Taste for Characteristics or Risk Factor Aversion? Evidence from Institutional Demand

Matthias Bank & Franz Insam & Jochen Lawrenz *

January 15, 2022

Abstract

We disentangle the relevance of risk factors versus stock characteristics in the cross-section of expected returns not only by analyzing covariance patterns in returns, but also by starting from economic first principles of price formation, i.e. by employing the information content of institutional portfolio holdings. Our main contribution is to show that excess demand from 13(f) institutional investors is strongly affected by ex-ante known stock characteristics but not so by risk factors. Furthermore, we find strong evidence that this characteristics-induced demand is pricing-relevant and therefore reflected in the cross-section of returns. Our results can explain recent evidence which shows little explanatory power for risk factors once tested on bias-corrected individual stock level. Robustness checks discard the concern that our results are driven by latent risk factors, poor factor proxies or that it can be easily arbitraged away. (142 words)

*Department of Banking & Finance, University of Innsbruck, Universitätsstrasse 15, 6020 Innsbruck, Austria, +43(0)512 50773100, corresponding email: jochen.lawrenz@uibk.ac.at

A. Introduction

Risk factor models are an integral part of canonical asset pricing theory and factor investing has become a well-established industry standard. Despite the enormous and prominent literature on multi-factor models, the academic research is still somewhat ambiguous with respect to the relative significance of factor- versus characteristics-based explanations of the cross-section of expected returns. Based on novel econometric approaches, a growing number of recent contributions cast renewed and serious doubts on risk-factor models by showing that stock characteristics carry substantial risk premia, while prominent risk factors turn out insignificant. Jegadeesh et al. (2019) conclude their analysis by stating that "*... none of the factors from those asset pricing models are associated with a significant risk premium after controlling for corresponding firm characteristics. [...], so it represents a puzzle that calls for further research.*". We add a piece to this puzzle by arguing that it will be insufficient to answer the question by focussing on the analysis of the covariance structure of returns alone, since – loosely speaking – even a characteristics-based demand for certain stocks can manifest itself empirically as a covariance pattern that, in turn, may be interpreted as a risk factor. Instead, we propose to exploit information on institutional investors' demand and show that institutional portfolio holdings are significantly driven by characteristics but not by factors and that their portfolio tilts carry over to the cross-section of expected returns.

Over the past decades, the literature has produced a staggering number of so-called characteristics-based anomalies, and not at least in response to Cochrane (2011)'s famous call to '*tame the factor zoo*', a rapidly growing body of literature attempts to establish order. Recent contributions by Kozak et al. (2018), or Kelly et al. (2019) discard characteristic anomalies and argue that only factor risk exposure matters by either invoking arbitrage arguments (Kozak et al., 2018) or by controlling for exposure to latent risk factors (Kelly et al., 2019). On the other hand, Chordia et al. (2015), Pukthuanthong et al. (2019), Jegadeesh et al. (2019), and Kim et al. (2021) find strong support for characteristics and against risk factors. Their novel appraisal of factor models is derived from econometric advancements to cope with the notorious errors-in-variables (EIV) problems. Being able to control the EIV problem and testing factor models on stock-level data leads them to reject many risk factors, which have long been well-known from studies based on portfolios as test assets. The apparent contradiction between the two strands of the literature reflects a deeper underlying identification dilemma, that has recently been stressed by e.g. Kozak et al. (2018) and Pastor et al. (2020). They emphasize that a sufficiently large characteristics-induced investor demand will generate empirical return patterns which can always mechanically be attributed to unobserved latent risk factors, but that such an interpretation will "*miss the underlying economics*".[1] We strongly agree with this assessment and argue that the attempt to disentangle taste for characteristics from aversion to risk factors should start from economic first principles of price determination, i.e. from modelling and analyzing investor demand. Our contribution does so by being – to the best of our knowledge – the first study that explains the relevance of factors vs. characteristics via an analysis of institutional portfolio holdings.

As the underpinning of our empirical analysis, we put forward a theoretical model where investors derive direct utility from stock/firm characteristics.[2] The resulting equilibrium pricing equation shows that beyond the usual covariance with state variables, the excess returns will depend on ex-ante known characteristics. We take this prediction to the data, whereby our empirical identification strategy proceeds in essentially three steps: (i) In line with much of the literature, the pricing equation can be tested within a Fama and MacBeth (1973) regression framework. Thereby, we follow the recent econometric improvements from Chordia et al. (2015), Pukthuanthong et al. (2019), Jegadeesh et al. (2019) to estimate factor and characteristics premia on individual stock level by implementing the mean beta approach of Pukthuanthong et al. (2019) as well as the instrumental variables approach recently proposed by Jegadeesh et al. (2019). (ii) Unlike prior literature, we then focus on institutional demand as dependent variable. More precisely, we measure excess demand for individual stocks by institutional 13(f)-filers and label it as their portfolio tilt. We regress stock-level portfolio tilts on factors and characteristics within a panel regression model in order to test if tilts are related to taste for certain characteristics or to risk factors. (iii) Even more importantly, in order to establish the economic link, we regress stock-level excess returns on investors' portfolio tilts within a panel regression to see if institutional investors' excess demand is pricing-relevant. If investors' demand is driven by characteristics and this excess-demand is pricing-relevant, than we conjecture the

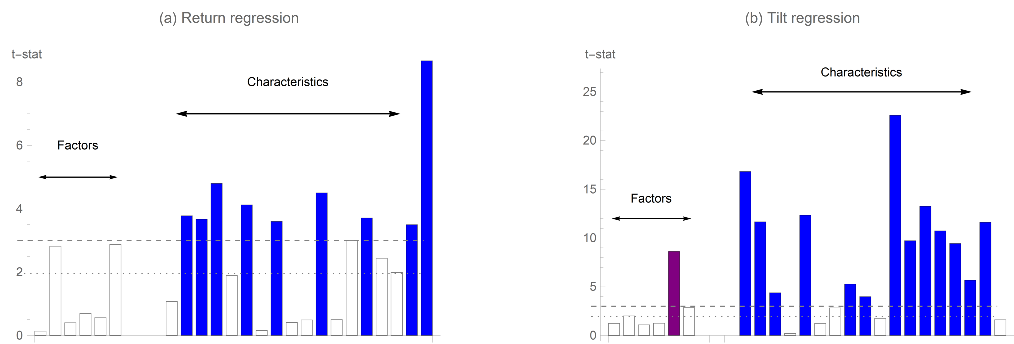
[1] Pastor et al. (2020), p. 7. In the same sense, Kozak et al. (2018) argue that any empirical 'horse race' between factors versus characteristics is futile if it is based on 'reduced-form' models, i.e. if it is done on the basis of return data alone or without any strong theoretical structure.

[2] Our theoretical model shares many similarities with Pastor et al. (2020). A notable difference is that Pastor et al. (2020) only consider the additional preference for ESG preferences, while our model is more general in the sense of allowing for an arbitrary number of characteristics. In this sense, our model is a straightforward extension to the Pastor et al. (2020) model.

link to be strongest when regressing the characteristics-induced return premium against the characteristics-induced excess demand.

We illustrate the core of our findings in Figure 1, which plots (absolute) regression t-stats for the six risk factors and 18 characteristics which we test on a comprehensive CRSP-sample (1975-2016, monthly data). The left panel reports results for excess returns as dependent variable, while the right panel shows results for institutional portfolio tilts. Following Harvey et al. (2016)'s suggestion to use a t-stat-hurdle of 3, only bars exceeding this threshold (indicated as the dashed line) are colored. Panel (a) shows that for excess returns none of the (usual) risk factors

Figure 1: Main result. We report t-stats for slope coefficients on risk factors and characteristics as bar chart from return regression (left panel) and tilt regression (right). Horizontal dashed lines indicate significance levels of the usual 1% level and the more strict level of $t = 3$ according to Harvey et al. (2016). Only t-stats exceeding 3 are reported as bold. From return regressions (left panel), we find that none of the 6 risk factors, but 9 out of 18 characteristics exceed $t = 3$. When regressing portfolio tilts of institutional investors on risk factors and characteristics, we find 1 out of 6 risk factors, but 13 out of 18 characteristics which exceed $t = 3$.



exceeds a t-stat of 3, while 9 out of 18 characteristics do so quite comfortably. This result is well in line with the recent literature and documents the poor performance of risk factors once they are tested on individual stock-level data.[3] While this result is not new to the literature, panel (b) highlights the novel contribution of our analysis. We take institutional investors' portfolio tilts as dependent variable and document a strikingly similar pattern. 13 out of 18 characteristics exceed a t-stat of 3 with values ranging up to 22, while there is only one single risk factor (rmw Robust-Minus-Weak) that turns out significant.[4] The similarity of results suggests that investors' demand strongly depends on various stock characteristics and that

[3]Under the common, but less strict 5% confidence level, the size and momentum factor would survive, while the number of significant characteristics would grow to 12 out of 18.

[4]On the 5%-level, we find momentum as well as size as (marginally) significant factors, whereas 14 out of 18 characteristics are significant.

these preferences carry over to the cross-section of expected returns. To rule out a spurious link, we apply various tests to corroborate the pricing-impact of institutional investors excess demand, and find strong support that our portfolio tilt measure explains excess returns. When regressing stock-level returns on portfolio tilts, the estimated coefficient turns out significantly negative. Besides confirming the pricing-impact, we also find that institutional investors' portfolios are (on average) tilted towards *under*performing stocks. Our negative relation confirms results by e.g. Edelen et al. (2016), who show that institutional investors have preferences to invest in the short leg of anomalies and thereby underperform ex post. We further decompose returns as well as portfolio tilts into factor and characteristics components and find that only the characteristics-induced tilt component is a significant explanatory variable, and that only the characteristics-induced excess return is significantly affected by the portfolio tilts. Taken together, our results provide novel demand-based evidence that taste for characteristics are significantly more important than aversion to risk factors in the explanation of the cross-section of expected returns.

We corroborate our findings by various robustness checks and discard several relevant concerns: First, we address the concern raised in e.g. Kelly et al. (2019) that characteristics may be driven by latent risk factors. We replace the six risk factors by those (four) latent factors identified by Kelly et al. (2019), and still find strong support for our result. Exactly the same set of characteristics remain strongly significant, while out of Kelly et al. (2019)'s four latent risk factors, only the first one is (marginally) significant. Second, we show that our analysis is robust to the risk-proxy hypothesis, which describes the concern that characteristics are (merely) better proxies for the true, unobserved risk factors. We do so by implementing the IV-mean estimator developed by Jegadeesh et al. (2019) and by allowing characteristics to be conditioning variables for factor loadings such as in e.g. Harvey and Liu (2021) and Hoechle et al. (2020). Within the latter approach, we complement our cross-sectional approach by panel models. Third, we address the argument raised in e.g. Kozak et al. (2018) that arbitrage forces should drive out the pricing-impact of characteristics. We show that the aggregate institutional portfolio tilts in our characteristics display strong persistence over long horizon which imposes high practical arbitrage restrictions and thereby prevent characteristics from being arbitraged away. Finally, we drop stocks with the smallest 20% (NYSE breakpoints) market capitalization and show that we find weaker but still comparable and consistent results within the large stock sample.

Our main novel contribution to the literature is the demand-based evidence for the relevance of stock/firm characteristics. Thereby, we restrict our attention to (nine) prominent aspects such as: Size, value, momentum, liquidity, volatility, systematic risk, distress, profitability and investment. We propose a novel identification strategy for these characteristics which consists in transforming them into (binary) indicator variables rather than treating them as metric variables. The indicator is defined by the event that at time t for stock i, the (metric) characteristic variable $\Phi_{i,j}$ was part of the α-quantile of the cross-sectional characteristics distribution $F(\Phi)$ for an extended period of the last T months, i.e. we define characteristics *labels* as $\phi_{i,j,t} = 1_{E_{i,j,t}}$, with event $E_{i,j,t} = \{\left(\Phi_{i,j,s} > F(\Phi)^{\alpha}_{i,j,s}\right)_{t-T \leq s \leq t-1}\}$. Essentially, the definition captures the idea that a stock is perceived as having a certain characteristic only if its realized value was among the top (bottom) quantile for several consecutive months, i.e. if the realized characteristic is a salient signal to investors both in the cross-section and over time. Applying the indicator to the top (α) and bottom ($1 - \alpha$) quantile for the nine characteristics above leads to the consideration of 18 binary characteristics. Our rationale for defining (binary) characteristics labels is fourfold: First, it is inspired by the usual portfolio sorts approach in the tradition of Fama and French (1993). Unlike for the factor construction, we do not form long-short portfolios, but attribute characteristics labels on individual stock level. Second, our construction follows agency-theoretic arguments such as the ones raised in Lakonishok et al. (1992) and more recently by Edelen et al. (2016) who emphasize that delegated asset managers prefer investments which are easy to communicate to clients and therefore they react to salient signals of individual stocks. Third, a growing strand of the literature, exemplified by Shiller (2017) and Hirshleifer (2020), stresses the role of narratives in financial markets. Investing is a social activity and narratives emerge from repetitive similar signals, leading to the manifestation that a certain stock is perceived as being, say, a particularly illiquid stock. Fourth, on the methodical level, our binary labels mitigate the above-mentioned concern that characteristics could be better proxies for the true betas. Jegadeesh et al. (2019) show that an estimator which is robust to this concern uses a time-series average of characteristics.[5] Therefore, since our characteristics labels are defined over the past T months (in the main analysis, we use $T = 6$ months), our results are less prone to this potential misattribution problem.

In particular against the background of delegated asset management, we use data from institutional investors portfolio holdings to test the relevance of characteris-

[5] See their IV mean-estimator in Proposition 2 and equation (25) in Jegadeesh et al. (2019).

tics labels. We use quarterly holdings data from institutional 13(f)-filings between 1990Q1 and 2015Q4. As empirical proxy for institutional excess-demand, we introduce portfolio *tilts*, which we operationalize as the percentage share of an individual stock i aggregated over all 13(f)-filers relative to its market capitalization (in percentage). Thus, tilt values above (below) one imply that institutional investors overweigh (underweigh) a particular stock relative to the market benchmark. In contrast to other measures such as the number of institutional investors, or the percentage of shares held by institutionals, which has been used in prior literature (see e.g. Edelen et al. (2016)), our proposed tilt measure has the advantage of being readily interpretable as excess demand and accounts for time-varying levels of aggregate institutional ownership.

Our contribution is related to a voluminous literature on factor identification, which Pukthuanthong et al. (2019) consider as the (arguably) most important topic in finance.[6] Equilibrium asset pricing literature predominantly relies on the idea, succinctly summarized by Cochrane (2005) that "*asset returns depend on how you behave, not who you are - on betas rather than characteristics [...]*".[7] But ever since the documentation of the size (Banz, 1981), book-to-market (Stattman, 1980), or leverage effect (Bhandar, 1988) and the development of the Fama and French (1993) three-factor model, there is ongoing discussion about the relevance of risk factors versus firm/stock characteristics. After early contributions by e.g. Daniel and Titman (1997) and Davis et al. (2000), this dispute more recently received renewed and reinforced interest spurred by a staggering proliferation of 'new' factors and anomalies. Harvey et al. (2016) count over 316 pricing relevant factors and characteristics in the literature. Green et al. (2017) identify 94 characteristics and find that only few of them provide independent information. Pukthuanthong et al. (2019) emphasize that not all contributions draw a sharp distinction between (known) characteristics and risk factor which leads to confusion in the usage of the terminology. Thus, two (related) strands have emerged in the literature that tackle the issue from different directions. First, different approaches have been proposed to reduce dimensionality. While Harvey et al. (2016) highlights the multiple-testing problem and suggests to increase the t-stat hurdle, papers such as Kelly et al. (2019), Light et al. (2017), Kim et al. (2021), Feng et al. (2020) use more advanced econometric techniques to link (a large number of) observed characteristics to (only few) unobservable latent factors. In particular, Kelly et al.

[6](See Pukthuanthong et al., 2019, p. 1575)
[7](Cochrane, 2005, p. 79)

(2019) argue by using instrumented principal component analysis that latent factor risk exposure rather than characteristics explain average returns. Second, in the spirit of Daniel and Titman (1997), a number of recent contributions such as Chordia et al. (2015), Jegadeesh et al. (2019), and Pukthuanthong et al. (2019) put their focus on the head-to-head comparison of risk factors versus characteristics by using different econometric improvements to overcome methodological problems in the Daniel and Titman (1997) analysis that have been discussed e.g. in Kozak et al. (2018) and Kim et al. (2021). In particular, Chordia et al. (2015) and Jegadeesh et al. (2019) propose ways to deal with the errors-in-variables problem which in turn allows estimation on individual stock level in contrast to the usual portfolio approach. Their results are significant challenges to risk factor models as they find little to no support for the common, well-established risk factors, but strong support for characteristics.

Any horse-race between factors and characteristics faces at least two challenges: First, Kozak et al. (2018) emphasize that arbitrageurs are supposed to eliminate any profit opportunities that arise from sentiment driven demand distortions which are not aligned with common factor covariances, and that mispricing can only persist if arbitrageurs are exposed to factor risks when trading against sentiment investors' demand distortions. Second, the finding of significant characteristics leaves open the possibility that they capture the statistical significance not due to being a genuine alternative explanatory variable, but due to being a better proxy for the (unobservable) true risk factor(s). This so-called risk-proxy hypothesis is addressed by various papers such Kelly et al. (2019), and Kim et al. (2021). However, approaches such as the instrumented PCA of Kelly et al. (2019) let the pendulum swing back to the other extreme in the sense that almost all characteristics are found to be (merely) proxies for latent risk factors, which can intuitively be understood by the fact that such econometric dimension-reduction techniques are prone to attributing any covariance patterns to latent risk factors. Kim et al. (2021) instead propose a PCA-based method which is designed to give characteristics maximal explanatory power for risk loadings before concluding that characteristics represent mispricing. With this refinement, they show that after controlling for the risk-proxy hypothesis, there is still strong support for characteristics. On a general level, although econometric identification strategies are constantly improving, the approach of relying on return data alone will always be subject to the dilemma recently discussed in Kozak et al. (2018) and Pastor et al. (2020) that also a common taste for characteristics can manifest itself as an omitted priced risk factor. Therefore, in order to not miss the underlying economics, it appears inevitable to turn to first principles of price determination which is to analyze investor demand directly. The recent review article by Brunnermeier et al. (2021) underscores the importance of intermediary asset pricing and asks to develop successful asset pricing models, which "*explain not only prices, but both prices and quantities, including portfolio holdings and flows*".[8] Our contribution speaks to this call.

