiplying (A4) with  $\mathbf{e}_1$ . In case of the portfolio sorts approach, we obtain the coefficient estimates for the “low” portfolio ( $\hat{\beta}_{low,k}$ ) by repeating the analysis of (A9) and (A10) with  $\mathbf{e}_1$  replacing  $\mathbf{e}_2$ . The resulting expressions for the portfolio sorts approach and the GPS-

model again coincide. This demonstrates  $\hat{\beta}_{low,k} \equiv \hat{\theta}_{low,k}$  ( $\forall k = 0, 1, \dots, K$ ) and, hence, completes the first part of the proof. ■

For the second part of the proof, we note that due to  $r_{high,t} = r_{low,t} + \Delta r_{p,t}$  the following corollary holds true:

**Corollary 2.**  $\hat{\beta}_{high,k} = \hat{\beta}_{low,k} + \hat{\beta}_{\Delta k} = \hat{\theta}_{low,k} + \hat{\theta}_{\Delta k}$  for all  $k = 0, 1, \dots, K$ . (A31)

Based on Corollary 2 and Proposition 2, and because of  $r_{pt} = N_p^{-1} \sum_{i=1}^N z_i^{(p)} r_{it}$  (for  $p$  = “low”, “high”), residual  $\hat{\varepsilon}_{\Delta t}$  in the second-step time-series regression (11) of the portfolio sorts approach is equal to

$$\hat{\varepsilon}_{\Delta t} = N_{high}^{-1} \sum_{i=1}^N z_i^{(high)} \hat{v}_{it} - N_{low}^{-1} \sum_{i=1}^N z_i^{(low)} \hat{v}_{it} \equiv N_{high}^{-1} V_{high,t} - N_{low}^{-1} V_{low,t} \quad (A32)$$

where  $\hat{v}_{it}$  is the (estimated) residual from pooled OLS regression (7). Replacing  $\hat{\varepsilon}_{\Delta t}$  in the Newey and West (1987) covariance matrix estimator (A8) for the coefficient estimates in (A10) by the respective expression in (A32) yields

$$\begin{aligned} \hat{\mathbf{S}}_T = & \sum_{t=1}^T \left( \frac{V_{high,t}}{N_{high}} - \frac{V_{low,t}}{N_{low}} \right)^2 \mathbf{x}'_t \mathbf{x}_t \\ & + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( \left( \frac{V_{high,\tau}}{N_{high}} - \frac{V_{low,\tau}}{N_{low}} \right) \left( \frac{V_{high,\tau-j}}{N_{high}} - \frac{V_{low,\tau-j}}{N_{low}} \right) (\mathbf{x}'_{\tau} \mathbf{x}_{\tau-j} \right. \\ & \left. + \mathbf{x}'_{\tau-j} \mathbf{x}_{\tau}) \right) \end{aligned} \quad (A33)$$

We now turn to the GPS-model with matrix  $\mathbf{Z}$  being defined as  $\mathbf{Z} = [\mathbf{I} \quad \mathbf{d}_{high}]$ . The  $(2 \times (K + 1))$ -dimensional column vector  $\mathbf{h}_{\tau}(\hat{\boldsymbol{\theta}})$  from (A5) in this case is equal to

$$\mathbf{h}_\tau(\hat{\boldsymbol{\theta}}) = (\mathbf{Z}' \otimes \mathbf{x}'_\tau) \hat{\mathbf{v}} = \begin{bmatrix} \sum_{i=1}^N \hat{v}_{i\tau} \\ x_{1\tau} \sum_{i=1}^N \hat{v}_{i\tau} \\ \vdots \\ x_{K\tau} \sum_{i=1}^N \hat{v}_{i\tau} \\ \sum_{i=1}^N z_i^{(high)} \hat{v}_{i\tau} \\ x_{1\tau} \sum_{i=1}^N z_i^{(high)} \hat{v}_{i\tau} \\ \vdots \\ x_{K\tau} \sum_{i=1}^N z_i^{(high)} \hat{v}_{i\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_\tau & \mathbf{x}'_\tau \\ \mathbf{x}'_\tau & \mathbf{0}' \end{bmatrix} \begin{bmatrix} V_{high,\tau} \\ V_{low,\tau} \end{bmatrix} \equiv \begin{bmatrix} \mathbf{x}'_\tau & \mathbf{x}'_\tau \\ \mathbf{x}'_\tau & \mathbf{0}' \end{bmatrix} \mathbf{V}_\tau \quad (\text{A34})$$

From (A34) and  $V_{high,\tau} + V_{low,\tau} \equiv V_\tau$  it follows for  $\hat{\boldsymbol{\Omega}}_j$  in Equation (A5) that

$$\hat{\boldsymbol{\Omega}}_j = \sum_{\tau=j+1}^T \begin{bmatrix} V_\tau V_{\tau-j} \mathbf{x}'_\tau \mathbf{x}_{\tau-j} & V_\tau V_{high,\tau-j} \mathbf{x}'_\tau \mathbf{x}_{\tau-j} \\ V_{high,\tau} V_{\tau-j} \mathbf{x}'_\tau \mathbf{x}_{\tau-j} & V_{high,\tau} V_{high,\tau-j} \mathbf{x}'_\tau \mathbf{x}_{\tau-j} \end{bmatrix} \quad (\text{A35})$$

As a result, matrix  $\tilde{\mathbf{S}}_T$  in (A5) can be written in block form as follows:

$$\tilde{\mathbf{S}}_T = \begin{bmatrix} \tilde{\mathbf{S}}_T^{(1,1)} & \tilde{\mathbf{S}}_T^{(1,2)} \\ \tilde{\mathbf{S}}_T^{(2,1)} & \tilde{\mathbf{S}}_T^{(2,2)} \end{bmatrix} \quad (\text{A36})$$

where

$$\begin{aligned} \tilde{\mathbf{S}}_T^{(1,1)} &= \sum_{t=1}^T V_t^2 (\mathbf{x}'_t \mathbf{x}_t) + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T (V_\tau V_{\tau-j} (\mathbf{x}'_\tau \mathbf{x}_{\tau-j} + \mathbf{x}'_{\tau-j} \mathbf{x}_\tau)) \\ \tilde{\mathbf{S}}_T^{(1,2)} &= \sum_{t=1}^T V_t V_{high,t} (\mathbf{x}'_t \mathbf{x}_t) + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T (V_\tau V_{high,\tau-j} \mathbf{x}'_\tau \mathbf{x}_{\tau-j} + \\ & V_{high,\tau} V_{\tau-j} \mathbf{x}'_{\tau-j} \mathbf{x}_\tau) \\ \tilde{\mathbf{S}}_T^{(2,1)} &= \sum_{t=1}^T V_t V_{high,t} (\mathbf{x}'_t \mathbf{x}_t) + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T (V_{high,\tau} V_{\tau-j} \mathbf{x}'_\tau \mathbf{x}_{\tau-j} + \\ & V_\tau V_{high,\tau-j} \mathbf{x}'_{\tau-j} \mathbf{x}_\tau) \\ \tilde{\mathbf{S}}_T^{(2,2)} &= \sum_{t=1}^T V_{high,t}^2 (\mathbf{x}'_t \mathbf{x}_t) + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T (V_{high,\tau} V_{high,\tau-j} (\mathbf{x}'_\tau \mathbf{x}_{\tau-j} + \mathbf{x}'_{\tau-j} \mathbf{x}_\tau)) \end{aligned}$$

Next, we rewrite matrix  $(\mathbf{Z}'\mathbf{Z})^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}$  in (A5) as

$$(\mathbf{Z}'\mathbf{Z})^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} N_{low}^{-1} (\mathbf{X}'\mathbf{X})^{-1} & -N_{low}^{-1} (\mathbf{X}'\mathbf{X})^{-1} \\ -N_{low}^{-1} (\mathbf{X}'\mathbf{X})^{-1} & (N_{low}^{-1} + N_{high}^{-1}) (\mathbf{X}'\mathbf{X})^{-1} \end{bmatrix} \quad (\text{A37})$$

and insert (A37) into the Driscoll and Kraay (1998) covariance matrix estimator of (A5) to obtain

$$\tilde{\mathbf{V}}(\hat{\boldsymbol{\theta}}) = ((\mathbf{Z}'\mathbf{Z})^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}) \tilde{\mathbf{S}}_T ((\mathbf{Z}'\mathbf{Z})^{-1} \otimes (\mathbf{X}'\mathbf{X})^{-1}) = \begin{bmatrix} \tilde{\mathbf{V}}^{(1,1)} & \tilde{\mathbf{V}}^{(1,2)} \\ \tilde{\mathbf{V}}^{(2,1)} & \tilde{\mathbf{V}}^{(2,2)} \end{bmatrix} \quad (\text{A38})$$

where

$$\tilde{\mathbf{V}}^{(1,1)} = N_{low}^{-2} (\mathbf{X}'\mathbf{X})^{-1} \left( \tilde{\mathbf{S}}_T^{(1,1)} - \tilde{\mathbf{S}}_T^{(1,2)} - \tilde{\mathbf{S}}_T^{(2,1)} + \tilde{\mathbf{S}}_T^{(2,2)} \right) (\mathbf{X}'\mathbf{X})^{-1}$$

$$\tilde{\mathbf{V}}^{(1,2)} = N_{low}^{-1} N_{high}^{-1} (\mathbf{X}'\mathbf{X})^{-1} \left( \tilde{\mathbf{S}}_T^{(1,2)} - \tilde{\mathbf{S}}_T^{(2,2)} \right) (\mathbf{X}'\mathbf{X})^{-1} - \tilde{\mathbf{V}}^{(1,1)}$$

$$\tilde{\mathbf{V}}^{(2,1)} = N_{low}^{-1} N_{high}^{-1} (\mathbf{X}'\mathbf{X})^{-1} \left( \tilde{\mathbf{S}}_T^{(2,1)} - \tilde{\mathbf{S}}_T^{(2,2)} \right) (\mathbf{X}'\mathbf{X})^{-1} - \tilde{\mathbf{V}}^{(1,1)}$$

$$\tilde{\mathbf{V}}^{(2,2)} = N_{high}^{-2} (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{S}}_T^{(2,2)} (\mathbf{X}'\mathbf{X})^{-1} - \tilde{\mathbf{V}}^{(1,1)} - \tilde{\mathbf{V}}^{(1,2)} - \tilde{\mathbf{V}}^{(2,1)}$$

According to the second part of Proposition 3, we have to show that  $\tilde{\mathbf{V}}^{(2,2)}$  in (A38) coincides with the Newey-West covariance matrix estimator in (A10) with  $\hat{\mathbf{S}}_T$  specified as in (A33). Therefore, we simplify the “sandwich” expressions in (A38) as follows

$$\begin{aligned} \tilde{\mathbf{S}}_T^{(1,1)} - \tilde{\mathbf{S}}_T^{(1,2)} - \tilde{\mathbf{S}}_T^{(2,1)} + \tilde{\mathbf{S}}_T^{(2,2)} &= \sum_{t=1}^T V_{low,t}^2 (\mathbf{x}'_t \mathbf{x}_t) + \\ &\quad \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( V_{low,\tau} V_{low,\tau-j} (\mathbf{x}'_{\tau} \mathbf{x}_{\tau-j} + \mathbf{x}'_{\tau-j} \mathbf{x}_{\tau}) \right) \\ \tilde{\mathbf{S}}_T^{(1,2)} - \tilde{\mathbf{S}}_T^{(2,2)} &= \sum_{t=1}^T V_{low,t} V_{high,t} (\mathbf{x}'_t \mathbf{x}_t) + \\ &\quad \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( (V_{low,\tau} V_{high,\tau-j}) \mathbf{x}'_{\tau} \mathbf{x}_{\tau-j} + (V_{low,\tau-j} V_{high,\tau}) \mathbf{x}'_{\tau-j} \mathbf{x}_{\tau} \right) \\ \tilde{\mathbf{S}}_T^{(2,1)} - \tilde{\mathbf{S}}_T^{(2,2)} &= \sum_{t=1}^T V_{high,t} V_{low,t} (\mathbf{x}'_t \mathbf{x}_t) + \\ &\quad \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( (V_{high,\tau} V_{low,\tau-j}) \mathbf{x}'_{\tau} \mathbf{x}_{\tau-j} + (V_{high,\tau-j} V_{low,\tau}) \mathbf{x}'_{\tau-j} \mathbf{x}_{\tau} \right) \end{aligned}$$

and insert the resulting expressions into  $\tilde{\mathbf{V}}^{(2,2)}$  from (A38). This finally yields

$$\tilde{\mathbf{V}}^{(2,2)} = (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{Q}}_T^A (\mathbf{X}'\mathbf{X})^{-1}$$

with

$$\begin{aligned} \tilde{\mathbf{Q}}_T^A &= \sum_{t=1}^T \left( \frac{V_{high,t}}{N_{high}} - \frac{V_{low,t}}{N_{low}} \right)^2 \mathbf{x}'_t \mathbf{x}_t \\ &\quad + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( \left( \frac{V_{high,\tau}}{N_{high}} - \frac{V_{low,\tau}}{N_{low}} \right) \left( \frac{V_{high,\tau-j}}{N_{high}} - \frac{V_{low,\tau-j}}{N_{low}} \right) (\mathbf{x}'_{\tau} \mathbf{x}_{\tau-j} \right. \\ &\quad \left. + \mathbf{x}'_{\tau-j} \mathbf{x}_{\tau}) \right) \end{aligned} \quad (\text{A39})$$

Since  $\tilde{\mathbf{Q}}_T^A$  in (A39) and  $\hat{\mathbf{S}}_T$  in (A33) coincide, this shows that  $\text{SE}(\hat{\theta}_{\Delta k}) = \text{SE}(\hat{\beta}_{\Delta k})$  for all  $k = 0, 1, \dots, K$ .

The last part of Proposition (3) claims that  $\tilde{\mathbf{V}}^{(1,1)}$  in (A38) coincides with the Newey-West covariance matrix estimator for the second-step time-series regression of the portfolio sorts approach applied to portfolio “low”. The respective Newey-West covariance estimator has been derived in Expression (A24) above with  $p = \text{“low”}$ . By replacing  $\tilde{\mathbf{S}}_T^{(1,1)} - \tilde{\mathbf{S}}_T^{(1,2)} - \tilde{\mathbf{S}}_T^{(2,1)} + \tilde{\mathbf{S}}_T^{(2,2)}$  with the corresponding term derived above, we finally obtain the following expression for  $\tilde{\mathbf{V}}^{(1,1)}$ :

$$\tilde{\mathbf{V}}^{(1,1)} = (\mathbf{X}'\mathbf{X})^{-1} \tilde{\mathbf{Q}}_T^{high} (\mathbf{X}'\mathbf{X})^{-1} \quad (\text{A40})$$

with

$$\tilde{\mathbf{Q}}_T^{high} = \sum_{t=1}^T N_{low}^{-2} V_{low,t}^2 \mathbf{x}_t' \mathbf{x}_t + \sum_{j=1}^H \omega_{j,H} \sum_{\tau=j+1}^T \left( N_{low}^{-2} V_{low,\tau} V_{low,\tau-j} (\mathbf{x}_\tau' \mathbf{x}_{\tau-j} + \mathbf{x}_{\tau-j}' \mathbf{x}_\tau) \right)$$