B. Methodology

B.1. Modeling demand towards stock and firm characteristics

As underpinning to our empirical approach, we put forward a theoretical model which allows for an arbitrary number of risk factors as well as particular tastes for known characteristics. We derive explicit demand functions and the equilibrium pricing equation. This section sketches the main steps, while Appendix A contains the detailed model derivation and description. We consider a market with K investors whose risky final wealth $W_1^k = W_0^k[1 + (r_f + \mathbf{x}_\mathbf{k}'(\boldsymbol{r} - r_f\mathbf{1}))]$ is derived from investing in N assets. W_0^k denotes initial wealth, r_f the risk-free rate, $\boldsymbol{r}$ the return vector of the $N-1$ risky assets, and $\mathbf{x}_\mathbf{k}'$ the holdings vector. We next depart from the standard modelling approach by specifying the following general utility function.

$$u^k(W_1^k, \boldsymbol{\phi}, \boldsymbol{Z}) = -e^{-a^k W_1^k + a^k \sum_j^J b_j^k \phi_j + a^k \sum_h^H c_h^k Z_h}, \tag{1}$$

Besides deriving utility from terminal wealth, we allow investors' utility to be affected by an arbitrary number of characteristics ϕ and risk factors Z. a^k measures constant absolute risk aversion, b_j^k is the (relative) attitude towards the asset's characteristic j, and c_h^k is the attitude towards risk factor Z_h impacting the asset returns. We impose no restriction on the sign of b_j^k, which means that investors can derive non-pecuniary utility or disutility. Note that characteristics ϕ_j are not assumed to be random variables but ex ante known quantities.[9] Determining the expected utility and solving the first-order condition, leads to the optimal portfolio

[8] Brunnermeier et al. (2021)p. 2144.

[9] Note further, that by setting b_j^k and c_h^k equal to zero for all investors, the model simply leads to the well-known Sharpe-Lintner-Mossin capital asset pricing model.

weights which can be expressed as

$$\mathbf{x_k^*} = \frac{\mathbf{\Sigma^{-1}}(\boldsymbol{\mu} - r_f\mathbf{1})}{a^k W_0^k} - \underbrace{\sum_{j=1}^{J} \frac{\mathbf{\Sigma^{-1}}\hat{b}_j^k \boldsymbol{\phi}_j}{a^k W_0^k}}_{\text{characteristics taste}} - \underbrace{\sum_{h=1}^{H} \frac{\mathbf{\Sigma^{-1}} c_h^k \mathbb{C}ov(\boldsymbol{r}, Z_h)}{a^k W_0^K}}_{\text{factor risk aversion}}. \tag{2}$$

The demand function in (25) consists of three parts. While the first term reflects the standard systematic market risk and the last part represents further factor risk in the usual sense of covariance, the middle part shows that demand is explicitly affected by investor k's taste for known characteristics. Agents with positive preferences for a particular stock characteristic thus receive non-pecuniary utility from investing into assets that exhibit a particular characteristic and are therefore willing to accept a lower expected returns. The model by Pastor et al. (2020) is a potential example for this type of characteristic preferences, which they interpret as preferences for green (i.e. sustainable) stocks. Furthermore, we build upon the argument in Pastor et al. (2020) that as long as the mass of agents that care about a characteristic is non-zero, this leads to differences in expected returns among stocks and alphas when regressing stock returns on factors. Hence, any relation between characteristics and returns is fundamentally driven by heterogenous investors' preferences and the resulting demand distortions. The demand function (25) represents also the theoretical basis of defining a relative portfolio tilt in stock i by (a subset of) investors. Aggregating demand over all K investors leads to the market weights vector $\mathbf{x_M} = \sum_{k=1}^{K} \mathbf{x_k^*} W_0^k / W_0^M$, where W_0^M denotes market capitalization, i.e. the aggregate invested wealth (in risky assets). By splitting investors into a subset of P professional investors and the remaining $R = K - P$ retail investors, we can decompose market weights and define (the vector of) professional investors' portfolio tilts $\mathbf{T}^P$ (by elementwise vector division, i.e. the Hadamard operator $\oslash$) as

$$\begin{aligned} \mathbf{x_M} W_0^M &= \sum_{k=1}^{P} \mathbf{x_k^*} W_0^k + \sum_{k=P+1}^{K} \mathbf{x_k^*} W_0^k \equiv \mathbf{x_P} W_0^P + \mathbf{x_R} W_0^R, \\ \mathbf{T}^P &= \mathbf{x_P} \oslash \mathbf{x_M}. \end{aligned} \tag{3}$$

From the demand function (25), we derive equilibrium expected returns which are determined by the following central pricing equation:

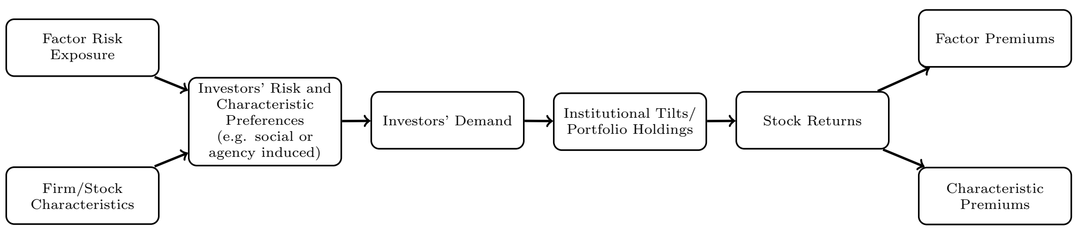
$$\boldsymbol{\mu} - r_f\mathbf{1} = \kappa_M \mathbf{\Sigma} \mathbf{x_M} + \sum_{j=1}^{J} \kappa_j \boldsymbol{\phi}_j + \sum_{h=1}^{H} \kappa_h \mathbb{C}ov(\boldsymbol{r}, Z_h). \tag{4}$$

Analogously to the demand function, equilibrium expected excess returns can be decomposed into three parts, where κ_M, κ_j, and κ_h aggregate the risk aversion and attitude parameters a, b, and c.[10] With appropriate rearrangements, equation (4) can be transformed into the beta representation $\mu_i - r_f = \beta_{iM}\hat{\kappa}_M + \sum_{h=1}^{H} \beta_{ih}\hat{\kappa}_h + \sum_{j=1}^{J} \kappa_j \phi_{ij}$, which shows that the excess return of asset i is a function of the asset's beta with the market, the H mimicking portfolios, the return contribution of the market plus the return contributions of the relevant characteristics j. Note that betas in this representation will have the interpretation of standard bivariate OLS betas, not multivariate regression betas. We consider the beta representation of equation (4) as basis for our empirical approach, which is outlined in the next sections.

B.2. Empirical Methodology

While the beta representation of (4) can be considered as the theoretical motivation for regressing returns on characteristics and factors, our empirical analysis strives for the more ambitious goal of identifying the relevance of characteristics versus risk factors in investors' revealed preferences and thus to shed light on the transmission channel from investors' preferences to the pricing of the cross-section. Figure (2) illustrates our identification strategy in a stylized way and shows our basic idea that investors have specific preferences with respect to factor risk exposure and known firm and stock characteristics. Investors' preferences are reflected in their demand and portfolio choices, which in turn may find their repercussion in stock returns and their corresponding risk and characteristics premia. We distinguish three steps in

Figure 2: In this figure we provide a stylized and simplified illustration of the proposed transmission channel from (institutional) investor's risk aversion and characteristic preferences into factor and characteristic related return premiums. Investors observe stock characteristics and assess the factor risk of stocks. Based on their internal risk aversion and preferences for characteristics this leads to demand for stocks/assets. Investors reveal their characteristic preferences and risk aversion through their observable portfolio holdings/tilts. Demand tilts may translate into factor and characteristic related return premiums.



[10] See the appendix for details.

our analysis: First, in line with a voluminous literature, we run Fama and MacBeth (1973) regression of monthly returns on factor betas and characteristics to directly test for the pricing relevance of factors vs. characteristics. Second, we run panel regressions of institutional demand tilts on factor betas and characteristics, thereby testing for revealed characteristic preferences and risk aversion. Third, we conduct panel regressions of quarterly returns on past demand tilts to test whether factor tilts are pricing relevant. This test aims at establishing the economic link, which explains if preference-induced demand manifests itself in return premia. While the results from the first step are not new to the literature, we are – to the best of our knowledge – the first to address steps two and three.

Fama and MacBeth (1973) Pricing Test and EIV Correction

First, we estimate return contributions of factors and characteristics by applying the Fama and MacBeth (1973) two-step regression approach. We run monthly cross-sectional regressions of stock-level excess returns r_{it}^{ex} on historical factor betas and binary characteristic labels. The set of risk factors includes the five factors according to Fama and French (2015) augmented with the momentum factor of Carhart (1997). The factor betas (i.e. historical factor risk exposures) for these six risk factors are themselves estimates from a first step rolling time-series regression. As characteristics, we include: Trading liquidity, market capitalization, book-to-market ratio, momentum, return-volatility, market beta, financial distress, asset-growth and operating profitability. We implement characteristics as a binary variable which we interpret as characteristics label whose construction is outlined in more detail in Section B.3 on data description.

The cross-sectional regression equation can be summarized as

$$r_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{h,t} \hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{j,t} \phi_{i,j,t-1} + \epsilon_{i,t} \quad \forall\ t = 1, ..., T \tag{5}$$

In this regression γ_{0t} is the unexplained alpha, γ_{ht} are the return contributions of the factors, and γ_{jt} are the return contributions of the binary characteristics labels of month t. The average premium associated to factor h and characteristic j is estimated as the time-series average of the individual factor and characteristic premium estimates.

It has long been recognized, that this two-step estimation procedure leads to a potential errors-in-variables bias. In order to avoid the portfolio sorts approach (which is known to introduce a factor structure itself) and to run our tests on

individual stock level, we implement two recently proposed EIV corrections. First, we follow the simple approach by Pukthuanthong et al. (2019) who replace the individual stock beta $\hat{\beta}_{i,h,t-1}$ by the equal weighted average beta of all stocks which were assigned within the same size-market factor category at time $t-1$. The mean beta is held constant for one calendar year starting in July. The procedure is repeated for all h factors. We refer to this EIV correction method as the mean beta approach. Second, we implement the recently developed instrumental variables approach of Jegadeesh et al. (2019), who propose to estimate betas from disjunct sets of odd and even calendar months. In the cross-sectional second stage regression odd month betas can be used as explanatory variables and even month betas as instruments when month t is an odd month and vice versa for even months. Jegadeesh et al. (2019) show that the measurement errors of odd and even month betas are uncorrelated – therefore the instrumental variables method yields consistent risk premium estimates.
We contrast both EIV corrections with simple (uncorrected) OLS results. As we will discuss in more detail in the results section (i.e. C.1), we find that our main results are not affected by the particular choice of EIV correction and are actually close to OLS results. We find that the IV approach by Jegadeesh et al. (2019) suffers from an increasing number of severe outliers when estimating the full model, thus we report results from the mean beta approach by Pukthuanthong et al. (2019) as our baseline specification.

Using the estimated return contributions from the Fama and MacBeth (1973) regression it is possible to decompose cross-sectional return variation into relative contributions of factor loadings/betas and characteristics. This was shown and implemented by Chordia et al. (2015) (recently refined by Raponi et al. (2020)) and we apply a slightly adapted version of their approach. In order to compute the model's explained return we use the individual monthly estimates $\hat{\gamma}$ instead of the full sample means as in Chordia et al. (2015). This adapted approach has the advantage that it allows to control for structural changes over the sample period, meaning that it allows for more variation in the explanatory power of factors vs. characteristics. Specifically, the excess return $\hat{r}_{it}^{ex}$ of stock i, which is explained by the model in month t is[11]

$$\hat{r}^{ex}_{it} = \gamma_{0t} + \underbrace{\sum_{h=1}^{H} \hat{\gamma}_{ht}\hat{\beta}_{i,h,t-1}}_{\hat{r}^{ex,beta}_{it}} + \underbrace{\sum_{j=1}^{J} \hat{\gamma}_{jt}\phi_{i,j,t-1}}_{\hat{r}^{ex,char}_{it}} \qquad \forall\ t = 1, ..., T \tag{6}$$

where $\hat{r}^{ex,beta}_{it}$ and $\hat{r}^{ex,char}_{it}$ are defined as the the factor and characteristic components of the model's explained return. Our goal of the decomposition is twofold. First, similar to Chordia et al. (2015) we compute from the above Equation (6) the cross-sectional variance of the explained/fitted returns $\mathbb{V}_{CS}(\hat{r}^{ex}_t)$ for each month t across all stocks i and additionally the cross-sectional variance of the beta $\mathbb{V}_{CS}(\hat{r}^{ex,beta}_t)$ and characteristic component $\mathbb{V}_{CS}(\hat{r}^{ex,char}_t)$. Following Chordia et al. (2015), we define ratios

$$\mathbb{VR}^{beta}_t = \frac{\mathbb{V}_{CS}(\hat{r}^{ex,beta}_t)}{\mathbb{V}_{CS}(\hat{r}^{ex}_t)}; \qquad \mathbb{VR}^{char}_t = \frac{\mathbb{V}_{CS}(\hat{r}^{ex,char}_t)}{\mathbb{V}_{CS}(\hat{r}^{ex}_t)}. \tag{7}$$

$\mathbb{VR}^{beta}_t$ and $\mathbb{VR}^{char}_t$ can be interpreted as the relative contribution of factor and characteristics loadings respectively to the overall explanatory power of the full model for each month t. The ratios do not necessarily add up to one, due to the potential covariation between the factor and beta component.[12]Second, we use the return decomposition in our third step – described in more detail below – where we assess the impact of investors demand on either component in isolation.

Portfolio Tilt Regressions

As second step, we investigate whether institutional demand tilts are related to characteristics or past factor risk exposure. We run a panel regression of individual

[11]Note that the main difference between the approach of Chordia et al. (2015) and our approach is that Chordia et al. (2015) estimate expected returns using the full sample time series averages of the estimated γ coefficients for the factors and characteristics. They argue that the time series averages of the γ coefficients are closer to the true return premia and therefore allow for a better approximation of the true variation in expected returns. Indeed we argue that using the full sample time series average of the γ coefficients does not allow to observe any structural changes over time. Therefore we decide to use the individual monthly estimates. Robustness checks of our results, where we apply the same method as Chordia et al. (2015) show that the insights of the return decomposition do not change.

[12]Note that Chordia et al. (2015) propose an errors in variables correction for the estimated factor loadings/betas. We decide not to incorporate this EIV correction. First the γ coefficients are already corrected for a potential EIV bias. Secondly as described by Chordia et al. (2015), the EIV bias leads to an overstatement of the cross-sectional variation in the true loadings. Therefore, we argue that using the non-EIV corrected betas likely underestimates the true contribution of the characteristics relative to the factors. Hence the non-EIV corrected results are a conservative estimate for the true contribution of characteristics relative to factors.

stock tilts (in the sense of Equation (3)) at quarterly frequency on binary stock characteristic labels and historical factor betas.[13] We argue, based on our theoretical model, that institutional investors' reveal their preferences with respect to characteristics and factor risk exposure through their portfolio disclosures. The explanatory variables include the binary characteristic labels from Section (B.3) and historical factor betas. We estimate the historical factor betas from univariate time series regressions of monthly stock returns over a 36-month rolling window on the five Fama and French (2015) factors and the Carhart (1997) momentum factor. Consistent with our theoretical model (see Section A.1), we use univariate factor betas. Thus, we estimate the model

$$T_i = \alpha_0 + \sum_{h=1}^{F} \theta_{h,t} \hat{\beta}_{i,h} + \sum_{j=1}^{J} \theta_{j,t} \phi_{i,j} + c_i + p_t + \epsilon_{i,t}. \tag{8}$$

Using a panel approach has the advantage to control for fixed effects. We include firm fixed effects (c_i, where $\mathrm{E}(c_i) = 0$) to account for unobserved heterogeneity and time fixed effects p_t to control for time trends.[14]

As robustness check and in order to be consistent with the return regression, we further implement the Fama and MacBeth (1973) two-step regression approach for portfolio tilts as dependent variable, i.e. we use the model

$$T_{i,t} = \alpha_0 + \sum_{h=1}^{H} \theta_{h,t} \hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \theta_{j,t} \phi_{i,j,t-1} + \epsilon_{i,t}, \qquad \forall t = 1, ..., T. \tag{9}$$

Linking Portfolio Tilts to Expected Returns and Return Premia

In the third and final step, we address how institutional investors' demand tilts are related to consecutive stock returns via a panel regression model. As the institutional demand tilt data has quarterly frequency, we compute quarterly returns from the monthly CRSP return data and match them to the quarterly 13(f)-filings data via the CUSIP identifier. We aim to address whether institutional demand tilts are related to consecutive quarterly stock returns. Therefore we regress the quarterly stock returns of quarter t on the institutional demand tilts of the prior quarter $t-1$.[15] Additionally we also control for firm fixed effects c_i to control

[13] The quarterly frequency is dictated by data availability for 13(f) filings.

[14] Controlling for time fixed effects in this model appears to be particularly reasonable, as e.g. over the last decades the overall market share of institutional investors increased significantly.

[15] For example this means that the return of the first quarter of a year (i.e. the return that is realized between the 01.01. and the 31.03.) is regressed on the tilt that is computed from the

for unobserved heterogeneity. We also include period fixed effects p_t to control for market fluctuations in excess returns. The panel regression model is:

$$r_{it}^{ex} = \alpha_0 + \beta \times T_{i,t-1} + c_i + p_t + \epsilon_{i,t} \tag{10}$$

B.3. Data Set and Variable Definition

Our empirical methodology relies on two main data types. First, institutional portfolio holdings data, and second, cross-sectional characteristics and return data.

Portfolio Holdings Data

Regarding the institutional investors' portfolio holdings we obtain quarterly portfolio holdings from institutional 13(f)-filings through the Thomson Reuters institutional ownership database via WRDS. We exclude bad filings data, whenever the aggregate number of stocks filed by institutions exceeds the number of shares outstanding for a particular stock. We compute institutional investors portfolio tilts for individual stocks according to Equation (3) and winsorize them at the 97% quantile. Corresponding descriptive statistics are provided in Table (13) of Appendix (B). The sample period is from the first quarter 1990 to the fourth quarter of December 2015. Within the full sample the mean institutional portfolio tilt is 0.705. When we remove the smallest 20% of stocks, the mean tilt is 1.07, thus close to the market weights. The standard deviation in this case is 0.361 showing substantial variation. We report further properties of the tilt measure in the context of robustness checks in Section D.3.

Return and Characteristics Data

We obtain market data from the CRSP database and financial statement data from the Compustat database. The data is available from 1962 onwards, while we choose our sample period from January 1975 to June 2016. We include all common stocks (share codes 10 and 11 in CRSP) and exclude financial firms. The set of risk factors[16] includes the market (*mkt*), small-minus-big (*smb*), high-minus-low (*hml*), conservative-minus-aggressive (*cma*) and robust-minus-weak (*rmw*) factors according to Fama and French (2015) augmented with the Carhart (1997) momentum factor (*mom*), which we obtain from Kenneth French's webpage.[17]

fourth quarter's 13(f) filings (the reporting date is the 31.12.).

[16]We label risk factors by lower-case letters to distinguish them from upper-case labelled characteristics.

[17]See `mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

For each firm in the sample we obtain and compute several firm and stock characteristics – including market capitalisation, price to book value, monthly return volatility, illiquidity, return-momentum, market beta, the Campbell et al. (2008) distress measure, asset growth and operating profitability. Table (1) summarizes the characteristic variables construction and the corresponding lag structure of the market and accounting data. The applied lag-structure follows the established approach in the literature to lag accounting data by 4 months and market data by 1 month (e.g. see Green et al. (2017)).

Table 1: This table provides an overview on the characteristics that we use as the basis for the construction of the characteristic dummy indicator variables, as described in Section (B.3). Column 1 contains variable labels. Column 2 describes characteristics as usually referenced within the asset pricing literature. Column 3 reports how the characteristics are measured. In Column 4 we report the data frequency of the corresponding characteristic. Column 5 shows the applied lag structure. Column 6 indicates the lower and upper quantile and the time window used in the construction of the indicator variable.

ϕ	**Characteristic**	**Characteristic Measure Φ**	**Data Frequency**	**Lag**	**Indicator**
SZ	Size	Market capitalization in Mio. US-Dollars	Monthly	1 month	30/70/6M
PB	Value/Growth	Price to book ratio	Quarterly	4 months	30/70/6M
MO	Momentum	6 month momentum; 1 month skipped ($t-7$ to $t-2$)	Monthly	1 month	30/70/-
IL	Trading Liquidity	Amihud (2002) Illiq measure	Monthly	1 month	30/70/6M
VO	Volatility	Monthly return volatility	Monthly	1 month	30/70/6M
BT	Market Beta	Rolling 24-month market beta	Monthly	1 month	30/70/6M
FD	Financial Distress	Campbell et al. (2008) distress measure	Quarterly accounting data, monthly and daily market data	4 month[18]	30/90/12M
OP	Profitability	Operating profitability according to Fama and French (2015)	Quarterly revenues, costs and book equity	4 months	30/70/12M
IV	Investments	Change in total assets similar to Fama and French (2015)	Quarterly total assets	4 months	30/70/12M

We restrict our sample to all stocks, where the full set of characteristics is observable. Particularly the Campbell et al. (2008) distress measure requires that several accounting measures and market data are completely observable, thus our sample only includes stocks that offer comprehensive data availability. The final sample on average includes 3,745 firms per month. In Table 13 of Appendix B we report descriptive statistics of the stock returns and the individual stock and firm characteristics.

[18]Computation similar to Campbell et al. (2008) with accounting data: 2 months lagged and market data: 1 month lagged; the resulting Campbell et al. (2008) distress measure is lagged by another 2 months to finally obtain a total 4 month lag for the accounting data

Consistent with our theoretical model from Section A.1, we translate all stock characteristics into binary dummy variables that identify the top and bottom stocks based on a characteristic sorting. More precisely, the indicator is defined by the event that at time t for stock i, the (metric) characteristic $\Phi_{i,j,t}$ was part of the α-quantile of the (contemporaneous) cross-sectional characteristics distribution $F(\Phi)_{j,t}$ for an extended period of the last T months, i.e. we define $\phi_{i,j,t}$ as

$$\phi_{i,j,t} = 1_{E_{i,j,t}}, \quad \text{with event:} \quad E_{i,j,t} = \left\{ \left(\Phi_{i,j,s} > F(\Phi)^{\alpha}_{i,j,s} \right)_{t-T \leq s \leq t-1} \right\}. \tag{11}$$

By construction, $\phi_{i,j,t}$ can be interpreted as characteristic *labels*, reflecting the idea that investors recognize stock aspects in particular if they are observed over a longer time-period and if the realization of that characteristic is sufficiently salient. Our base case parameter choice therefore refers to the 30% and 70%-quantile over six consecutive months. As indicated in column 6 of Table 1, we deviate from this choice only for distress FD, profitability OP, and investment IV, for which we consider the last 12 month since the data frequency is quarterly, and in the case of FD, we restrict the upper quantile to 90% as empirically only a small fraction of all firms is in financial distress.[19] Regarding the Amihud (2002) Illiq-measure, the monthly return volatility, market capitalisation, price to book ratio and the rolling 24-month market beta, we split the sample at the 50% quantile.

We check for the dependence structure among the characteristic labels, by analyzing their variance inflation factors (VIF) which we report in Table (14) of Appendix B. We observe that the characteristic dummies for high liquidity IL^H and large market-capitalization SZ^H are highly correlated. Green et al. (2017) propose to exclude characteristics that have a variance inflation factor that exceeds 3. We therefore run separate empirical analysis, where we exclude the liquidity characteristic identifiers.[20] We find that none of our main empirical findings are sensitive to including the liquidity characteristic.

To better understand the dynamic properties of the characteristic dummies, we calculate the number of times an individual stock changes its status of being assigned a given characteristic. We report these results as histograms in Figure 7 in Appendix B. For the size characteristic SZ the mean change of status is less than 3, consistent with the fact that the ordering of market capitalization has little

[19]Since momentum MO is already defined over the last six months (i.e. within $t-7$ to $t-2$), no (further) time window criteria is necessary.

[20]We decide to exclude the liquidity characteristic identifiers, as the size characteristic is more prominent in the literature and typically has higher data reliability compared to the trading data.

variation over time. In contrast, for the momentum characteristic MO the mean number of changes is 20.

Merging the Datasets

We merge the monthly characteristic dummies and historical factor betas from the monthly CRSP and Compustat sample to the quarterly portfolio holdings data via the CUSIP identifier.[21] The stocks from the asset pricing sample are (almost) completely in the sample of the institutional 13(f) portfolio holdings. Thus, we can rule out that our results are driven by data gaps between the two datasets. For each quarter, we match the characteristic dummies and factor betas that are available at the beginning of a quarter (i.e. for example for Q4 we use the characteristic dummies and factor betas computed by the end of September) to assure that our results are not driven by any forward looking bias. We are aware of and take into account the reporting delay in 13(f) filings, but since our interest is not in exploiting 13(f) information in a real-time trading setting, it imposes no constraint on our analysis.

C. Empirical Results

C.1. The Pricing of Stock Characteristics

This section starts by reporting results from the two-stage Fama and MacBeth (1973) cross-sectional regression procedure with the mean beta EIV-correction from Pukthuanthong et al. (2019), as represented in Equation (5). Note, that according to our model from Section A.1, we use the univariate factor betas in the second stage of the Fama and MacBeth (1973) regression. This offers the interpretation that the estimated coefficients correspond to orthogonalized return contributions of factors and characteristics.

Overall, the results, which are summarized in Table 2, show that stock characteristics carry large and significant return premiums, while most of the risk factors are not priced. This findings is well in line with recent results by Chordia et al. (2015), Pukthuanthong et al. (2019) and Jegadeesh et al. (2019). In our results, smb is significant with a t-stat of 2.82, and the momentum factor mom is significant with $t = -2.87$. However, if we follow Harvey et al. (2016)'s suggestion to use a t-stat hurdle of 3, no risk factor is significant. Furthermore, according to the

[21] We thereby follow e.g. Koijen and Yogo (2019), who match 13(f) filings data and the CRSP data set via the CUSIP identifier.

factor protocol proposed by Pukthuanthong et al. (2019), a necessary condition for qualifying as risk factor is the fact that a proposed factor carries a positive risk premium. According to our results only *smb* carries a significantly positive risk premium.

Table 2: In this table we report the empirical estimation results for the Fama and MacBeth (1973) regressions according to Equation (5): $r_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{h,t}\hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{j,t}\phi_{i,j,t-1} + \epsilon_{i,t} \; \forall \; t = 1, ..., T$. The dependent variable is the monthly excess market return of individual stocks in percent. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. As described in Section (B.2) we apply the EIV correction from Pukthuanthong et al. (2019). In this approach stocks are sort into deciles of market-capitalistion and factor betas. The individual factor beta is then replaced by the average beta of the corresponding size-beta category. The reported estimates in the table are the time-series averages of the monthly slope-coefficient estimates. The sample period is January 1975 to May 2016. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%;

	(1) **Full Model**	(2) **excl. IL**	(3) **CAPM**	(4) **FF3**	(5) **FF5**	(6) **FFC6**	(7) **IL**	(8) **PB**	(9) **MO**	(10) **VO**	(11) **BT**	(12) **SZ**	(13) **FD**	(14) **IV**	(15) **OP**
$const$	0.629	0.672	1.128	1.014	0.998	0.995	0.853	0.955	0.976	1.018	0.998	0.762	0.989	0.982	0.965
	(3.15) ***	(3.44) ***	(5.95) ***	(6.05) ***	(5.96) ***	(6.00) ***	(4.99) ***	(5.79) ***	(6.18) ***	(5.73) ***	(5.51) ***	(4.47) ***	(5.42) ***	(6.05) ***	(6.02) ***
mkt	0.014	-0.009	-0.113	-0.156	-0.130	-0.203	-0.043	-0.145	-0.245	-0.192	-0.213	0.001	-0.164	-0.185	-0.220
	(0.14)	(-0.09)	(-0.71)	(-1.15)	(-0.95)	(-1.49) **	(-0.348)	(-1.09)	(-1.86) *	(-1.51)	(-1.59)	(0.012)	(-1.29)	(-1.39)	(-1.68) **
smb	0.142	0.142		0.211	0.237	0.186	0.152	0.183	0.172	0.172	0.187	0.135	0.165	0.171	0.188
	(2.82) ***	(2.80) ***		(2.38) **	(2.90) ***	(2.40) **	(2.44) **	(2.43) **	(2.42) **	(2.54) ***	(2.45) **	(2.27) **	(2.58) ***	(2.27) **	(2.56) ***
hml	-0.025	-0.025		0.068	0.094	0.004	0.015	-0.021	-0.067	0.001	0.033	0.016	0.010	0.017	0.014
	(-0.40)	(-0.39)		(0.88)	(1.24)	(0.060)	(0.20)	(-0.29)	(-0.94)	(0.010)	(0.45)	(-0.21)	(0.14)	(0.22)	(0.19)
cma	0.032	0.033			0.002	0.007	0.000	0.019	-0.003	0.016	0.016	0.000	0.023	-0.000	0.10
	(0.69)	(0.70)			(0.03)	(0.14)	(0.02)	(0.39)	(-0.08	(0.33)	(0.33)	(0.02)	(0.46)	(-0.00)	(0.20)
rmw	0.024	0.024			0.061	0.056	0.058	0.043	0.036	0.040	0.065	0.053	0.066	0.066	0.038
	(0.56)	(0.56)			(1.12)	(1.07)	(1.20)	(0.84)	(0.73)	(0.80)	(1.35)	(0.107)	(1.37)	(1.31)	(0.78)
mom	-0.230	-0.232				-0.364	-0.303	-0.35	-0.317	-0.310	-0.300	-0.263	-0.301	-0.341	-0.369
	(-2.87) **	(-2.89) ***				(-3.64) ***	(-3.24) ***	(-3.53) ***	(-3.54) ***	(-3.55) ***	(-3.30) ***	(-3.02) *	(-3.45) ***	(-3.52) ***	(-3.81) ***
IL^L	-0.067						-0.038								
	(-1.07)						(-0.32)								
IL^H	0.339						0.542								
	(3.78) ***						(3.67) ***								
PB^L	0.229	0.247						0.390							
	(3.67) ***	(3.90) ***						(4.20) ***							
PB^H	-0.320	-0.327						-0.304							
	(-4.80) ***	(-4.89) ***						(-3.71) ***							

Table continued:

	(1) **Full Model**	(2) **excl. IL**	(3) **CAPM**	(4) **FF3**	(5) **FF5**	(6) **FFC6**	(7) **IL**	(8) **PB**	(9) **MO**	(10) **VO**	(11) **BT**	(12) **SZ**	(13) **FD**	(14) **IV**	(15) **OP**
MO^L	-0.192	-0.191							-0.096						
	(-1.89) **	(-1.87) **							(-0.67)						
MO^H	0.309	0.299							0.251						
	(4.12) ***	(3.97) ***							(2.71) ***						
VO^L	-0.011	-0.035								-0.111					
	(-0.16)	(-0.52)								(-0.95)					
VO^H	-0.543	-0.475								-0.055					
	(-3.60) ***	(-3.12) ***								(-0.22)					
BT^L	-0.020	-0.010									-0.017				
	(-0.41)	(-0.21)									(-0.27)				
BT^H	0.029	0.024									0.048				
	(0.49)	(0.41)									(0.67)				
SZ^L	0.455	0.610										0.650			
	(4.50) ***	(5.74) ***										(4.20) ***			
SZ^H	-0.036	-0.083										0.030			
	(-0.50)	(-1.02)										(0.26)			
FD^L	0.163	0.152											0.007		
	(3.00) ***	(2.73) ***											(0.06)		
FD^H	0.765	0.792											0.909		
	(3.71) ***	(3.83) ***											(2.90) ***		
IV^L	0.267	0.268												0.304	
	(2.44) **	(2.46) ***												(1.87) *	
IV^H	-0.135	-0.144												-0.179	
	(-1.99) **	(-2.11) **												(-2.21) **	
OP^L	-0.251	-0.252													-0.124
	(-3.50) ***	(-3.51) ***													(-0.97)
OP^H	0.419	0.422													0.343
	(8.67) ***	(8.66) ***													(5.82) ***

In stark contrast to the absence of significance for risk factors, 13 out of 18 characteristics are significant at conventional levels, and ten even pass the stricter hurdle of having t-stats beyond three. In particular, we find that the stock characteristics high-illiquidity IL^H, low price-to-book value PB^L (i.e. value stocks), positive momentum MO^H, small market-capitalization SZ^L, distress FD, low asset growth IV^L, and high operating-profitability OP^H carry significant positive characteristic premiums. In contrast low-illiquidity IL^L, high price-to-book value PB^H (i.e. growth stocks), negative momentum MO^L, high return-volatility VO^H, high asset growth IV^H and low operating-profitability OP^L earn significant negative characteristic premiums. Except for the positive risk premium related to financial distress risk (see e.g. Campbell et al. (2008)), our findings are well in line with findings of the earlier empirical asset pricing literature. Recall from above, that we prefer to use univariate betas in the second stage estimation. However, when using multivariate betas, our results are qualitatively identical, albeit even slightly stronger as neither the momentum factor *mom* nor the size factor *smb* turn out strongly significant. We report slope coefficient estimates from using multivariate betas (only for the full model) as column (4) in Table 16 in Appendix C.

Results in Table 2 are based on the mean beta EIV correction by Pukthuanthong et al. (2019), which is a simple and pragmatic approach. We check our results by implementing the more sophisticated IV approach by Jegadeesh et al. (2019) and report results in Table 16 in appendix C. We estimate univariate factor betas from daily returns of disjunct sets of odd and even calendar months. We estimate the factor betas using daily returns over a rolling window of 720-trading days, which corresponds to the 36-month rolling window used by Jegadeesh et al. (2019).[22] Overall, we find almost identical results as with the mean beta approach. The only notable difference is that with the IV approach, also the *smb* factor turns out insignificant, thereby even strengthening our findings.[23] The reason for not relying on the IV approach as our baseline specification is the fact that it tends to produce an increasing number of extreme outliers in the second stage regression as the number of factors is augmented.[24] With the caveat of outliers in the

[22]Note already, that in a later section (i.e. section D.1), we will estimate the Fama and MacBeth (1973) regression with betas from monthly data instead of a daily frequency. Results remain qualitatively unchanged.

[23]One could speculate that the fact that the mean beta approach produces a significant *smb* factor could be due to deriving the mean beta from sorting on size. However, we were rerunning the analysis with mean betas from sorting on Price-Book ratios as well as on past price performance and found almost identical results (results not reported, available on request).

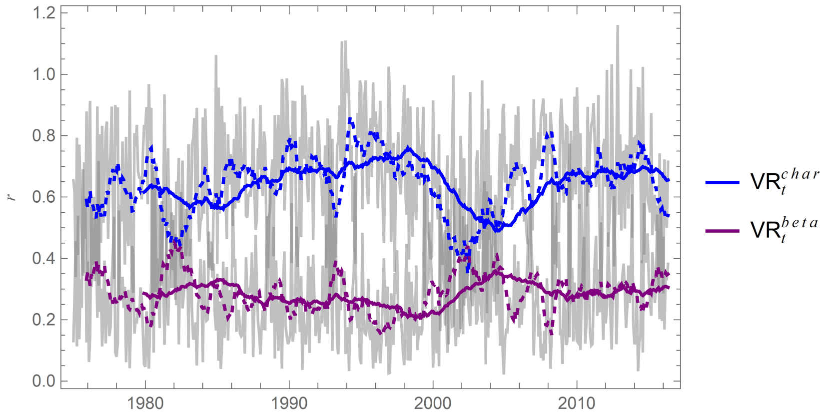
[24]Jegadeesh et al. (2019) note in their footnote 17, that they drop monthly risk premia estimates if they are larger than six standard deviation from the average raw return, which in their analysis

second stage, the IV approach largely confirms our results. Finally, in line with Jegadeesh et al. (2019) and Chordia et al. (2015), we contrast our findings with results from OLS estimation. We report results in Table 16 in Appendix C, which show that the EIV correction is not crucial and that even OLS results are close to our overall finding. We find that *smb* looses significance, so only *mom* (with $t = -2.44$) remains a relevant risk factor, while all the results from mean beta or IV correction are virtually unchanged in the simple OLS estimation, with 14 out of 18 characteristics as significant out of which 9 exhibit t-stats beyond 3.