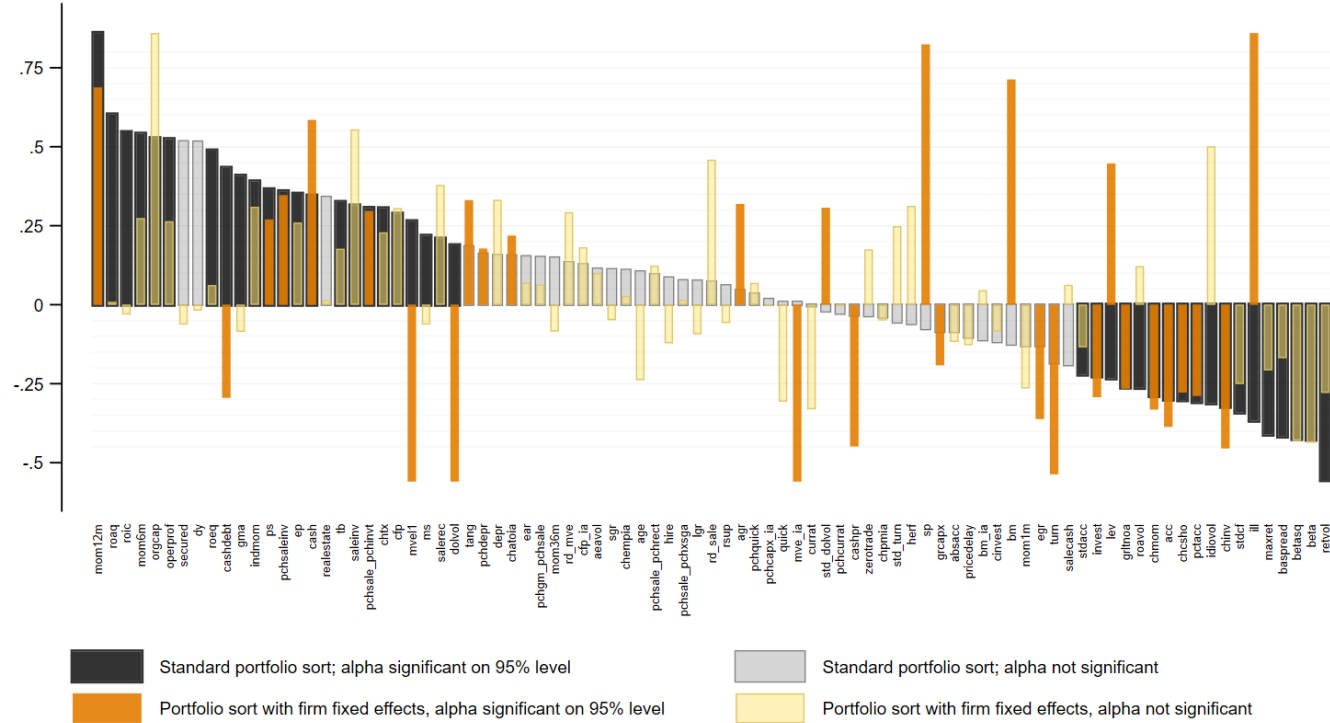
Since  $\tilde{\mathbf{Q}}_T^{high}$  in (A40) and  $\hat{\mathbf{S}}_T$  for portfolio  $p = \text{“low”}$  in (A23) are identical, this demonstrates that  $\text{SE}(\hat{\theta}_{low,k}) = \text{SE}(\hat{\beta}_{low,k})$  for all  $k = 0, 1, \dots, K$ . This completes the proof of Proposition 3. ■

## References

- Driscoll, John C., and Aart C. Kraay, 1998, Consistent covariance matrix estimation with spatially dependent panel data, *Review of Economics and Statistics* 80, 549-560.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.
- Sydsaeter, Knut, Arne Strom, and Peter Berck, 2000, *Economists' Mathematical Manual* (Springer: Berlin and Heidelberg) 3rd edn.

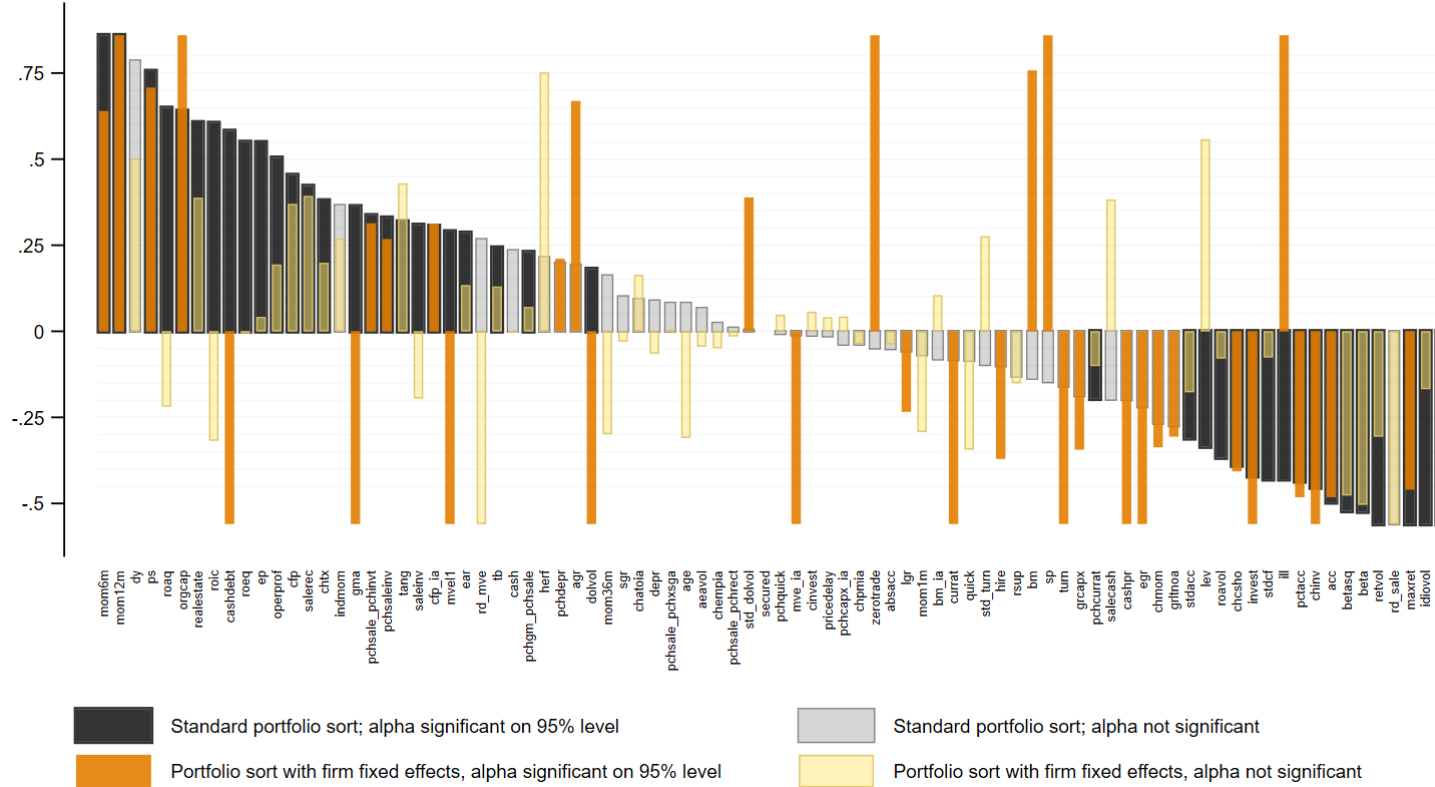
## Internet Appendix B: Figures

Figure IA.1: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (5 portfolios) – Fama and French three-factor model