Determining the significance of risk premia estimates for factors versus characteristics is standard in the literature to establish the importance of the corresponding explanatory variable. However, an arguably even cleaner test of the relative significance of either candidates is to perform the expected return decomposition according to Chordia et al. (2015), which quantifies the relative contribution of the factor loadings (i.e. betas) and characteristics to the expected (i.e. explained) stock return. We therefore determine the proportional cross-sectional variance of factors and characteristics ($\mathbb{VR}_t^{beta}$ and $\mathbb{VR}_t^{char}$) according to equation (7) and plot their time-variation in Figure 3 for the monthly values, as well as for the 1-year (dashed) and 5-year (solid) moving average over the sample period 01/1975–06/2016. Over the entire sample, the average fraction of expected return variation that is explained by risk factors is 28.56%, while the corresponding fraction for characteristics is 64.08%. Thus, on average characteristics explain the major share of cross-sectional variation in expected returns. The graph shows substantial time variation on a monthly frequency, but averaging over a rolling 1-year and 5-year window reveals that this pattern appears to be robust over time. On the basis of the 1-year MA, only two periods in the early 80s and 2000s display a close to equal decomposition, while the 5-year MA shows only little time variation and a consistently larger explanatory fraction of characteristics. This finding is important in particular against the background of contributions which show that there is vanishing evidence for anomalies after 2003. Green et al. (2017) find that in particular for non-microcap stocks, evidence for characteristics-based predictability drops sharply around 2003. When restricting our analysis to post-2003 data,

is the case for approximately 3% of the observations. In our specification with six risk factors (and by using the same 6-σ criterion), we find that up to 40% of monthly factor premia estimates would have to be excluded. (To be precise: The fraction of observations that are beyond six standard deviations are 22.7% for *mk*, 19.9% for *smb*, 39.6% for *hml*, 42.8% for *cma*, 27.1% for *rmw*, and 20.7% for *mom*.) It is hard to believe that monthly risk premia can jump by a factor of 6, while at the same time it also appears inappropriate to drop such a substantial subset of observations. Therefore, we consider the IV approach with a larger number of risk factors as impractical.

Figure 3: In this figure we plot the time series of the cross-sectional variation, which is explained by factor-risk exposure (i.e. betas) and characteristics over time. As shown by Chordia et al. (2015) due to the correlation between factor betas and characteristics the two ratios do not necessarily have to add up to 1. The mean over the full sample period for the explanatory power of characteristics is 64.08%, while the corresponding mean is 28.56% for the factor betas. The dotted and the solid line respectively show the moving average of the factor/beta explained component over a 12 month and 60 month rolling window.



we can confirm the decline in statistical significance of point estimates of return contribution from characteristics.However, when looking at the variance decomposition, we indeed see a drop in characteristics-induced variation in the interval 1998-2002, but which is only temporary. For the post-2003 sample (i.e. from 01/2003–06/2016), we find an average variance ratio of 65.50% for characteristics and 29.70% for factors which is almost identical to the entire sample. Thus, we find no evidence of a declining importance of characteristics from the perspective of the variance decomposition.

C.2. Revealed Preferences From Institutional Portfolio Tilts – Factors or Characteristics

In this section, we shift our focus on institutional demand as dependent variable. We run the panel regression model of Equation (8), which attempts to identify if tastes for characteristics or aversion to risk factors are first-order determinants of institutional portfolio holdings. The dependent variable in the regression model is the relative market tilt of institutional portfolio managers. Our results, which are summarized in Table 3, show that institutional managers' tilts for individual stocks are strongly related to stocks characteristics. These results emerge robustly from the univariate regressions (Columns 4 to 9), as well as in our regression model that includes all characteristics simultaneously.

Table 3: In this table we report the empirical estimation results for the institutional investors portfolio tilt panel regressions according to Equation (8): $T_i = \alpha_0 + \sum_{f=1}^{F} \theta_{ft} \hat{\beta}_{if} + \sum_{j=1}^{J} \theta_{jt} \phi_{ij} + c_i + p_t + \epsilon_{it}$. The dependent variable is the demand tilt, winsorized at the 97 % quantile. The explanatory variables include the univariate factor betas with respect to the market, SMB, HML, CMA, RMW and momentum factor, estimated over a rolling window of 36 months and winsorized at the 1 % and 99 % quantile. Additionally, we include the characteristic labels/indicator dummies, which identify whether a particular stock belongs to the top/bottom quantile according to the corresponding characteristic sort (see Section B.3). The regression coefficients can be interpreted as the change in percentage points in the average stock tilt, when multiplied with 100. The regression includes firm fixed and time fixed effects. Standard errors are estimated according to Driscoll and Kraay (1998). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%;

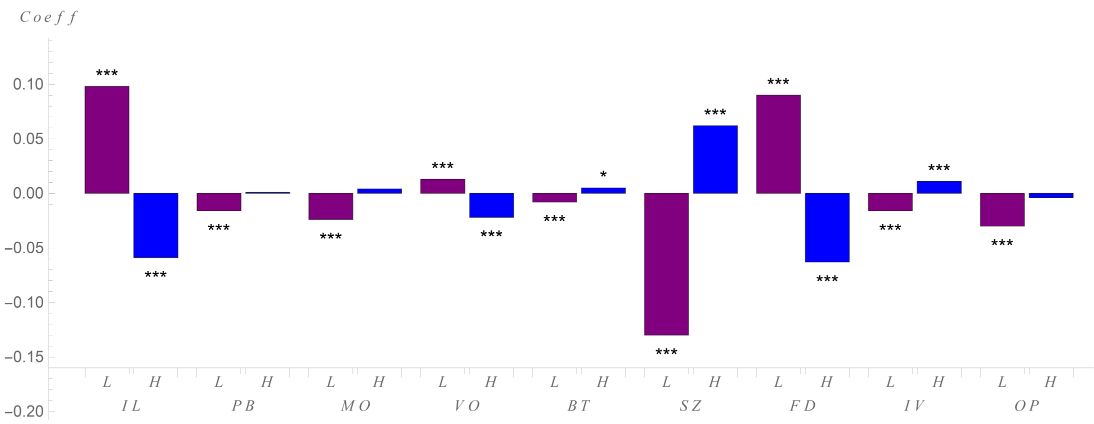
	(1) **Full Model**	(2) **excl. IL**	(3) **Char**	(4) **IL**	(5) **PB**	(6) **MO**	(7) **VO**	(8) **BT**	(9) **SZ**	(10) **FD**	(11) **IV**	(12) **OP**
const	0.557	0.558	0.558	0.569	0.597	0.610	0.606	0.603	0.567	0.555	0.601	0.608
	(67.46) ***	(70.60) ***	(68.78) ***	(108.13) ***	(162.40) ***	(147.37) ***	(156.51) ***	(154.10) ***	(84.57) ***	(100.63) ***	(160.84) ***	(151.72) ***
mkt	0.002	0.003										
	(1.24)	(1.58)										
smb	0.002	0.003										
	(2.00) **	(2.16) **										
hml	0.001	0.001										
	(1.09)	(0.93)										
cma	0.001	0.001										
	(1.26)	(1.12)										
rmw	0.006	0.006										
	(8.63) ***	(8.33) ***										
mom	0.004	0.004										
	(2.86) ***	(3.21) ***										
IL^L	0.098		0.097	0.176								
	(16.82) ***		(16.95) ***	(20.19) ***								
IL^H	-0.059		-0.058	-0.145								
	(-11.66) ***		(-11.43) ***	(-16.25) ***								
PB^L	-0.016	-0.021	-0.016		-0.078							
	(-4.38) ***	(-5.75) ***	(-4.43) ***		(-19.66) ***							
PB^H	0.001	0.005	0.001		0.052							
	(0.21)	(0.95)	(0.23)		(8.35) ***							
MO^L	-0.024	-0.022	-0.024			-0.041						
	(-12.35) ***	(-12.04) ***	(-11.87) ***			(-17.16) ***						

Table continued:

	(1) **Full Model**	(2) **excl. IL**	(3) **Char**	(4) **IL**	(5) **PB**	(6) **MO**	(7) **VO**	(8) **BT**	(9) **SZ**	(10) **FD**	(11) **IV**	(12) **OP**
MO^H	0.004	0.005	0.004			0.008						
	(1.25)	(1.63)	(1.20)			(2.17) ***						
VO^L	0.013	0.014	0.013				0.036					
	(2.83) ***	(3.40) ***	(2.89) ***				(8.01) ***					
VO^H	-0.022	-0.030	-0.024				-0.093					
	(-5.28) ***	(-6.86) ***	(-5.32) ***				(-19.87) ***					
BT^L	-0.008	-0.007	-0.006					-0.010				
	(-3.98) ***	(-3.72) ***	(-2.37) ***					(-3.19) ***				
BT^H	0.005	0.006	0.002					-0.003				
	(1.76) *	(2.08) **	(0.55)					(-0.78)				
SZ^L	-0.130	-0.156	-0.130						-0.192			
	(-22.60) ***	(-19.01) ***	(-22.35) ***						(-20.44) ***			
SZ^H	0.062	0.107	0.063						0.165			
	(9.72) ***	(14.17) ***	(9.78) ***						(17.18) ***			
FD^L	0.090	0.101	0.091							0.150		
	(13.26) ***	(14.69) ***	(13.56) ***							(19.34) ***		
FD^H	-0.063	-0.071	-0.064							-0.157		
	(-10.73) ***	(-11.34) ***	(-10.67) ***							(-18.05) ***		
IV^L	-0.016	-0.018	-0.017								-0.058	
	(-9.44) ***	(-10.49) ***	(-9.61) ***								(-21.86) ***	
IV^H	0.011	0.015	0.011								0.049	
	(5.67) ***	(7.32) ***	(5.92) ***								(14.33) ***	
OP^L	-0.030	-0.033	-0.030									-0.070
	(-11.61) ***	(-12.60) ***	(-11.61) ***									(-20.29) ***
OP^H	-0.004	-0.001	-0.003									0.029
	(-1.61)	(-0.57)	(-1.35)									(9.40) ***
Within R^2	17.54 %	16.51 %	17.38 %	12.20 %	7.92 %	6.69 %	6.97 %	6.03 %	13.33 %	11.01 %	6.98 %	7.02 %
p-value Hausmann	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	389,971	389,971	389,971	389,971	389,971	389,971	389,971	389,971	389,971	389,971	389,971	389,971

Reflecting the results from the previous section, we find that from the set of risk factors, *smb* and *mom* are again significant at conventional levels. Additionally, *rmw* turns out as strongly significant for portfolio tilts. However, when imposing the t-stat hurdle of 3, only *rmw* survives. Thus, only one out of the six common risk factors can robustly explain portfolio tilts. Results for the set of characteristics are markedly different. 13 out of 18 characteristics comfortably exceed a t-value of 3, and 15 out of 18 are significant at conventional levels. Results are qualitatively identical, albeit even slightly stronger if we exclude the illiquidity characteristics due to the results from the variance inflation factor (see second column). When estimating univariate relationships, we find substantially stronger coefficients, reflecting the fact that individual characteristics can hardly be isolated. To facilitate the qualitative interpretation of the regression results, we illustrate the estimated coefficients (from column 1 in Table 3) in Figure 4 as barcharts which allows for the immediate visual assessment in which characteristics institutional portfolios are over- (positive bars) or underweighted (negative bars). Note that the regression coefficients (when multiplied with 100) on the characteristics dummies can be interpreted as the percentage change in the average stock tilt in case that a particular stock carries the corresponding label.

Figure 4: In this figure, we show the estimated coefficients from column 1 in Table 3 as barchart plot. Positive (negative) bars indicate over- (under-)weighted tilts. Regression coefficients (when multiplied with 100) on the characteristics dummies can be interpreted as the percentage change in the average stock tilt in case that a particular stock carries the corresponding label. Statistical significance is indicated as stars above/below bars.



Across the 18 high and low characteristic labels, the strongest coefficients (and t-stats) can be found for illiquidity, size and distress. Investors are strongly tilted towards liquid, large and financially-healthy stocks. IL^L, SZ^H, and FD^L have coefficients of 0.098, 0.062, and 0.09 respectively, indicating that stock tilts are roughly up to 10% larger when stocks carry the corresponding characteristic. In contrast, significant and substantial negative tilts (or underweighting) can be found for stocks that are illiquid (IL^H, with coefficient -0.059), are prone to financial distress (FD^H, -0.063), and display low operating profitability (OP^L, -0.03). Further statistically significant negative tilts are obtained for stocks with low valuation ratios (PB^L, -0.016), low price continuation (MOL, -0.024), high volatility (VO, -0.022), and low asset growth (IV^L, -0.016) although coefficients are economically smaller. Taken together, the results reveal the pervasive pattern that institutional investors tilt their portfolios towards stock characteristics which can be considered as good aspects and tilt away from allegedly weak stock characteristics.

In contrast to the return regression, we employ panel regressions which allows the inclusion of fixed effects. However, as robustness check, we estimate the tilt regression also within the Fama and MacBeth (1973) approach, i.e. according to the model in (9) to make it directly comparable to the return regression. Due to the data availability of 13(f) filings, the tilt regression is conducted on a quarterly frequency. Therefore, we re-run the above return regression also on a quarterly frequency (and with the mean beta EIV correction). Results are summarized in Table 17 in Appendix D. Results for quarterly *return* regressions (as reported in row (3) of Table 17) are qualitatively identical, albeit at lower significance levels which is likely due to the lower number of observations. Within the Fama and MacBeth (1973) *tilt* regressions, we find overall a substantially higher significance level for all variables (as compared to panel regression). It implies that we do find that characteristics as well as risk factors turn out to be significant with large t-values. The size characteristic e.g. carries a t-stat of -86.2, the market factor mkt is significant with $t = 16.19$. However, these high significance levels can be attributed to the lack of firm fixed effects. Once we remove firm fixed effects from the panel regression (see row (2) in Table 17), almost identical results to the Fama and MacBeth (1973) approach are obtained, which is plausible, as the Fama and MacBeth (1973) method is in theory close to a panel regression without firm fixed effects.[25] We interpret the results as strong support to rely on results from panel regression *with* fixed effects as reported in our main Table 3.

Our results are related to and well in line with findings of Edelen et al. (2016), who show that institutional investors seem to trade contrary to anomaly prescriptions in the sense of rather buying overvalued stocks that are part of the short leg.

[25]See e.g. Cochrane (2005).

Similar to Edelen et al. (2016), we observe that our demand tilts turn out to be unrelated to the investment suggestions by the academic anomaly literature, such as the well-established finding that high-beta stocks and high asset growth stocks earn below average returns. Further empirical results suggest that sophisticated investors should exploit the size and value premium by investing into small-cap and value stocks, as they historically provided above average returns. As argued by Edelen et al. (2016) high price-to-book value, high asset growth and large market capitalization create an appearance of a good investment, despite that these stocks empirically provide below average returns. Our findings strongly suggest that institutional investors pay attention to stock characteristics, regardless of the empirically observed return patterns associated to these characteristics, and that they use stock characteristics as tangible identifiers to select stocks and construct their portfolios. While a full analysis of the rationale behind this observed behavior is beyond the scope of this paper, a potential reason may be found in information processing arguments and agency relationships. Stock characteristics are salient and readily available pieces of information which can be easily used to classify stocks into strong and weak investments. Such an uncontroversial classification can in turn be used as legitimation against investors or superiors to justify investment decisions.

Our empirical approach is distinct to Edelen et al. (2016) as it allows to address the marginal tilt effects within a multivariate regression setting.[26] Thus, in our full regression model we additionally control for the past factor risk exposure of individual stocks. Therefore we compute univariate factor betas for *mkt*, *smb*, *hml*, *cma*, *rmw* and the momentum factor *mom* over a rolling window of 36 months, winsorized at the 1% and 99% quantile to remove outliers, and include these betas as additional regressors. Thereby we check whether the relation of demand tilts to characteristics is orthogonal to controlling for past factor risk exposure. This regression framework allows to address whether investors have tastes for known characteristics besides aversion with respect to risk factors. We argue that institutional investors reveal their preferences for factor risk exposure and stock characteristics through their portfolio holdings. With this regression we also address the challenge raised by Pastor et al. (2020), who state that from an economic perspective it is necessary to address whether any observed characteristic excess return relation is related to fundamental tastes or (latent/unobserved) factor risks.

[26] Applying a *multi*variate regression approach with respect to demand tilts addresses the general call for multivariate methods in empirical asset pricing, as e.g. emphasized by Cochrane (2005) and Green et al. (2017).

The results from our full regression model show that institutional portfolio holdings are significantly related to stock characteristics, despite controlling for factor risk exposure. The findings are closely comparable to the univariate regressions and the regression without the factor beta controls. This finding can be interpreted in the sense that institutional investors appear to have tastes/preferences for certain stock characteristics, which is orthogonal to the stocks' factor risk exposure.

C.3. Linking Institutional Tilts and Expected Returns

The previous two sections have shown that excess returns as well as institutional investors' portfolio tilts are significantly related to characteristics, which offers the tentative interpretation that pervasive patterns in investors' demand drives the cross-section of returns. We analyse this alleged link more rigorously in this section by regressing excess returns on portfolio tilts via the panel regression model from Equation (10). Results are reported in Table 4.