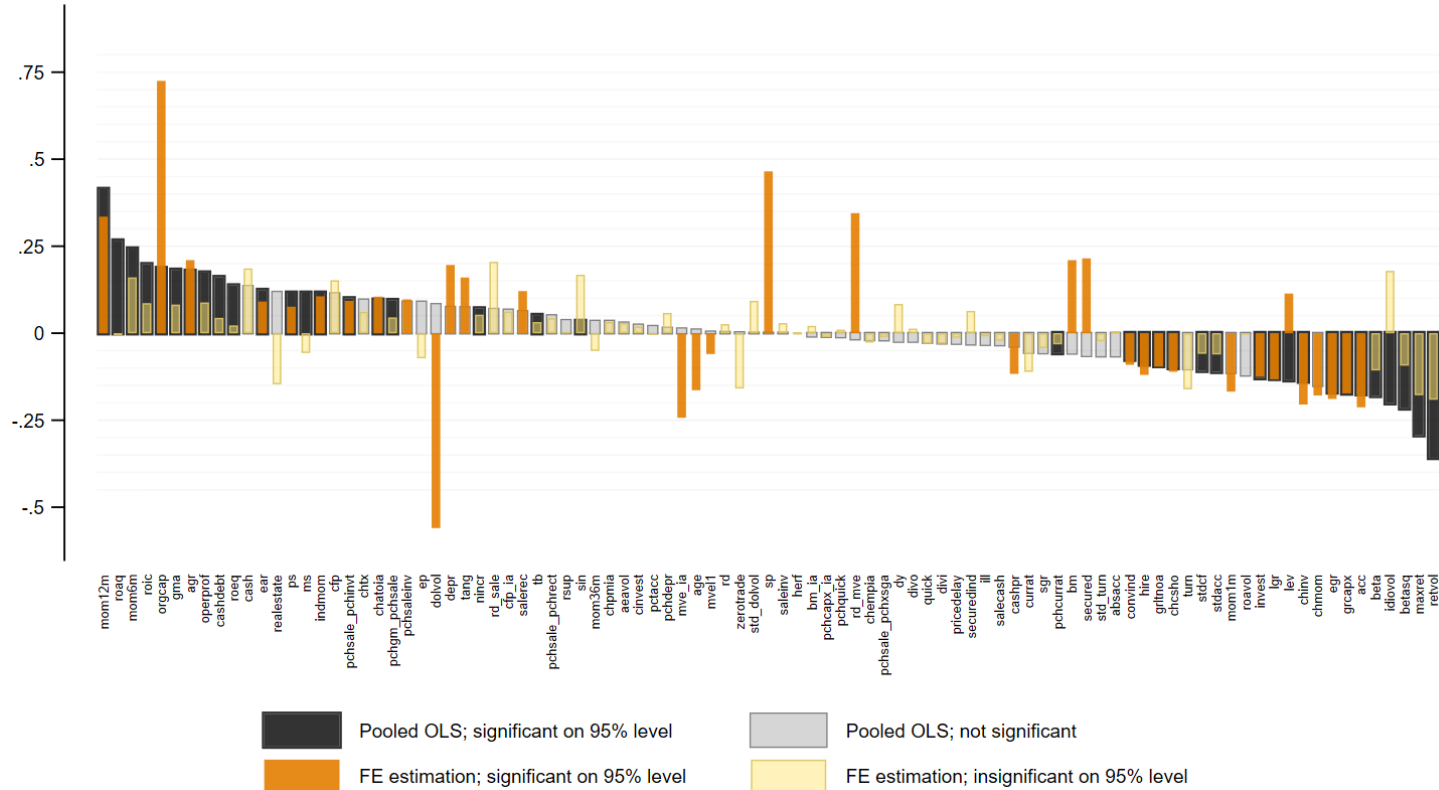
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-quintiles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top quintile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t]$  with RMRF (market excess return), SMB (small minus big), and HML (high minus low) representing monthly Fama and French (1993) factor returns. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

Figure IA.2: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (10 portfolios) – Fama and French three-factor model



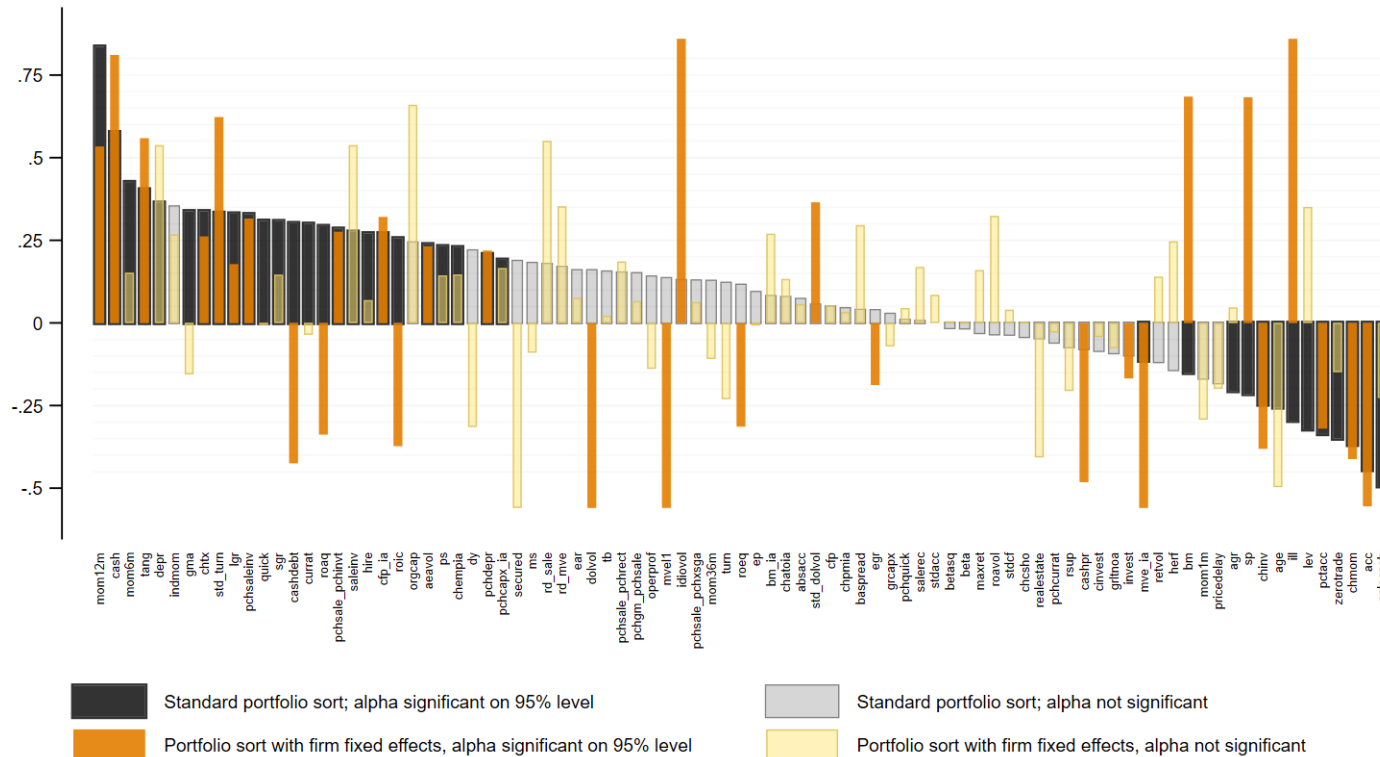
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-deciles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top decile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t \quad SMB_t \quad HML_t]$  with RMRF (market excess return), SMB (small minus big), and HML (high minus low) representing monthly Fama and French (1993) factor returns. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

**Figure IA.3: GPS models with pooled WLS, weighted FE, all variables standardized to zero mean, unit standard deviation – Fama and French three-factor model**