Table 4: In this table we report the results according to the panel regression model from Equation (10): $r_{it}^{ex} = \alpha_0 + \beta \times T_{i,t-1} + c_i + p_t + \epsilon_{i,t}$. The dependent variable is the quarterly excess stock return. The explanatory variable is the institutional tilt of the prior quarter computed according to 3. Standard errors are estimated according to Driscoll and Kraay (1998). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	Dependent variable r_{it}^{ex}	
Constant	0.095	
	(8.27)	***
$T_{i,t-1}$	-0.101	
	(-6.37)	***
Within-R^2	13.18 %	
Observations	381,655	

The coefficient on lagged tilts turns out negative with a t-stat of more than six. Thus, it is statistically highly significant and the economic magnitude is large. A stock in which institutional investors are overweighted by 1%-point is associated with a next quarters' return that is lower by 0.101%-points. To assess the economic magnitude, recall from the tilt regression (i.e. equation (8), Table 3) that e.g. the high liquidity characteristic (IL^L) carries a coefficient of 0.098 (in the multivariate model), which means that the tilt in very liquid stocks is higher by 9.8%-points. Thus, it translates to a subsequent return which is lower by 1.77%-points per quarter, or by 7.11% annually, which is economically significant.

Therefore we conclude that institutional demand tilts have a strong impact on asset prices and consecutive stock returns. The negative sign indicates that institutions are on average tilted towards the wrong, i.e. underperforming side of ex ante expected stock returns. This finding is actually consistent to the results in Edelen et al. (2016) who document that over a one year horizon, stock returns are negatively related to institutional investor demand, which they interpret as evidence against the sophisticated investor hypothesis.[27] Our results support the interpretation that institutional investors are tilting their portfolio to a number of favorably perceived stock characteristics and that they are willing to accept a lower than average return for holding those stocks that fit their preferences.
Although institutional investors are obviously an important and sizable group of investors, our analysis cannot take into account demand tilts of retail investors. Still, theoretical arguments as discussed in e.g. Pastor et al. (2020) show, that even if only a subgroup of investors has preferences for stock characteristics, then this imbalance is sufficient to induce the manifestation of a relationship between average returns and stock characteristics. Therefore, independent of the underlying rationale, our results show that institutional demand tilts turn out to be strong enough to distort asset prices, which leads support to our proposed transmission channel.

We take our analysis one step further by decomposing returns as well as tilts into characteristic and factor-risk induced components. The previous result has documented that the overall portfolio tilt is strongly associated to price changes. If it is true that characteristics are predominantly important for portfolio tilts and that this preference-induced demand translates to return patterns, then we should expect to see this relation in particular from the characteristics-induced component of the portfolio tilts. Thus, in order to decompose institutional stock tilts into characteristic and factor-risk related components, we first run quarterly cross-sectional Fama and MacBeth (1973) regressions of individual stock tilts on

[27]Note that Edelen et al. (2016) proxy institutional demand by the changes in the aggregate number of institutional investors. We consider our portfolio tilt measure to be a superior proxy for demand as it accounts for the position size of the institutional investors' portfolio. However, we run a robustness check, where we regress quarterly returns on the lagged change in the number of institutional stock owners. Comparable to Edelen et al. (2016) we find that there is a horizon effect of changes in the number of institutional stock owners on stock returns, where the effect becomes significantly negative when expanding the lag between changes in the number of institutional stock owners and future returns from one quarter to one year. Indeed this horizon effect is not observable when using the tilt variable. It therefore appears that this documented horizon effect of Edelen et al. (2016) is special to using the number of institutional stock owners as an explanatory variable. The results for this robustness check can be provided upon request.

factor betas and stock/firm characteristics as in (9). Similarly to the variance decomposition, we then define the respective tilt part that is explained by factors and characteristics for each individual stock i in each quarter t as:

$$\hat{T}_{i,t}^{beta} = \sum_{f=1}^{F} \hat{\theta}_{f,t}\hat{\beta}_{i,f,t-1}, \qquad \hat{T}_{i,t}^{char} = \sum_{j=1}^{J} \hat{\theta}_{j,t}\phi_{i,j,t-1} \quad \forall\ t = 1, ..., T. \tag{12}$$

For the lagged beta estimate of stock i, we use the EIV correction according to the mean beta approach of Pukthuanthong et al. (2019) as well as the IV estimator from Jegadeesh et al. (2019). Betas from the IV approach are estimated from disjunct sets of odd and even months over a 720-day rolling estimation window. We include the factor betas for the five Fama and French (2015) factors augmented with the Carhart (1997) momentum factor.[28] Similar to the Fama and MacBeth (1973) return regressions, we compute time-series averages of the slope coefficients and report them in Column (1) of Table 17 of Appendix D.
Since coefficients are determined within an EIV-corrected regression, they will not suffer from a systematic bias.[29] Therefore we use fitted tilt parts $\hat{T}_{i,t}$ as explanatory variables in the next step of a predictive panel regression of future quarterly stock returns on the tilt parts explained by factors and characteristics.[30] This panel regression includes firm fixed effects c_i and time-fixed effects p_t:[31]

$$r_{i,t}^{ex} = \alpha + \gamma_{beta}\,\hat{T}_{i,t-1}^{beta} + \gamma_{char}\,\hat{T}_{i,t-1}^{char} + c_i + p_t + \epsilon_{i,t} \tag{13}$$

We combine the decomposition of tilts as in (12) with the return decomposition

[28] For the cross-sectional regression of quarter t we use the factor betas and stock/firm characteristics that are observable at the end of the preceding quarter. This means that for example for the fourth quarter we use the factor betas and characteristics that are observable at the end of September.

[29] To check for potential biases, we conducted a robustness check where we run a Monte Carlo Simulation with randomly drawn tilts from the variance-covariance matrix of the estimates from Equation (9). The robustness check confirms our results and indicates that a fitted tilts do not introduce an EIV problem.

[30] We only include stocks in the sample that have a full trading record within quarter t. Therefore stocks that delist e.g. in February are not included in the first quarter, but are included in the fourth quarter of the preceding year.

[31] We include firm fixed effects to control for unobserved heterogeneity. The Hausman (1978) -type specification test proposed by Hoechle et al. (2020) strongly rejects the null-hypothesis that the random effects assumption holds. Therefore we decide to include firm-fixed effects. We include time-fixed effects in the regression as they allow to control for market-wide return fluctuations that occur within a specific quarter (i.e. movements of the market risk factor). By controlling for time-fixed effects we can estimate the marginal effects of institutional factor/characteristic tilts on future excess returns. Neglecting time-fixed effects would lead to biased estimates and standard errors.

already introduced in Equation (6),[32] which we use in running a panel regression of the factor/characteristic return part on the factor tilt component and the characteristic tilt component. Again we include firm fixed effects c_i and time-fixed effects p_t:

$$\hat{r}_{it}^{ex,beta} = \alpha + \gamma_{beta}\, \hat{T}_{i,t-1}^{beta} + \gamma_{char}\, \hat{T}_{i,t-1}^{char} + c_i + p_t + \epsilon_{i,t} \tag{14}$$

$$\hat{r}_{it}^{ex,char} = \alpha + \gamma_{beta}\, \hat{T}_{i,t-1}^{beta} + \gamma_{char}\, \hat{T}_{i,t-1}^{char} + c_i + p_t + \epsilon_{i,t} \tag{15}$$

We report results in Table 5.

Table 5: In Column 1 we report the results of the panel regression according to Equation (13). In Columns 2 and 3 we report the estimation results of the panel regression according to Equation (14). Standard errors are estimated according to Driscoll and Kraay (1998). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	(1)		(2)		(3)	
			Dependent variable			
	r_{it}^{ex}		$\hat{r}_{it}^{ex,beta}$		$\hat{r}_{it}^{ex,char}$	
constant	0.044		0.014		-0.009	
	(3.95)	***	(1.73)	*	(-1.96)	*
$\hat{T}_{i,t-1}^{beta}$	0.013		0.087		-0.002	
	(0.15)		(1.41)		(-0.09)	
$\hat{T}_{i,t-1}^{char}$	-0.165		-0.009		-0.038	
	(-7.47)	***	(-1.43)		(-2.97)	***
Time-Fixed Effects	Yes		Yes		Yes	
Firm-Fixed Effects	Yes		Yes		Yes	
Within-R^2	17,06%		55.46%		13.17 %	
Observations	389,992		389,992		389,992	

Column (1) of Table (5) reports the estimation result for the regression from (quarterly) excess returns on the factor- and characteristic-induced tilt component. We find that only the characteristic-induced part is significant. In comparison to the aggregate regression, the magnitude of the coefficient as well as the t-stat are larger which leads strong support to our conjecture that characteristics-induced demand is pricing-relevant but not demand due to risk factor aversion. Column

[32] Note, that contrary to the return decomposition in Section C.1, we use *quarterly* frequency here since the tilt measure is only available on a quarterly basis. To check if return regression results change with the lower frequency, we report the time-series averages of the estimated quarterly coefficients in Column (3) of Table 17 in Appendix D. Overall the results are comparable to the monthly regression results presented in Table 2.

(2) reports results when we regress only the risk factor-related return component $\hat{r}_{it}^{ex,beta}$ on the tilt decomposition. Neither tilt component is significant. Vice versa, when regressing the characteristics-related return part $\hat{r}_{it}^{ex,char}$ on the tilt decomposition (Column 3), we find the same pattern, namely no significance for the risk factor-induced tilt, while the characteristics-induced tilt is significant again. As robustness test, we check if the EIV correction influences results. We run the same analysis by implementing the IV EIV correction of Jegadeesh et al. (2019) and report results in Table 18 in Appendix E. We find almost identical results. Taken together, these results appear as strong support for the interpretation that the preferences with respect to characteristics are revealed in portfolio tilts of institutional investors and that these preferences are pricing-relevant in the cross-section of equity returns.

D. Robustness Checks

D.1. Latent Factor Risks

An important concern is the possibility that characteristics turn out strongly significant not because they are generic explanatory variables but because they are proxies for some underlying latent risk factor. Although we control for six well-established risk factors, it is still conceivable that yet unidentified factors are driving results. Kelly et al. (2019) is a recent contributions which argues that a large number of characteristics should not be interpreted as anomalies but rather as dynamic latent factor loadings. Kelly et al. (2019) derive latent risk factors from the cross-section of returns through an augmented variant of principal components analysis. They show that it is sufficient to consider a low-dimension model with only four latent factors, which they use as their baseline model. It is thus a natural robustness check to test if our results survive the inclusion of the latent risk factors as identified by Kelly et al. (2019). We obtain the latent risk factor data from the internet appendix of Kelly et al. (2019) for the period 1975 to July 2014. We include them in our Fama and MacBeth (1973) regression framework as substitutes for the six risk factors.

A direct comparison to our previous results is complicated by the fact that the latent factor data has a monthly frequency and ends in 2014. To overcome this problem, we first re-run the first stage regression by computing betas from *monthly* observations for the five Fama and French (2015) and the Carhart (1997) momentum factor over a rolling window of 48 months. In line with the preceding section, we use the mean beta approach of Pukthuanthong et al. (2019) and the IV method by Jegadeesh et al. (2019) to accommodate the EIV problem. We report results from the mean beta correction in Table 6 (and include the results from the IV method as Table 19 in Appendix F). In both tables, column (1) reports slope coefficients form our base specification as in Section C.2 (sample period 01/1975–06/2016) but with monthly betas. Column (2) repeats the same analysis as in column (1) but only for the restricted sample period up to 2014 to be directly comparable to the specification in column (3) which replaces the six risk factors by the four latent Kelly et al. (2019) factors $F1$–$F4$.

Table 6: In this table we report the empirical estimation results for the Fama and MacBeth (1973) regressions according to Equation (5): $r_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{h,t} \hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{j,t} \phi_{i,j,t-1} + \epsilon_{i,t} \ \forall \ t = 1, ..., T$. The dependent variable is the monthly excess market return of individual stocks in percent. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. The factor betas are estimated with respect to the five Fama and French (2015) factors augmented with the Carhart (1997) momentum factor. Additionally, we also use the four latent risk factors provided by Kelly et al. (2019) (F1, F2, F3, F4). As described in Section (B.2) we apply the EIV correction from Pukthuanthong et al. (2019). In this approach stocks are sort into deciles of market-capitalistion and factor betas. The individual factor beta is then replaced by the average beta of the corresponding size-beta category. In contrast to the results of Table **??** the factor betas are estimated from monthly (not daily) returns over a 48-month rolling window, as the latent factors from Kelly et al. (2019) are only available at quarterly frequency. The reported estimates in the table are the time-series averages of the monthly slope-coefficient estimates. The sample period is January 1975 to July 2014. For comparison reasons column *FF Six Factor 2016* provides results for the period January 1975 to May 2016. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%;

	Dependent variable r_{it}^{ex}						
	(1) **6-Factor-2016**		(2) **6-Factor-2014**			(3) **Latent-4-Factor**	
const	0.837		0.823		*const*	0.812	
	(4.01)	***	(3.83)	***		(3.72)	***
mkt	-0.096		0.051		$F1$	0.378	
	(-1.27)		(-0.66)			(2.50)	**
smb	0.068		0.069		$F2$	0.073	
	(1.47)		(1.44)			(0.477)	
hml	-0.007		0.011		$F3$	0.022	
	(-0.13)		(0.203)			(0.39)	
cma	0.017		0.080		$F4$	0.013	
	(0.53)		(0.24)			(0.27)	
rmw	0.006		0.017				
	(0.17)		(0.45)				
mom	-0.134		-0.129				
	(-2.11)	**	(-1.95)	*			
IL^L	-0.064		-0.063		IL^L	-0.070	
	(-0.95)		(-0.92)			(-1.04)	
IL^H	0.246		0.242		IL^H	0.238	
	(2.65)	***	(2.53)	**		(2.47)	**
PB^L	0.181		0.195		PB^L	0.187	

Table 6 continued:

	(1)		(2)			(3)	
	(2.88)	***	(3.07)	***		(2.92)	***
PB^H	-0.242		-0.262		PB^H	-0.259	
	(-3.57)	***	(-3.78)	***		(-3.74)	***
MO^L	-0.115		-0.001		MO^L	-0.096	
	(-1.07)		(-0.92)			(-0.89)	
MO^H	0.238		0.256		MO^H	0.242	
	(3.04)	***	(3.14)	***		(3.03)	***
VO^L	-0.017		-0.051		VO^L	-0.044	
	(-0.23)		(-0.67)			(-0.57)	
VO^H	-0.333		-0.322		VO^H	-0.335	
	(-1.92)	*	(-1.84)	*		(-1.95)	*
BT^L	-0.058		-0.049		BT^L	-0.046	
	(-1.06)		(-0.87)			(-0.83)	
BT^H	-0.078		-0.079		BT^H	-0.085	
	(-1.03)		(-1.03)			(-1.10)	
SZ^L	0.414		0.463		SZ^L	0.463	
	(4.00)	***	(4.40)	***		(4.57)	***
SZ^H	-0.093		-0.097		SZ^H	-0.103	
	(-1.30)		(-1.31)			(-1.43)	
FD^L	0.063		0.055		FD^L	0.057	
	(1.18)		(0.99)			(1.02)	
FD^H	0.599		0.668		FD^H	0.674	
	(2.78)	***	(3.01)	***		(3.03)	***
IV^L	0.105		0.083		IV^L	0.066	
	(0.78)		(0.59)			(0.47)	
IV^H	-0.164		-0.159		IV^H	-0.164	
	(-2.13)	**	(-2.01)	**		(-2.05)	**
OP^L	-0.293		-0.244		OP^L	-0.240	
	(-3.81)	***	(-3.10)	***		(-3.15)	***
OP^H	0.363		0.385		OP^H	0.379	
	(7.77)	***	(8.14)	***		(8.02)	***
Observations	**498**		**474**			**474**	

First, by comparing results form column (1) with the previous results in Table 2, we note that changing the data frequency (from daily to monthly beta calculation) yields almost identical results among characteristics, while the size factor *smb* looses its significance. Shortening the sample period to 2014 has almost no impact on results. Second and more importantly, column (3) shows that replacing the six (predetermined) risk factors by the latent risk factors $F1$–$F4$ has almost no impact on the results. Not only do characteristics still show high significance, but it is also the same set of characteristics with the same magnitude in the slope coefficients which explain excess returns. Third, when implementing the EIV correction according to the IV approach (Table 19 in Appendix F), we again find almost identical results.[33] We interpret the results as strong evidence that char-

[33]With IV EIV correction, even *mom* is no longer significant, while *cma* and *rmw* carry weak

acteristic return premiums are not attributable to latent/unobserved factor risks. Results in column (3) further show that only $F1$, i.e. the first out of the four Kelly et al. (2019) factors carries a significant (t-stat of 2.5) and positive return premium. According to the factor identification protocol of Pukthuanthong et al. (2019), a true pervasive risk factor should be related to the covariance matrix of returns and have an associated risk premium. The latent risk factors $F2$ to $F4$ do not meet this requirement and would not be classified as true pervasive risk factors in this sense.

By construction, latent risk factors per se have no direct economic interpretation. To investigate their meaning, we report the correlation matrix between the latent risk factors of Kelly et al. (2019) and the six FFC factors in Table (7). We

Table 7: In this table we report the correlation matrix of the five Fama and French (2015) factors (*mkt*, *smb*, *hml*, *rmw*, *cma*, *mom*), the Carhart (1997) momentum factor and the four latent risk factors of Kelly et al. (2019) (F1, F2, F3, F4).

Pairwise Correlations				
	$F1$	$F2$	$F3$	$F4$
mkt	0.070	0.839	0.350	-0.071
smb	0.127	0.654	-0.271	0.145
hml	0.195	-0.409	0.317	0.453
cma	0.056	-0.413	0.057	0.409
rmw	-0.265	-0.477	0.373	0.160
mom	-0.692	0.028	-0.317	0.255

observe that the first latent risk factor has a strong negative correlation with the momentum factor and therefore can be interpreted as an inverse variant of the momentum factor. Hence, the finding of a positive risk premium documented for the first latent risk factor $F1$ is consistent with the negative risk premium that we find for the momentum factor *mom*. The second latent risk factor $F2$ strongly correlates with the market factor *mkt*. The finding that the second latent risk factor does not have a significant risk premium lines up with the insignificant factor premiums that we find for the market risk. With respect to the relative contribution of factor loadings and characteristics to the overall variation in average returns we find that in the model including the latent risk factors, characteristics still explain the largest fraction of the overall variation with 74.52 %, while factors only explain 17.55 %. Despite that the latent risk factors of Kelly et al. (2019) are constructed from an advanced version of principal components analysis and

significance (t-values of 1.66 and 1.57 respectively).

are therefore theoretically better adapted to the cross section of returns compared to the factors of Fama and French (2015) and Carhart (1997), characteristics still explain the major fraction of the overall variation in average returns.