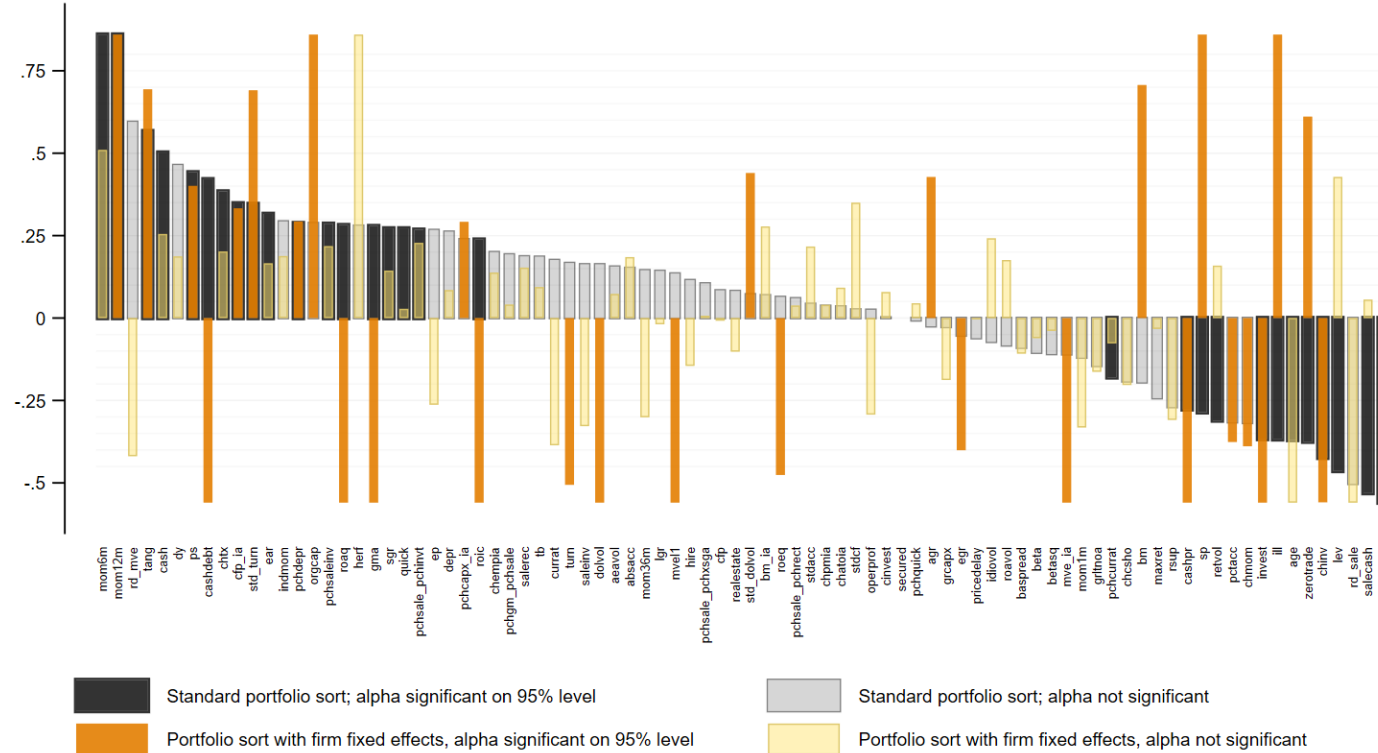
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t \quad SMB_t \quad HML_t]$  with RMRF (market excess return), SMB (small minus big), and HML (high minus low) representing monthly Fama and French (1993) factor returns. All firm characteristics of Gu, Kelly, and Xiu (2020, 2021) are cross-sectionally standardized with vector  $\mathbf{z}_{it}$  including a constant and the standardized firm characteristic. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

**Figure IA.4: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (5 portfolios) – Fama and French five-factor model**



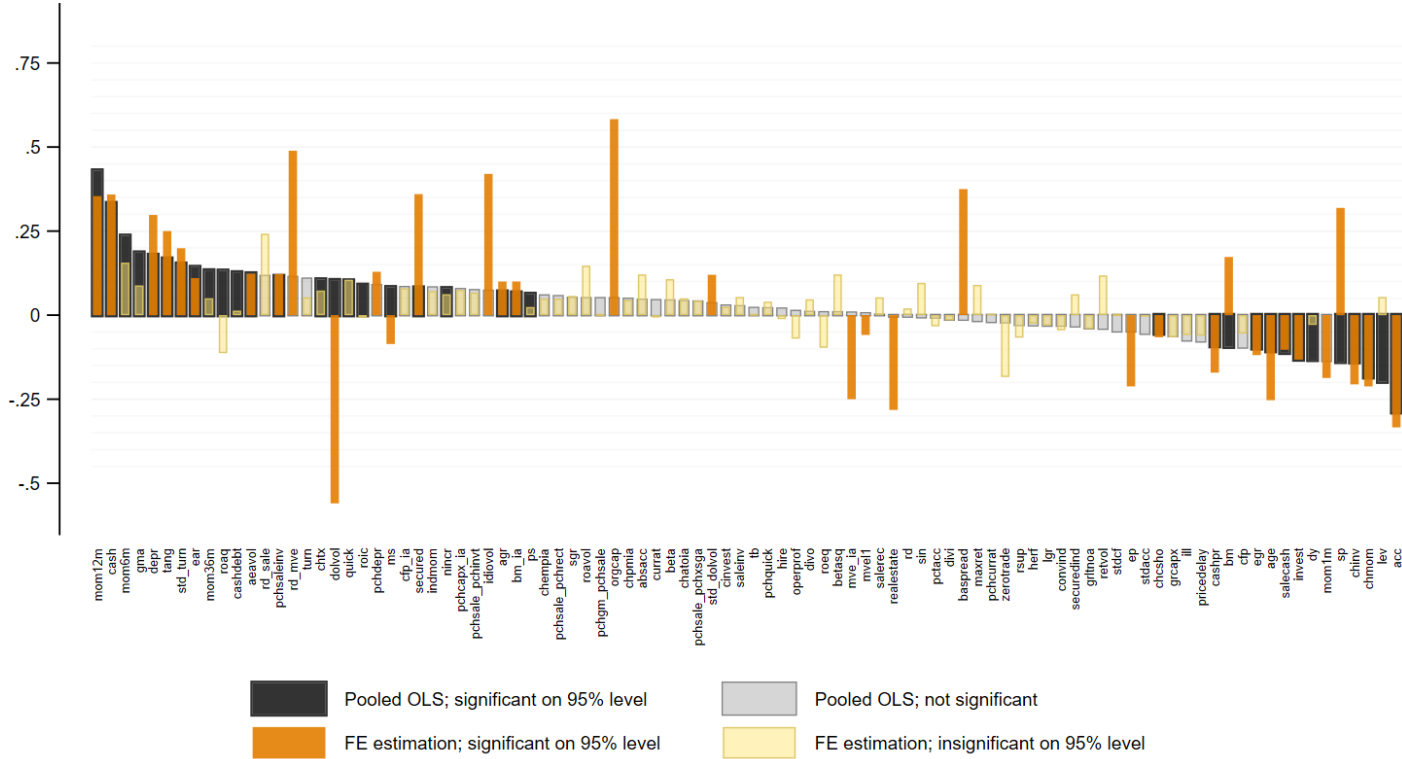
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-quintiles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top quintile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad \text{RMRF}_t \quad \text{SMB}_t \quad \text{HML}_t \quad \text{RMW}_t \quad \text{CMA}_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), RMW (robust minus weak), and CMA (conservative minus aggressive) representing monthly Fama and French (2015) factor returns. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

Figure IA.5: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (10 portfolios) – Fama and French five-factor model



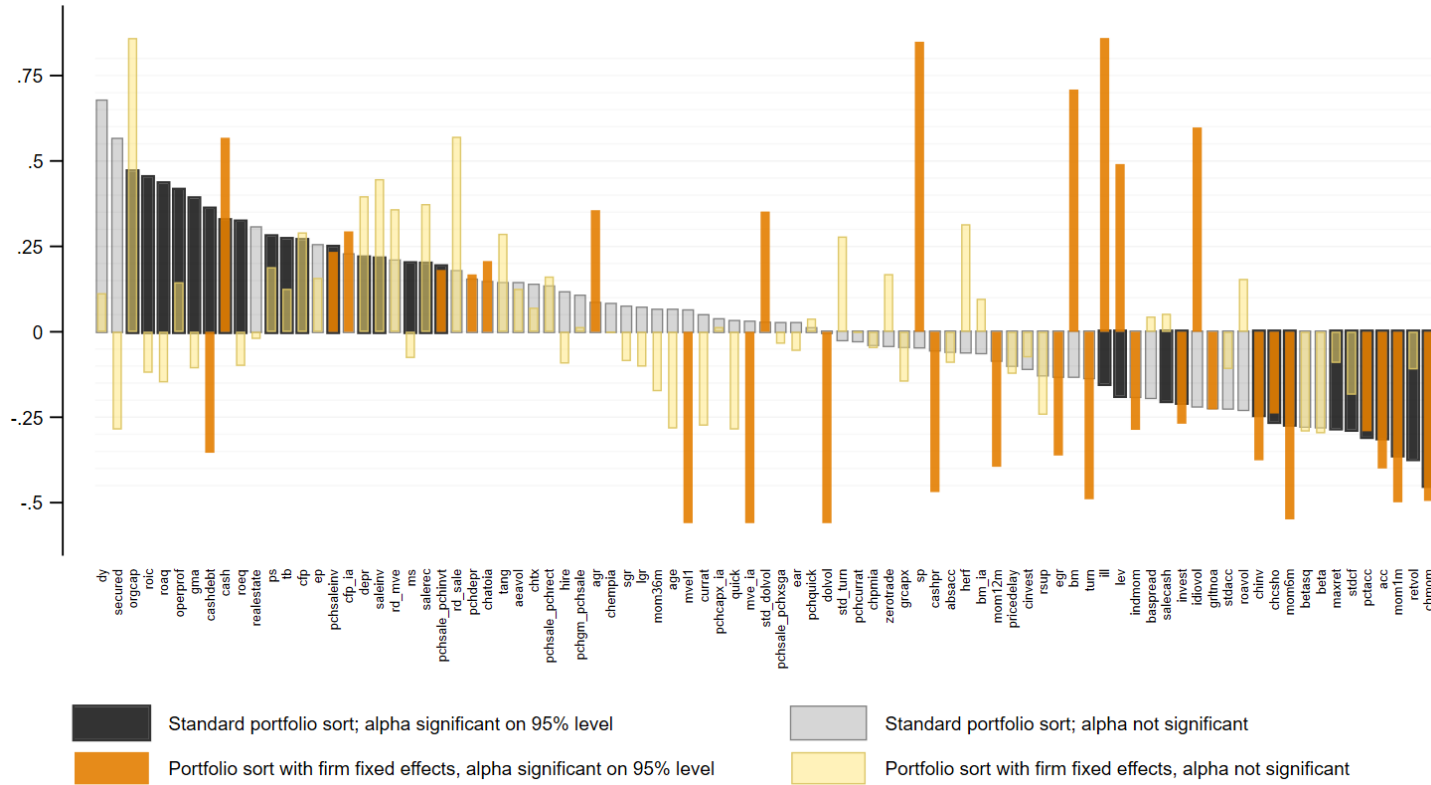
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-deciles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top decile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t \quad SMB_t \quad HML_t \quad RMW_t \quad CMA_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), RMW (robust minus weak), and CMA (conservative minus aggressive) representing monthly Fama and French (2015) factor returns. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

**Figure IA.6: GPS models with pooled WLS, weighted FE, all variables standardized to zero mean, unit standard deviation – Fama and French five-factor model**



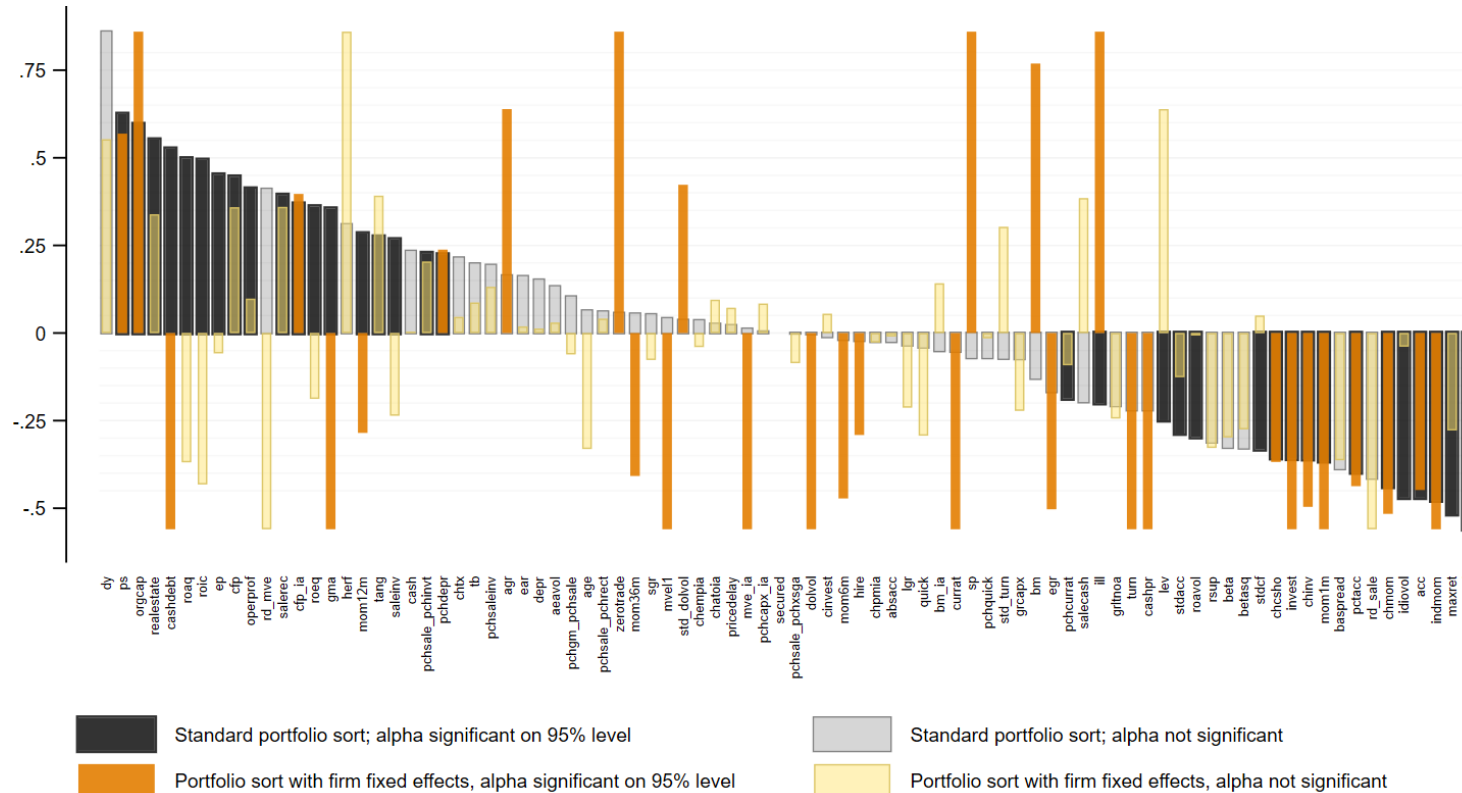
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \text{ } RMRF_t \text{ } SMB_t \text{ } HML_t \text{ } RMW_t \text{ } CMA_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), RMW (robust minus weak), and CMA (conservative minus aggressive) representing monthly Fama and French (2015) factor returns. All firm characteristics of Gu, Kelly, and Xiu (2020, 2021) are cross-sectionally standardized with vector  $\mathbf{z}_{it}$  including a constant and the standardized firm characteristic. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

Figure IA.7: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (5 portfolios) – Carhart (1997) four-factor model