D.2. Risk Proxy Hypothesis

A closely related concern is the possibility that the significance of characteristics is derived from being a better proxy for true future betas than the estimated past betas themselves, which is labelled the 'risk proxy hypothesis'. Under this hypothesis, characteristics pick up explanatory power not because of capturing (yet) unidentified factors, but because of being a better empirical proxy for the future realization of the known factors. Although mitigated by the correction of the EIV bias, betas estimated from past data may still contain measurement error which makes them poorer predictors of future betas than characteristics. Jegadeesh et al. (2019) address the risk proxy from a rigorous perspective and derive an explicit estimator (labelled as IV mean-estimator) under the assumption that characteristics anticipate innovations in beta.[34] Essentially, their IV mean-estimator uses characteristics as instruments in the first-step regression and is based on time-series averages of returns as well as characteristics. We implement their approach to check the risk proxy hypothesis in our analysis, and report results in Table 8.[35]

As in Jegadeesh et al. (2019), we use Hansen-Hodrick standard errors with 12 lags to account for the overlapping data structure. We use the 3-factor model only, as the IV approach tends to produce a too large number of outliers as discussed in footnote 24. The results show very strong similarity with our base results. As risk factors, we only find *smb* to be significant (but not passing $t = 3$), while we have 14 out of 18 strongly significant characteristics. Thus, after controlling for the risk proxy hypothesis in the sense of the IV mean-estimator of Jegadeesh et al. (2019), we still find strong support for the generic influence of characteristics.
The strong robustness of our results to the risk proxy concern can, at least partially, be attributed to the fact that our characteristic variables are dummy variables that are defined from an already extended period of time. Thus, the time-averaging in the IV mean-estimator approach does not substantially alter their impact.

[34]See Jegadeesh et al. (2019), Section 5.1, in particular Proposition 2, p. 289.

[35]As mentioned earlier, the IV approach suffers from a strongly increasing number of outliers in the monthly slope coefficient estimates in particular when using an increasing number of risk factors. Therefore, we restrict our analysis to the three FF risk factors *mkt*, *smb*, and *hml*.

Table 8: In this table we report the test of the risk-proxy hypothesis according to Jegadeesh et al. (2019). The dependent variable is the average return over month t to month $t + 11$ in percent. The explanatory variables include separate betas for odd and even calendar months and the characteristic labels/indicator dummies. We run a two stage least squares regression with separate betas for odd and even calendar months. Characteristics and odd/even month betas serve as instrumental variables for even/odd month regressor betas. Standard errors are estimated according to Hansen-Hodrick with 12 lags. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%;

	Dependent variable r_{it}^{ex}	
$const$	0.634	(3.69)***
mkt	-0.299	(-1.08)
smb	0.432	(2.41)**
hml	0.053	(0.46)
IL^L	-0.074	(-5.89)***
IL^H	0.322	(4.05)***
PB^L	0.221	(2.90)***
PB^H	-0.345	(-5.62)***
MO^L	-0.065	(-0.85)
MO^H	0.133	(2.44)**
VO^L	-0.061	(-2.14)**
VO^H	-0.068	(-0.81)
BT^L	-0.027	(-0.74)
BT^H	-0.057	(-2.02)**
SZ^L	0.470	(3.09)***
SZ^H	0.181	(2.19)**
FD^L	0.116	(1.78)*
FD^H	1.123	(5.70)***
IV^L	0.390	(3.41)***
IV^H	-0.403	(-4.94)***
OP^L	-0.126	(-2.50)***
OP^H	0.238	(4.24)***

The risk proxy hypothesis can also be tackled from the perspective of allowing characteristics to be conditioning variables for factor loadings, such as in Harvey and Liu (2021) and Hoechle et al. (2020). Although their perspective is on time-varying betas, this approach is also suitable to test the concern if the significance of characteristics is completely absorbed into the interaction term with betas. Thus, we implement the GPS regression model according to Hoechle et al. (2020), which can be summarizes as:

$$r_{i,t}^{ex} = (\boldsymbol{z}_{i,t-1} \otimes \boldsymbol{x}_t)\boldsymbol{\theta} + c_i + v_{i,t}. \tag{16}$$

The vector $\boldsymbol{z}_{i,t} = [1 \; z_{2,i,t} \; ... \; z_{M,i,t}]$ contains a constant and firm characteristics $z_{m,i,t}(m = 2, ..., M)$. In our case these firm characteristics correspond to ϕ. The second vector $\boldsymbol{x}_t = [1 \; x_{1,t} \; ... \; x_{K,t}]$ consists of a constant and a set of risk factors $x_{k,t}(k = 1, ..., K)$, which vary over time, but not across firms, which in our implementation corresponds to the six risk factors above. Through the Kronecker product all possible multiplicative combinations of the vectors $\boldsymbol{z}$ and $\boldsymbol{x}$ enter the regression, thus the regression model consists of $M \times (K+1)$ explanatory variables. Vector $\boldsymbol{\theta}$ contains the coefficient estimates. Firm fixed effects are captured by c_i with $\mathrm{E}(c_i) = 0$. The strictly exogenous error term is denoted by $v_{i,t}$.[36]

By construction, the panel approach is different from cross-sectional regressions in estimating factor loadings instead of determining the set of variables which carry a risk premium. Still, we can focus on the fact if characteristics turn out to yield an independent contribution to mispricing, i.e. if characteristics themselves also seem to have incremental power in explaining returns even after allowing them to conditioning betas. We report results from the panel regression model in Table 20 in Appendix G. The first column reports contributions from characteristics to the total alpha, while columns 2–7 report contributions to the risk-factor loadings indicated on the top of the column. While generally speaking, a substantial number of interaction terms are strongly significant (and therefore also support the finding of time-varying loadings, as emphasized e.g. by Harvey and Liu (2021)), the more interesting finding for our purposes is the fact that 11 out of 18 characteristics

[36] Regression model (16) can be estimated with pooled OLS (without firm-fixed effects c_i) or with firm fixed effects. Hoechle et al. (2020) show that by estimating the regression model with pooled OLS, it is possible to exactly replicate portfolio alphas and betas from the standard portfolio sorts approach. As pooled OLS is only consistent if the random effects (RE) assumption holds, it follows that the standard portfolio sorts implicitly rests on the RE assumption. In order to address whether the RE assumption is likely to hold empirically, Hoechle et al. (2020) provide a specification test, which is similar to Hausman (1978). The null hypothesis of this Hausman (1978) type specification test is that the random effects assumption holds. If the null-hypothesis has to be rejected, this suggests that the GPS model should be estimated including firm fixed effects.

exhibit a statistically and economically substantial contribution to Jensen-alpha. We interpret this finding as strong evidence that although stock characteristics can condition factor loadings, they are still associated with a substantial contribution to mispricing. In this sense, the evidence is consistent with the main findings of the previous sections.

D.3. Limits to Arbitrage

The findings of our return and tilt regressions in the previous sections strongly suggest that return premiums related to ex-ante known stock and firm characteristics exist and that they are driven by institutional investors portfolio preferences. This result immediately faces the usual concern, why such apparent 'anomalies' are not arbitraged away. Among many other authors, Kozak et al. (2018) argue that contrary to sentiment-investors, arbitrageurs are able to use leverage and short positions to neutralized near-arbitrage opportunities, from which they conclude that taste-based demand distortions should have a limited pricing impact. This section provides three arguments why it appears unlikely that our results can easily be arbitraged away. First, univariate risk premia are difficult to isolate. Second, squared Sharpe ratios are not excessive. And third, portfolio tilts behind characteristic premia are highly persistent over time.

A first argument why these characteristic related return premiums are not arbitraged away is that the return of each individual stock is a multidimensional conglomerate of several factor and characteristic premiums. In order to isolate an individual characteristic-related return premium it is necessary to hedge out multiple risk factors and stock characteristics simultaneously. Arguably the number of stocks that can be used to build up a hedge portfolio becomes very small when sorting stocks according to multiple stock characteristics. Furthermore, since we define dichotomous characteristics labels (i.e. high vs. low quantiles) that will carry opposite risk premia (such as in the case of PB and OP where PB^L and OP^H carry positive premia, while PB^H and OP^L carry negative premia), it would be impossible to simultaneously go long and short in the same set of stocks to exploit the difference.
Thus, individual slope coefficients are not readily interpretable as return premia which can easily be exploited by a practical trading strategy. Instead, the evidence should better be understood as strong evidence that the cross-section of returns is strongly related to tastes for certain characteristics rather than risk aversion.

Therefore we argue that arbitrage focused on one particular characteristic premium is practically limited due to the large number of characteristics that influence the overall stock return.

Second, even under the unlikely assumption that individual characteristic premia were investable, it can shown that these premia are certainly not constant, but exhibit considerable time-series volatility. Already Fama (1976) pointed out that the estimated slope coefficients from the monthly Fama and MacBeth (1973) regressions can be interpreted as portfolio returns,[37] for which Sharpe ratios are a meaningful risk-return metric. Since it can be shown that the squared Sharpe (SR^2) ratio is related to an upper bound on the variance of the stochastic discount factor, Ross (1976) suggested to use two times the SR^2 of the market portfolio as an upper bound for the absence of near-arbitrage opportunities. This suggestion is taken up in Kozak et al. (2018) to identify arbitrage opportunities and we follow their line of reasoning to calculate the SR^2 for our characteristic premia. Results are reported in Table (9) for annualized SR^2 of the monthly factor and characteristic return contributions and the SR^2 of the excess market return.

Table 9: In this table we report the monthly squared Sharpe-Ratios (column SR^2) of the time series of the estimated return premiums from Table **??**. Additionally, we report the proportional squared Sharpe-Ratios of the characteristic premia relative to the squared Sharpe-Ratio of the excess market return (column SR^2 Ratio). According to Ross (1976) and Kozak et al. (2018) proportional Sharpe-Ratios above the boundary of 2 can be interpreted as near-arbitrage opportunities.

	SR^2	SR^2 Ratio		SR^2	SR^2 Ratio		SR^2	SR^2 Ratio
			$Market$	0.0224	1			
IL^L	0.0041	0.18	VO^L	0.0003	0.01	FD^L	0.0169	0.76
IL^H	0.0269	1.20	VO^H	0.0266	1.19	FD^H	0.0297	1.32
PB^L	0.0308	1.37	BT^L	0.0007	0.03	IV^L	0.0146	0.65
PB^H	0.0468	**2.09**	BT^H	0.0006	0.03	IV^H	0.0107	0.48
MO^L	0.0084	0.37	SZ^L	0.0332	1.48	OP^L	0.0268	1.19
MO^H	0.040	1.78	SZ^H	0.0008	0.04	OP^H	0.1488	**6.64**

Table (9) reveals that only the high growth PB^H and the high profitability characteristic OP^H show SR^2 multiples against the market portfolio of more than 2. High growth has a SR^2 ratio of 2.09 and is therefore only marginally beyond the upper boundary. Furthermore it carries a negative premium, which would require

[37] More recently, this interpretation is also emphasised by Back et al. (2013, 2015) and Fama and French (2020).

short positions to exploit it as near-arbitrage opportunity.[38] Only the high profitability characteristic carries a positive premium and a SR^2 multiple of more than 6 which is substantially beyond the Ross (1976) upper bound. Thus only one out of 18 characteristics represents a near-arbitrage opportunity. Vice versa, we find that 17 out of 18 characteristics are unlikely to be unexploited arbitrage opportunities. Furthermore, the evidence that most of the characteristic premia do not represent arbitrage opportunities has to be interpreted in view of the finding that characteristics explain the main share of the cross-sectional return variation. Combining both findings, it is unlikely that our results imply that the market is inefficient or offers arbitrage opportunities. Instead, our results rather offer the interpretation that the observed characteristic premia are the result of strong characteristic related preferences of (institutional) investors, that distort market prices and finally translate into stock returns that line up with characteristics.

Third, as final piece of evidence in this section, we analyze the persistence in demand tilts of institutional investors. The analysis is motivated by the reasoning that if a characteristic carries a substantial premia that can in principle be arbitraged away, and if institutional investors are sophisticated arbitrageurs, then their portfolio tilt should reflect their arbitrage activity over time. Putting it differently: Given that institutional investors account for a majority of the market share, the fact that portfolio tilts are highly persistent over time can be interpreted as an indication that there is not sufficient arbitrage (i.e. smart) money which will exploit the premium.

We investigate portfolio tilt persistence by aggregating the stock-level tilts over all stocks that carry a particular characteristic label. Similar to Equation (3), we compute the aggregate institutional characteristic demand tilt in characteristic j at time t as:

$$T^P_{j,t} = \frac{\mathbf{x_t^P} \cdot \boldsymbol{\phi}'_{j,t}}{\mathbf{x_t^M} \cdot \boldsymbol{\phi}'_{j,t}}, \tag{17}$$

where $\boldsymbol{\phi}'_{j,t}$ is again the I-vector of dummies which indicates if asset i carries characteristic j at time t. The scalar product with the weights vectors aggregates over all stocks, so $T^P_{j,t}$ summarizes how strongly institutional investors over- or underweight stocks of a particular characteristic class relative to the corresponding market weights of these stocks at time t. In Table (10), we report the (time-series) average aggregate characteristic tilts for all 18 characteristics. Additionally we

[38] Among others, e.g. Frazzini and Pedersen (2014) argue that large groups of investors face restrictions to short selling.

report pairwise significance tests for the comparison of the top and bottom characteristic portfolio tilts.[39] The third column reports p-values from an Augmented-Dickey-Fuller test. We conduct the analysis for the full sample, as well as the sample that excluded the smallest 20% of stocks (labelled Non-Micro). We further illustrate the time-series pattern of the characteristics in the full sample by plotting the quarterly aggregate characteristic tilt in Figure 5, where the blue solid (purple dashed) line refers to the low (high) manifestation of each characteristic. We include the same figure for the non-micro sample as Figure 8 in Appendix H.

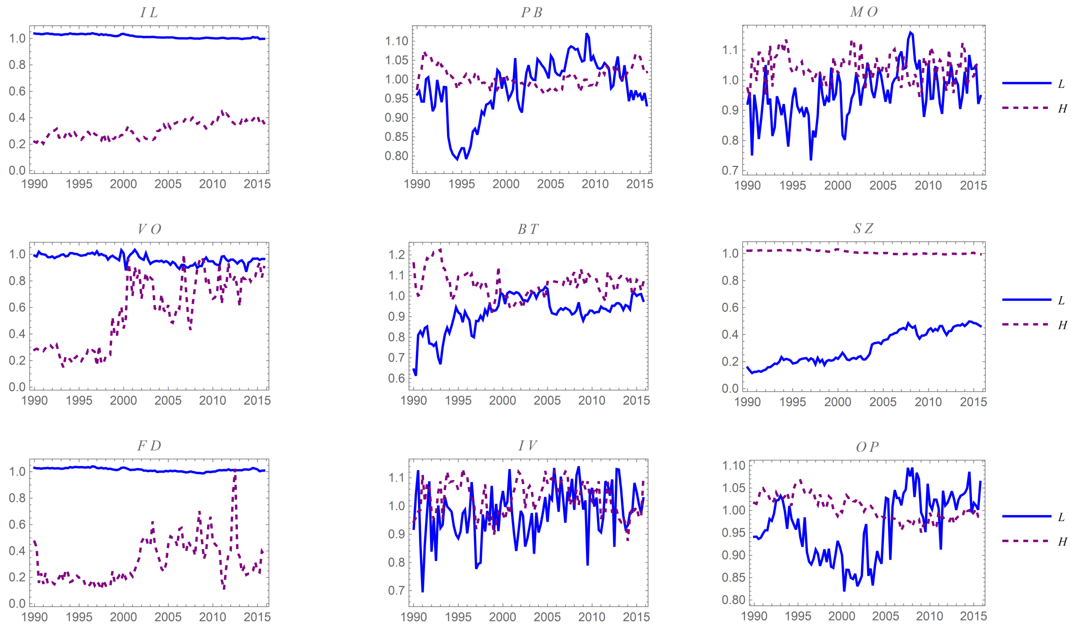
Table 10: In this table we report in column $Mean$ the average aggregate characteristic tilt according to Equation (17). In column $Difference$ we report the average tilt difference between the top and bottom characteristic sorts, with a corresponding t-test to test for the significance of the difference. We want to mention that a Mann-Whitney-U-test yields similar results as the t-test. In the last column $ADF\ p$ we report p-values from an augmented Dickey-Fuller unit root test. We report separate results for the sample that includes all stocks and the sample without microcap stocks. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%;)

	Full Sample			Non-Micro		
	$Mean$	Difference H-L	ADF p	$Mean$	Difference	ADF p
IL^L	1.015	-0.705	0.610	1.005	-0.202	0.494
IL^H	0.031	(112.10)***	0.207	0.803	(14.78)***	0.496
PB^L	0.975	0.029	0.203	1.002	0.006	0.146
PB^H	1.005	(-3.71)***	0.047**	1.008	(-1.08)	0.083*
MO^L	0.957	0.082	0.000***	0.994	0.060	0.000***
MO^H	1.039	(-8.87)***	0.000***	1.054	(-7.73)***	0.000***
VO^L	0.965	-0.398	0.0014***	0.924	0.089	0.001***
VO^H	0.566	(15.28)***	0.203	1.014	(-8.95)***	0.000***
BT^L	0.919	0.141	0.011**	0.937	0.142	0.042**
BT^H	1.059	(-13.45)***	0.0029***	1.079	(-15.88)***	0.037**
SZ^L	0.307	0.704	0.814	1.013	-0.013	0.611
SZ^H	1.011	(-58.65)***	0.638	0.999	(0.89)	0.454
FD^L	1.016	0.697	0.005***	1.015	0.168	0.541
FD^H	0.318	(-42.51)***	0.314	0.847	(-10.83)***	0.174
IV^L	0.976	0.051	0.000***	1.029	0.017	0.000***
IV^H	1.028	(-4.83)***	0.000***	1.046	(-1.77)*	0.000***
OP^L	0.965	0.039	0.034**	1.032	-0.034	0.073*
OP^H	1.004	(-5.26)***	0.182	0.998	(4.86)***	0.130

In the full sample, we find statistically significant differences for all characteristics. From the numerical values, in particular liquidity IL, size SZ and distress FD show the largest tilt differences. Excluding micro-caps has a strong impact on the results, as the difference in size SZ tilts is no longer significant and differences for IL and FD are substantially smaller. Together this insinuates that size, liquidity, and distress are characteristics which together are important among the

[39] Results are from t-tests. Similar results are obtained from Mann-Whitney-U-Tests.