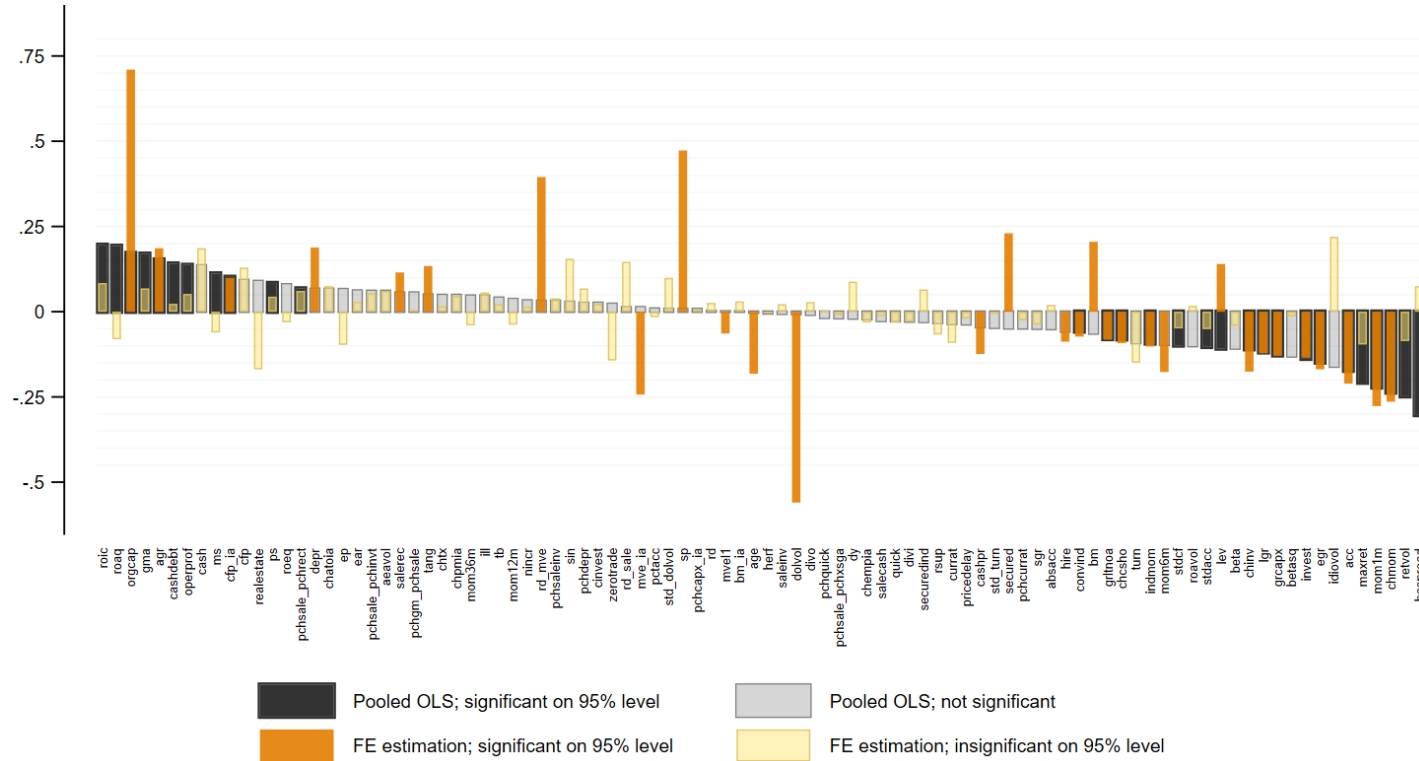
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-quintiles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top quintile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t \quad SMB_t \quad HML_t \quad MOM_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), and MOM (winners minus losers) representing monthly Carhart (1997) factor returns. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

**Figure IA.8: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (10 portfolios) – Carhart (1997) four-factor model**



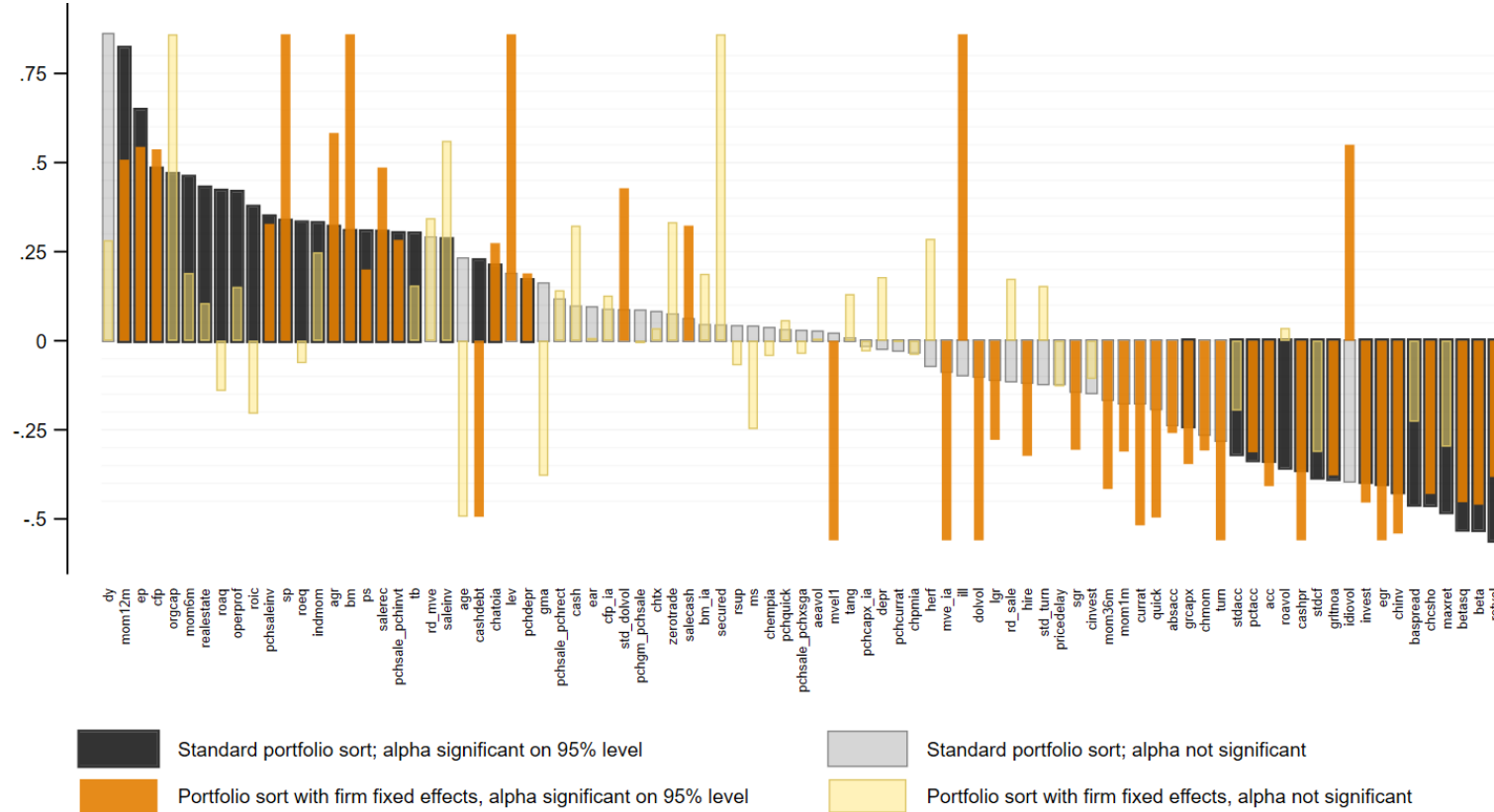
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-deciles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top decile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad \text{RMRF}_t \quad \text{SMB}_t \quad \text{HML}_t \quad \text{MOM}_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), and MOM (winners minus losers) representing monthly Carhart (1997) factor returns. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

**Figure IA.9: GPS models with pooled WLS, weighted FE, all variables standardized to zero mean, unit standard deviation – Carhart (1997) four-factor model**



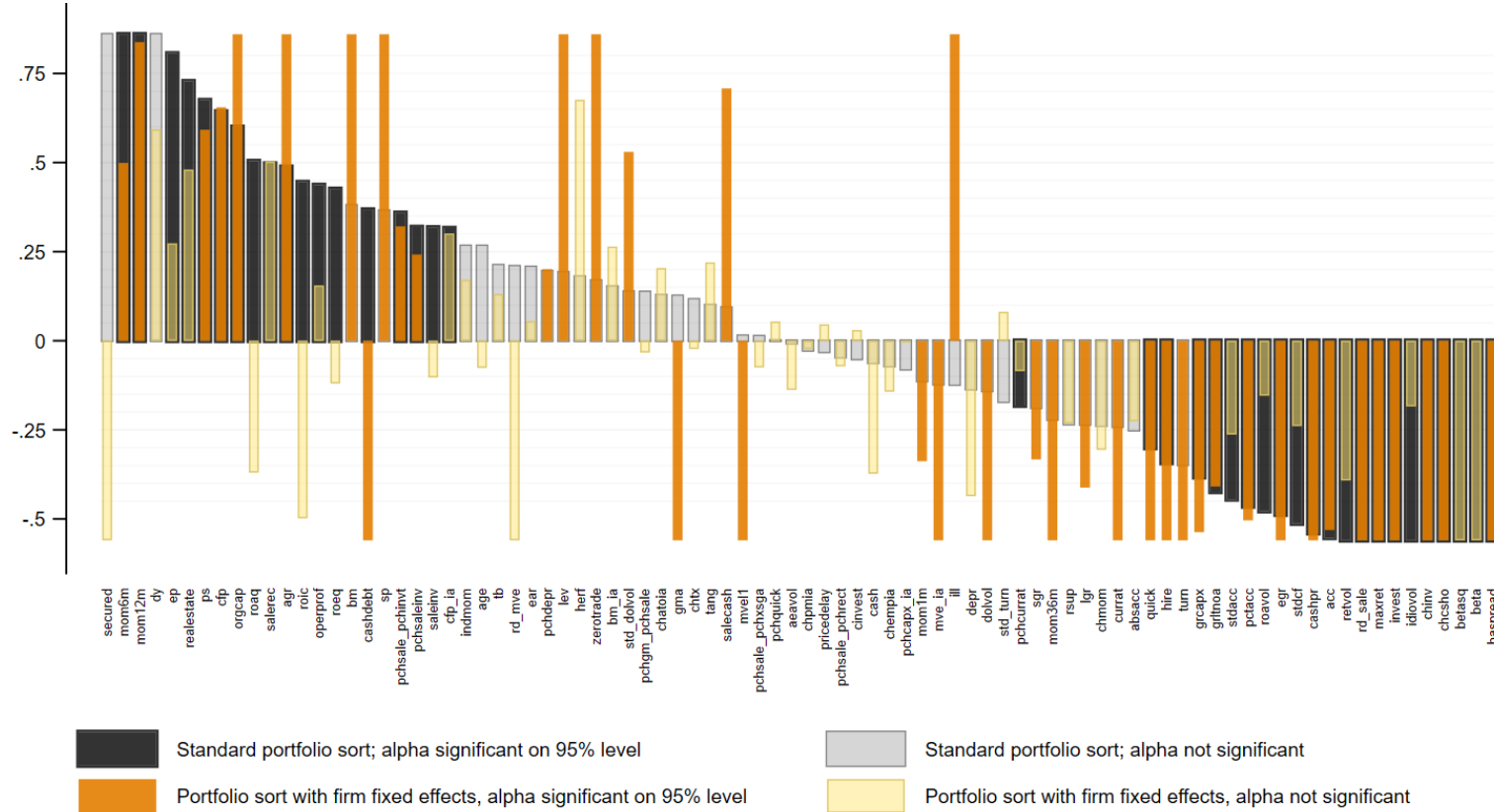
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t \quad SMB_t \quad HML_t \quad MOM_t]$  with RMRF (market excess return), SMB (small minus big), HML (high minus low), and MOM (winners minus losers) representing monthly Carhart (1997) factor returns. All firm characteristics of Gu, Kelly, and Xiu (2020, 2021) are cross-sectionally standardized with vector  $\mathbf{z}_{it}$  including a constant and the standardized firm characteristic. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

Figure IA.10: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (5 portfolios) – Market model



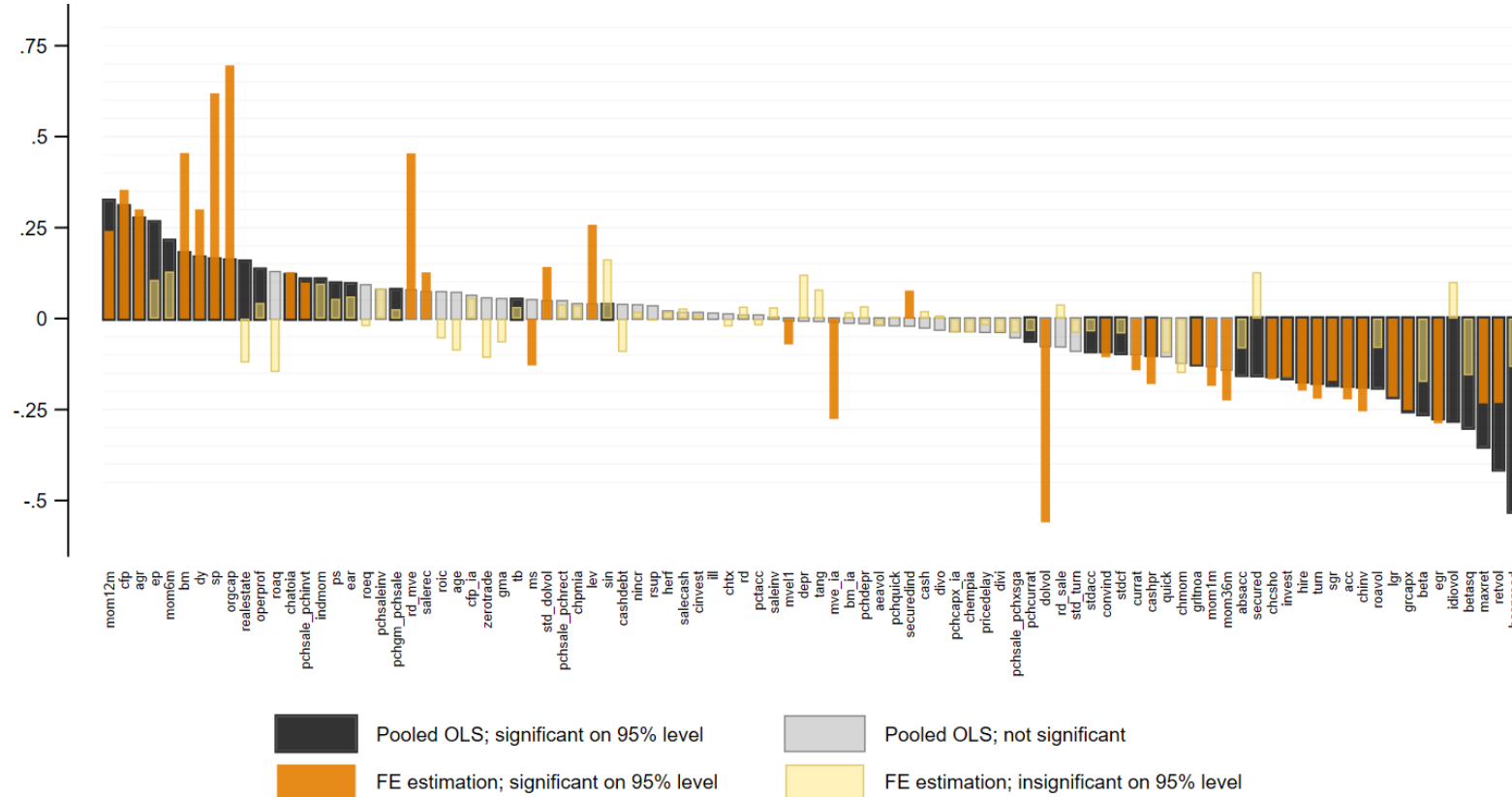
This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-quintiles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top quintile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t]$  with RMRF representing the market excess return. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

Figure IA.11: GPS models with pooled WLS, weighted FE, and top-versus-bottom portfolio sorts (10 portfolios) – Market model



This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). The sample only includes observations which belong to either the top- or bottom-deciles of observations sorted on the respective firm characteristic with vector  $\mathbf{z}_{it}$  including a constant and a dummy which is one for observations in the top decile of the respective firm characteristic. Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t]$  with  $RMRF_t$  representing the market excess return. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.

Figure IA.12: GPS models with pooled WLS, weighted FE, all variables standardized to zero mean, unit standard deviation – Market model



This figure displays results obtained from estimating GPS models either with pooled WLS or a fixed effects estimator testing the predictive power of each individual firm characteristic of the set of 94 asset pricing predictors used in Gu, Kelly, and Xiu (2020, 2021). Vector  $\mathbf{x}_t$  is set to  $\mathbf{x}_t = [1 \quad RMRF_t]$  with RMRF representing the market excess return. All firm characteristics of Gu, Kelly, and Xiu (2020, 2021) are cross-sectionally standardized with vector  $\mathbf{z}_{it}$  including a constant and the standardized firm characteristic. Each bar compares the coefficient estimates from standard portfolio sorts (black and gray) to coefficient estimates obtained from estimating a GPS model with FE (orange and yellow) for one of the 94 asset pricing predictors. In all models, observation weights are set equal to the beginning-of-time  $t$  value-weights of the stocks. Statistical inference is based on Driscoll and Kraay (1998) standard errors with a lag-length of three months. The sample period is July 1963 to December 2019.