Figure 5: In this figure, we plot the time series for aggregate characteristics tilts $T^{P}_{j,t}$ from equation (17). Each panel reports the low (high) manifestation of the characteristic as blue solid (purple dashed) line. Sample includes all stocks. Quarterly frequency. Time period 1990Q1 – 2015Q4.



small stocks. We explore the role of excluding small stocks in the following and final robustness subsection.

D.4. Non-micro Sample

In this section we provide robustness checks for the tilt and return regressions, where we exclude microcap stocks. Similar to Green et al. (2017) we define all stocks that are smaller than the 20% NYSE size-quantile as microcap stocks. Overall this reduces our sample by approximately 50% of observations.

In column 3 of Table (11) we report results for the panel tilt regressions without microcap stocks. The results for the sample without microcap stocks are comparable to the results for the full sample. Institutional investors' portfolio tilts are strongly related to stock characteristics.

Table 11: In this table we report the empirical estimation results for the Fama and MacBeth (1973) regressions according to Equation (5), where we *exclude microcap* stocks: $r_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{h,t} \hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{j,t} \phi_{i,j,t-1} + \epsilon_{i,t}\ \forall\ t = 1, ..., T$. The dependent variable is the monthly excess market return of individual stocks in percent. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. The factor betas are estimated with respect to the five Fama and French (2015) factors augmented with the Carhart (1997) momentum factor. For the factor betas we apply the EIV correction procedure from Pukthuanthong et al. (2019). The reported estimates in the table are the time-series averages of the monthly slope-coefficient estimates. We provide two separate results. As the exclusion of microcap stocks reduces the number of stocks in the sample by approximately 50 %, selecting high and low investment stocks leads to a small number of stocks in the respective categories in the period from 1975 to 1985, due to lacking data on total assets at quarterly frequency. Therefore, we provide results for two sample periods. The first sample period is January 1975 to June 2016, but excludes the high/low investment characteristic. The second sample is from January 1985 to June 2016 and includes the full set of characteristics. In column three we provide results according to Equation 8), where we *exclude microcap*¡ stocks: $T_i = \alpha_0 + \sum_{f=1}^{F} \theta_{ft} \hat{\beta}_{if} + \sum_{j=1}^{J} \theta_{jt} \phi_{ij} + c_i + p_t + \epsilon_{it}$. The dependent variable is the demand tilt, winsorized at the 97 % quantile. The explanatory variables include the univariate factor betas with respect to the market, SMB, HML, CMA, RMW and momentum factor, estimated over a rolling window of 36 months and winsorized at the 1 % and 99 % quantile. Additionally, we include the characteristic labels/indicator dummies, which identify whether a particular stock belongs to the top/bottom quantile according to the corresponding characteristic sort (see Section B.3). The regression coefficients can be interpreted as the change in percentage points in the average stock tilt, when multiplied with 100. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	(1)			(2)			(3)		
				Dependent variable:					
	Return $r_{i,t}^{ex}$			Return $r_{i,t}^{ex}$			Tilt $T_{i,t}$		
const	0.794	(3.90)	***	0.845	(4.75)	***	0.949	(107.9)	***
mkt	-0.16	(-1.05)		-0.099	(-0.67)		-0.004	(-1.13)	
smb	-0.134	(-1.67)		0.003	(0.04)		-0.00	(-0.28)	
hml	-0.081	(-0.80)		-0.45	(-0.49)		0.001	(0.46)	
cma	0.107	(1.59)		0.095	(1.64)		0.006	(2.96)	***
rmw	0.026	(0.40)		0.069	(1.16)		0.008	(6.80)	***
mom	-0.15	(-1.26)		-0.176	(-1.59)		0.003	(1.22)	
IL^L	0.052	(0.76)		0.002	(-0.03)		0.024	(4.85)	***
IL^H	-0.07	(-1.13)		-0.009	(-0.17)		-0.040	(-10.62)	***
PB^L	0.054	(0.90)		0.110	(2.15)	**	0.003	(0.68)	
PB^H	0.025	(0.34)		-0.064	(-1.02)		0.014	(2.93)	***
MO^L	-0.074	(-0.77)		-0.050	(-0.64)		-0.026	(-6.77)	***
MO^H	0.222	(2.53)	**	0.246	(3.27)	***	0.004	(0.91)	
VO^L	0.026	(0.36)		-0.036	(-0.59)		-0.006	(-1.79)	*
VO^H	-0.57	(-3.37)	***	-0.582	(-4.15)	***	-0.021	(-5.42)	***
BT^L	0.064	(1.05)		0.038	(0.70)		-0.010	(-3.11)	***
BT^H	-0.008	(-0.107)		0.014	(-0.23)		0.028	(6.63)	***
SZ^L	0.076	(1.15)		0.084	(1.57)		0.011	(2.78)	***
SZ^H	-0.078	(-1.11)		-0.077	(-1.31)		0.005	(0.96)	
FD^L	0.054	(1.05)		0.053	(1.14)		0.048	(9.03)	***
FD^H	0.116	(0.77)		0.106	(0.83)		-0.096	(-11.80)	***
IV^L	0.013	(0.170)	***	-	-		0.005	(1.94)	*
IV^H	-0.128	(-1.51)	***	-	-		0.022	(6.91)	***
OP^L	-0.143	(-2.18)	**	-0.121	(-2.18)	**	0.018	(4.69)	***
OP^H	0.177	(3.13)	***	0.183	(3.67)	***	0.001	(0.32)	
Observations	377			497			173,633		
Within-R^2							6.93 %		

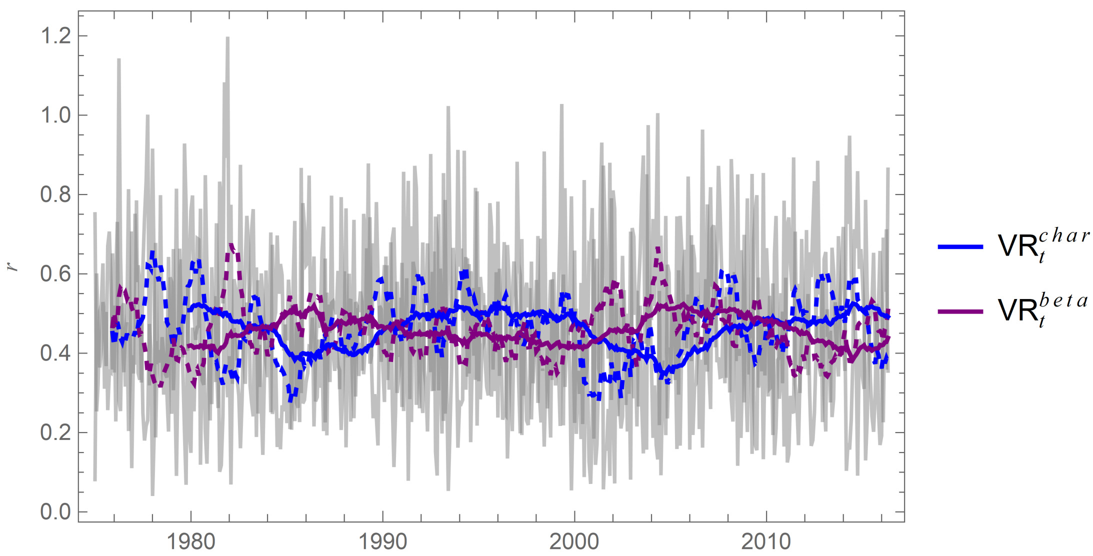
	(1)	(2)	(3)

Table continued:

Also for the cross-section of stock returns characteristics are pricing relevant, as we find significant and economically large characteristic premiums for the momentum, volatility and profitability characteristics. We argue that the finding that the characteristics illiquidity, size and financial distress are not pricing relevant in the sample without microcap stocks directly results from excluding the smallest stocks in the sample. Excluding approximately 50 % of the smallest stocks from the dataset necessarily affects the findings on the size characteristic premium. Furthermore, financial distress and illiquidity is also higher among small stocks. Therefore, the characteristic premiums are directly affected by excluding microcap stocks.

Further evidence that characteristics are pricing relevant for the cross-section of stock returns is provided in Figure (6).

Figure 6: In this table we plot the time series of the cross-sectional variation, which is explained by factor-risk exposure (i.e. betas) and characteristics over time. As shown by Chordia et al. (2015) due to the correlation between factor betas and characteristics the two ratios do not necessarily have to add up to 1. The mean over the full sample period for the explanatory power of characteristics is 46.34 %, while the corresponding mean is 45.72 % for the factor betas. The dotted and the solid line respectively show the moving average of the factor/beta explained component over a 12 month and 60 month rolling window.



In comparison to the full sample, the cross-sectional return variation which is explained by factor risk-exposure, is larger for the sample without microcap stocks.

Indeed, factor risk exposure and characteristics explain an equal proportion of approximately 45 % of the cross-sectional variation in stock returns. This finding holds throughout the full sample period. Contrasting with the findings of Green et al. (2017) we do not document any deterioration in the characteristic-explained part.

Table 12: In Column 1 we report the results of the panel regression according to Equation (13), where we *exclude microcap* stocks. In Columns 2 and 3 we report the estimation results of the panel regression according to Equation (14). Coefficients are reported in percent***. Standard errors are estimated according to Driscoll and Kraay (1998). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	(1)		(2)		(3)		(4)	
	Dependent variable							
	r_{it}^{ex}		r_{it}^{ex}		$\hat{r}_{it}^{ex,beta}$		$\hat{r}_{it}^{ex,char}$	
constant	0.130		0.037		-0.022		-0.019	
	(7.17)	***	(3.02)	***	(1.81)	*	(-8.10)	***
$\hat{T}_{i,t-1}$	-0.081							
	(-4.80)	***						
$\hat{T}_{i,t-1}^{beta}$			-0.000		0.035		0.027	
			(0.00)		(1.41)		(1.88)	*
$\hat{T}_{i,t-1}^{char}$			-0.046		0.009		-0.008	
			(-2.15)	**	(0.85)		(0.027)	
Time-Fixed Effects	Yes		Yes		Yes		Yes	
Firm-Fixed Effects	Yes		Yes		Yes		Yes	
Within-R^2	17.97%		22,20%		50.89%		11.76 %	
Observations	171,822		154,939		154,939		154,939	

Finally, in Table (22) we provide results on how demand tilts are related to the cross-section of stock returns. Similar to the results for the full sample we find that institutional portfolio tilts are negatively related to consecutive stock returns (see column (1)). The decomposition of tilts into a factor-induced and characteristic-induced demand component shows that only characteristic-induced demand tilts are relevant for the cross-section of stock returns, whereby the corresponding coefficient is significant at the 5 % level. The finding in column (3), shows a slightly significant and positive coefficient for factor related demand tilts on consecutive characteristic related returns. Indeed, the factor related tilt component is zero in the excess return regression of column (2), which is a more parsimonious model, as it does not require any first-step decomposition of the left-hand variable (i.e. the returns). We therefore argue that the results of column (2) are more robust evidence of a relationship between characteristic-induced demand tilts and consecutive stock returns. Despite that overall the estimated coefficients are smaller for the non-microcap sample, compared to the full-sample, we conclude that institutional demand tilts are also pricing relevant for the cross-section of non-microcap returns.

E. Conclusion

It is remarkable that a very recent contribution of Harvey and Liu (2021) concludes that even after 50 years of research in asset pricing, there is still no canonical factor model. Even worse, there seems to be inconclusive evidence with respect to the relevance of risk factors versus stock characteristics. Recent contributions, based on novel econometric advancements, find surprisingly little support for factors that have long been considered as well-established risk-based explanatory variables. One reason for these contradictory results is an underlying identification dilemma which makes it hard to discriminate between risk factors and characteristics only on the basis of return data, as recently stressed by Kozak et al. (2018) or Pastor et al. (2020).

Our contribution wants to help resolve this dilemma by resorting to institutional investor holdings data. Our main result is to show that – in line with recent literature – the common (six) risk factors find little empirical support when tested on a sample of individual stocks in EIV-corrected cross-sectional regressions. Instead characteristics are associated with significant return premia and explain roughly 2/3 of the variation in expected returns. Furthermore, we find no evidence that this variance decomposition declined in the most recent decades. To assess the importance of characteristics, we propose a novel definition which treats characteristics as binary indicators and thereby captures the idea that only salient facts of stocks will be recognized by investors. Our main novel contribution is to define portfolio tilts of individual institutional investors (derived from 13(f) filings) and to show that institutional demand is strongly related to characteristics but not to risk factors. We further provide evidence that the institutional demand tilt is negatively related to expected returns, thereby confirming related evidence that professional investors appear to fail to exploit alleged anomalies.

Overall, the evidence from our analysis strongly suggests that institutional investors' demand is rather driven by taste for characteristics than aversion to risk factors and that this pattern is reflected in the cross-section of expected returns.

A. Model description

A.1. Modeling demand towards stock and firm characteristics

Consider K investors indexed by $k = 1, \ldots, K$ investing in $i = 0, 1, \ldots, N$ assets, where asset 0 is the risk-free asset, returning rate r_f. Starting wealth of each investor is W_0^k. We assume that returns are jointly normally distributed, which directly results from asset prices being normally distributed $\mathbf{p} \sim \mathcal{N}(\bar{\mathbf{p}}, \mathbf{\Sigma})$. Additionally, each asset i can be characterized by a set of characteristics $\phi_{ij} \in \{0, 1\}$, where $j = 1, \ldots, J$. We interpret ϕ as an indicator variable. For example, ϕ may capture if an asset is distressed or not. Thus, $\phi_{ij} = 1$ would imply that the asset i is distressed whereas $\phi_{ij} = 0$ would mean that the asset i is not distressed. In that sense the indicator variable ϕ can be interpreted as a binary label of important asset characteristics.Asset returns are additionally impacted by H risk factors, indexed by $h = 1, \ldots, H$. The risky final wealth of investor k, who invests the portfolio weights $\mathbf{x_k}$, is given by:

$$W_1^k = W_0^k[1 + (r_f + \mathbf{x}_\mathbf{k}'(\boldsymbol{r} - r_f\mathbf{1}))]. \tag{18}$$

Now, each investor k is assumed to be risk averse and having a preference according to the j-th characteristic: ϕ_j for $j = 1, \ldots, J$. To be explicit, we define the utility function in terms of (final) wealth and characteristics in the following way:

$$u^k(W_1^k, \boldsymbol{\phi}, \boldsymbol{Z}) = -e^{-a^k W_1^k + a^k \sum_j^J b_j^k \mathbf{x}_\mathbf{k}' \phi_j + a^k \sum_h^H c_h^k Z_h}, \tag{19}$$

where a^k measures the parameter of constant absolute risk aversion, b_j^k is the (relative) attitude towards the asset's characteristic j, and c_h^k is the attitude towards risk factor Z_h impacting the asset returns. Thus, asset characteristics can provide non-pecuniary utility/disutility to investors. According to our definition positive preferences for a particular characteristic correspond to negative values of b_j^k (i.e. $b_j^k < 0$). Observe that we make no statement about the sign of the b_j^k since positive as well as negative values are allowed. We assume that the H risk factors are jointly normal distributed with variance-covariance matrix $\mathbf{\Sigma}_Z$ as well.

In a (perfect) world under the assumption that there are no asset characteristics and without additional risk factors, the model leads to the well known Sharpe-Lintner-Mossin capital asset pricing model since $\phi_j = 0 \ \forall \ j$ and $Z_H = 0 \ \forall \ h$. The same result applies if the b_j^k and c_h^k would be zero for all investors. Thus, the

CAPM is just a special case of our model. Note that the characteristics are not assumed to be random variables, since they are known at time $t = 0$. They simply constitute exogenously given inputs to the model.

From the assumptions the expected utility is given by

$$\begin{aligned}
\mathbb{E}[u^k(W_1^k, \boldsymbol{\phi}, \boldsymbol{Z})] &= \mathbb{E}\left[-e^{-a^k W_1^k + a^k \sum_j^J b_j^k \mathbf{x_k}' \phi_j + a^k \sum_h^H c_h^k Z_h}\right] && (20)\\
&= -e^{-a^k V^k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\phi}, \boldsymbol{Z})} && (21)
\end{aligned}$$

where the value function V^k is defined as:

$$\begin{aligned}
V^k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\phi}, \boldsymbol{Z}) &= W_0^k[1 + (r_f + \mathbf{x_k}'(\boldsymbol{\mu} - r_f \mathbf{1}))] - \frac{a^k}{2}(W_0^k)^2 \mathbf{x_k}' \boldsymbol{\Sigma} \mathbf{x_k} && (22)\\
&- \sum_{j=1}^{J} W_0^k \hat{b}_j^k \mathbf{x_k}' \boldsymbol{\phi}_j - \sum_{h=1}^{H} c_h^k \mathbb{E}[Z_h] - \sum_{h=1}^{H} \frac{a^k (c_h^k)^2}{2} \mathbb{V}ar(Z_h) && (23)\\
&- \sum_{h=1}^{H} \mathbf{x_k}' c_h^k W_0^k \mathbb{C}ov(\boldsymbol{r}, Z_h) + \sum_{h=1}^{H-1} \sum_{l=h+1}^{H} c_h^k c_l^k \mathbb{C}ov(Z_h, Z_l) && (24)
\end{aligned}$$

Note that we scale the attitude towards the characteristic by the wealth level such that $b_j^k \equiv W_0^k \hat{b}_j^k$, whereby $\hat{b}_1^k$ is a scaled parameter of the attitude towards characteristic j. This transformation simplifies the expression of the characteristic preferences in the following optimization problem.[40] Maximizing with respect to the portfolio weights $\mathbf{x_k}$ and setting the result to zero delivers the optimal asset weights:

$$\mathbf{x_k^*} = \frac{\boldsymbol{\Sigma^{-1}}(\boldsymbol{\mu} - r_f \mathbf{1})}{a^k W_0^k} - \sum_{j=1}^{J} \frac{\boldsymbol{\Sigma^{-1}} \hat{b}_j^k \boldsymbol{\phi}_j}{a^k W_0^k} - \sum_{h=1}^{H} \frac{\boldsymbol{\Sigma^{-1}} c_h^k \mathbb{C}ov(\boldsymbol{r}, Z_h)}{a^k W_0^k} \tag{25}$$

Note that, according to our definition b_j^k is negative, if an investor has a positive preference for characteristic ϕ_j. Hence, if $b_j^k < 0$, the individual optimal demand ceteris paribus increases if asset i exhibits characteristic j. Agents with positive preferences for a particular stock characteristic thus receive non-pecuniary utility from investing into assets that exhibit a particular characteristic and are therefore accepting lower expected returns. According to Pastor et al. (2020) preferences for "green" stock investments provide a potential example for this type of characteristic preferences.

[40] It is sufficient to maximize the exponent, as the exponential function is monotonically increasing in the exponent.

A.2. Equilibrium expected returns

Summing Equation (25) over K investors for each stock and dividing by the aggregated wealth W_0^{total} gives the weights of the market portfolio:

$$\mathbf{x_M} = \frac{\sum_{k=1}^{K} \mathbf{x_k^*} W_0^k}{W_0^{total}} \tag{26}$$

Using our expression for the optimal asset weights $\mathbf{x_k^*}$ and defining $A = \sum_{k=1}^{K} \frac{1}{a^k}$, $B_j = \sum_{k=1}^{K} \frac{\hat{b}_j^k}{a^k}$, $C_h = \sum_{k=1}^{K} \frac{c_h^k}{a^k}$, $\kappa_M = \frac{W_0^{total}}{A}$, $\kappa_h = \frac{C_h}{A}$, and $\kappa_j = \frac{B_j}{A}$ we get the following pricing equation:

$$\boldsymbol{\mu} - r_f \mathbf{1} = \kappa_M \mathbf{\Sigma} \mathbf{x_M} + \sum_{j=1}^{J} \kappa_j \boldsymbol{\phi}_j + \sum_{h=1}^{H} \kappa_h \mathbb{C}ov(\boldsymbol{r}, Z_h) \tag{27}$$

We reach our final asset pricing equation for asset i by setting $\hat{\kappa}_M = \kappa_M \sigma_M^2$ and $\hat{\kappa}_h = \kappa_h \psi \mathbb{V}ar(r_h)$, where portfolio return r_h is perfectly correlated with Z_h, i.e. $r_h = Z_h/\psi$, such that $\mathbb{C}ov(\boldsymbol{r}, Z_h) = \psi \mathbb{C}ov(\boldsymbol{r}, r_h)$:

$$\mu_i - r_f = \beta_{iM} \hat{\kappa}_M + \sum_{h=1}^{H} \beta_{ih} \hat{\kappa}_h + \sum_{j=1}^{J} \kappa_j \phi_{ij} \tag{28}$$

Equation (28) is our main theoretical result: the excess return of asset i is a function of the asset's beta with the market, the H mimicking portfolios, the return contribution of the market plus the return contributions of the relevant characteristics, given the asset exhibits the respective characteristic j. Note that $\beta_{iM} = \frac{\mathbb{C}ov(r_i, r_M)}{\mathbb{V}ar(r_M)}$ and $\beta_{ih} = \frac{\mathbb{C}ov(r_i, r_h)}{\mathbb{V}ar(r_h)}$ have the interpretation of standard bivariate OLS betas, not multivariate regression betas.

B. Descriptive statistics

Table 13: This table provides descriptive statistics for the individual excess stock returns and the stock and firm characteristics, which are the basis for the dummy-variable assignment procedure. The descriptive statistics cover the mean, median, standard deviation (σ), the 25% and 75% percentiles. Additionally, we provide summary statistics on the average, minimum and maximum number of firms that are included in our sample in each month. The sample period is January 1975 to May 2016.

Variable	Mean	Median	σ	25% Percentile	75% Percentile
Panel A. Return and Stock Characteristics					
Excess Return	0.010	-0.004	0.193	-0.075	0.075
Size	2,021.715	115.467	12,618.60	24.747	631.879
B/M-ratio	0.758	0.584	7.470	0.313	1.014
Momentum	0.082	0.024	0.521	-0.170	0.238
Trading Liquidity	1.074	0.011	25.690	0.001	0.140
Volatility	0.036	0.028	0.033	0.018	0.044
Market Beta	0.808	0.779	0.781	0.281	1.687
Financial Distress	0.009	0.000	0.078	0.000	0.000
Profitability	-0.116	0.042	26.81	-0.012	0.080
Investments	0.038	0.010	1.059	-0.023	0.049
Panel B. Tilt measure					
Full Sample	0.705	0.692	0.490	0.236	1.151
Small Stocks	0.236	0.141	0.262	0.033	0.353
Large Stocks	1.070	1.150	0.361	0.875	1.341

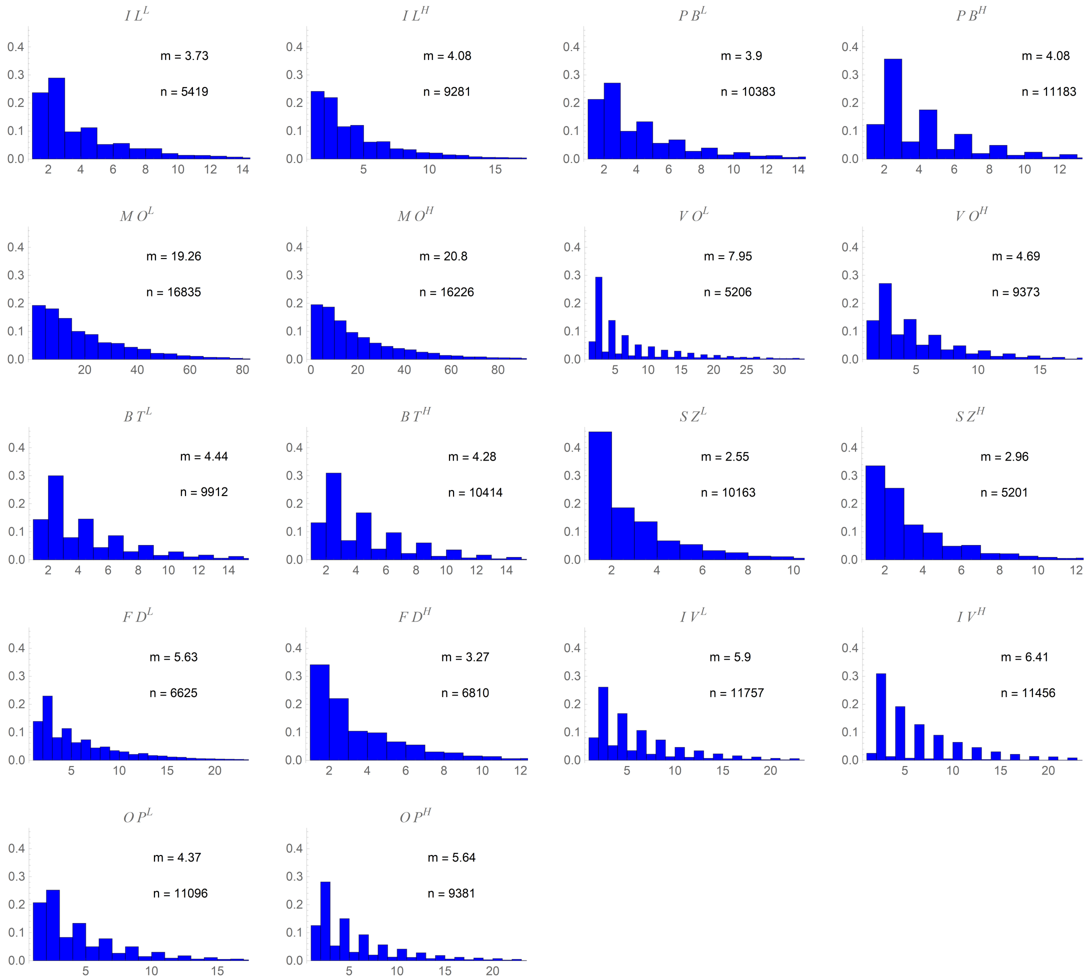
Table 14: In this table we report variance inflation factors (VIFs) for each of the 18 characteristic labels/dummies as described in Section B.3. The VIFs stem from a multivariate panel regression of monthly excess stock returns on the set of 18 characteristic dummies. Additionally we report VIFs after excluding the two trading liquidity identifiers, which are based on the Amihud (2002) illiquidity measure.

	VIF			VIF			VIF	
	(1)	(2)		(1)	(2)		(1)	(2)
SZ^H	3.11	1.77	MO^L	1.32	1.32	OP^H	1.16	1.16
IL^L	2.72	-	VO^H	1.32	1.27	PB^L	1.15	1.18
SZ^L	1.89	1.38	VO^L	1.25	1.24	IV^L	1.11	1.16
FD^L	1.86	1.84	MO^H	1.24	1.57	BT^L	1.10	1.10
IL^H	1.8	-	PB^H	1.19	1.19	BT^H	1.07	1.07
FD^H	1.32	1.32	OP^L	1.19	1.19	IV^H	1.05	1.05

Table 15: In panel A and Panel B of this table we report descriptive statistics for the regressor betas in the regression model of Equation (5), which are based on the EIV-correction procedure from Jegadeesh et al. (2019). The betas are univariate factor betas with respect to the market ($MKTRF$), small-minus-big (SMB), high-minus-low (HML), conservative-minus-aggressive (CMA), robust-minus-weak (RMW) and high minus low momentum (MOM) factors. The betas are estimated from daily returns over the preceeding 720 trading days. We apply the correction according to Dimson and Mussavian (1999) to the daily beta estimates. In panel C of this table we report descriptive statistics for the factor betas, which are estimated based on the EIV-correction procedure of Pukthuanthong et al. (2019). In this approach stocks are sort into deciles of market-capitalistion and factor betas. The individual factor beta is then replaced by the average beta of the corresponding size-beta category.

	β_{mkt}	β_{smb}	β_{hml}	β_{cma}	β_{rmw}	β_{mom}
Panel A. Statistics EIV-corrected Betas						
Mean	0.919	0.720	-0.683	-0.978	-0.766	0.099
Median	0.874	0.607	-0.519	-0.875	-0.598	0.039
Maximum	4.730	7.35	2.845	9.436	12.99	3.95
Minimum	-0.732	-4.635	-9.866	-7.647	-9.612	-3.089
Std. Dev.	0.606	0.972	1.246	1.550	1.628	0.882
Observations	1,831,539	1,831,539	1,831,539	1,831,539	1,831,539	1,831,539
Panel B. Statistics Even-/Odd-Month Betas						
Even-Month Betas						
Mean	0.918	0.693	-0.753	-1.099	-0.789	0.159
Median	0.871	0.577	-0.604	-0.957	-0.649	0.085
Maximum	4.125	5.017	4.769	7.262	6.921	4.526
Minimum	-1.586	-3.101	-6.642	-9.194	-8.999	-4.199
Std. Dev.	0.666	1.111	1.350	1.720	1.890	1.011
Observations	928,083	928,083	928,083	928,083	928,083	928,083
Odd-Month Betas						
Mean	0.905	0.693	-0.598	-0.840	-0.720	0.007
Median	0.864	0.565	-0.495	-0.740	-0.559	-0.0168
Maximum	3.417	5.147	4.670	7.750	6.120	4.427
Minimum	-1.897	-2.509	-5.377	-8.461	-9.051	-3.746
Std. Dev.	0.659	1.095	1.309	1.619	1.828	0.933
Observations	930,220	930,220	930,220	930,220	930,220	930,220

Figure 7: In this figure, we report histograms for the number of changes individual stocks change the status of being assigned a particular characteristic. Each panel contains the number n of stocks from the full sample that were assigned this characteristic at least once in the sample period. Each panel reports the mean number m of status changes for that characteristic label.



C. Alternative EIV correction and OLS Results

Table 16: In this table we report the empirical estimation results for the standard OLS regressions and the EIV corrected Fama and MacBeth (1973) regressions according to Equation (5): $r_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{h,t}\hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{j,t}\phi_{i,j,t-1} + \epsilon_{i,t}\ \forall\ t = 1, ..., T$. The dependent variable is the monthly excess market return of individual stocks in percent. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. For the EIV-correction we apply the instrumental variable approach as proposed by Jegadeesh et al. (2019), which uses separate factor betas for odd and even months. The factor betas are estimated with respect to the five Fama and French (2015) factors augmented with the Carhart (1997) momentum factor. The reported estimates in the table are the time-series averages of the monthly slope-coefficient estimates. Within the EIV-corrected approach we exclude all coefficient estimates that exceed the boundary of six standard-deviations around the corresponding mean factor/characteristic portfolio return. For the volatility characteristic and the distress characteristic we exclude observations that are six times larger than their corresponding OLS coefficient estimates. Both described criterions are proposed by Jegadeesh et al. (2019). The sample period is January 1975 to June 2016. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	(1) OLS		(2) EIV – FF3		(3) EIV – FFC6		(4) Multivariate β	
const	0.643		0.701		0.520		0.785	
	(3.33)	***	(3.03)	**	(1.55)		(3.92)	***
mkt	-0.018		0.104		0.189		-0.006	
	(-0.134)		(0.25)		(0.32)		(-0.07)	
smb	0.081		0.35		0.393		0.086	
	(1.46)		(1.76)	*	(1.11)		(1.78)	*
hml	-0.065		0.61		0.219		0.030	
	(-0.72)		(2.25)	**	(0.46)		(0.58)	
cma	0.074		-		1.039		0.074	
	(1.42)		-		(3.25)	***	(1.13)	
rmw	0.040		-		0.584		0.006	
	(0.756)	***	-		(1.87)	*	(0.756)	
mom	-0.245		-		-0.696		-0.116	
	(-2.44)	**	-		(-1.24)		(-1.40)	
IL^L	-0.070		-0.095		-0.092		-0.086	
	(-1.17)		(-1.55)		(-1.55)		(-1.37)	
IL^H	0.338		0.374		0.367		0.281	
	(3.90)	***	(4.41)	***	(4.32)	***	(3.14)	***
PB^L	0.247		0.238		0.234		0.200	
	(4.00)	***	(3.76)	***	(3.75)	***	(3.23)	***
PB^H	-0.308		-0.330		-0.323		-0.299	
	(-4.91)	***	(-4.79)	***	(-4.81)	***	(-4.61)	***
MO^L	-0.198		-0.288		-0.240		-0.113	
	(-2.03)	**	(-3.57)	***	(-2.76)	***	(-1.10)	
MO^H	0.361		0.296		0.293		0.284	
	(5.44)	***	(4.43)	***	(4.51)	***	(3.79)	***
VO^L	-0.022		-0.048		-0.051		-0.022	
	(-0.34)		(-0.65)		(-0.74)		(-0.31)	
VO^H	-0.531		-0.53		-0.513		-0.431	
	(-3.76)	***	(-3.57)	***	(-3.59)	***	(-2.74)	***
BT^L	-0.014		-0.046		-0.046		-0.093	
	(-0.289)		(-0.91)		(-0.92)		(-1.87)	*
BT^H	0.047		0.054		0.057		-0.072	
	(0.84)		(0.90)		(0.93)		(-1.06)	
SZ^L	0.418		0.410		0.397		0.456	
	(4.27)	***	(4.02)	***	(3.97)	***	(4.42)	***
SZ^H	-0.024		-0.062		-0.067		-0.073	
	(-0.35)		(-0.87)		(-0.947)		(-1.01)	
FD^L	0.173		0.141		0.137		0.121	
	(2.38)	**	(2.43)	**	(2.44)	**	(2.22)	**
FD^H	0.773		0.807		0.796		0.814	
	(3.17)	***	(3.87)	***	(3.85)	***	(3.92)	***
IV^L	0.252		0.045		0.059		0.211	
	(2.38)	**	(0.58)		(0.75)		(1.78)	*
IV^H	-0.128		-0.124		-0.113		-0.181	
	(-1.94)	*	(-2.26)	**	(-2.10)	**	(-2.63)	***
OP^L	-0.246		-0.291		-0.292		-0.248	

Table continued:				
	(1)	(2)	(3)	(4)
	(-3.51) ***	(-4.43) ***	(-4.54) ***	(-3.47) ***
OP^H	0.422	0.414	0.415	0.383
	(8.93) ***	(8.34) ***	(8.64) ***	(8.05) ***
Sample Period	1975-2016	1975-2016	1975-2016	1975-2016

D. Alternative Tilt and Return Regressions

Table 17: In this table we report in column (1) results for the quarterly Fama and MacBeth (1973) regressions in order to decompose the tilts according Equation (**??**): $T_{it} = \tau_{0t} + \sum_{h=1}^{H} \tau_{h,t}\hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \tau_{j,t}\phi_{i,j,t-1} + \epsilon_{i,t} \; \forall \, t = 1, ..., T$. The dependent variable is the institutional tilt. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. In column (2) we report results for the Equation (8): $T_i = \alpha_0 + \sum_{f=1}^{F} \theta_{ft}\hat{\beta}_{if} + \sum_{j=1}^{J} \theta_{jt}\phi_{ij} + p_t + \epsilon_{it}$, whereby we exclude firm fixed effects. The dependent variable is the institutional tilt. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. The dependent variable is the demand tilt, winsorized at the 97 % quantile. The explanatory variables include the univariate factor betas with respect to the market, SMB, HML, CMA, RMW and momentum factor, estimated over a rolling window of 36 months and winsorized at the 1 % and 99 % quantile. Additionally, we include the characteristic labels/indicator dummies Standard errors are estimated according to Driscoll and Kraay (1998). In column (3) we report results according to Equation (**??**): $\hat{r}_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{ht}\hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{jt}\phi_{i,j,t-1} \; \forall \, t = 1, ..., T$. The dependent variable are quarterly returns in percent. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. For the factor beta estimates we apply the EIV correction procedure of Pukthuanthong et al. (2019). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%.

	(1) FMB		(2) Panel		(3) Return	
	Dependent variable:					
	Tilt $T_{i,t}$		Tilt $T_{i,t}$		Return $r_{i,t}^{ex}$	
const	0.706				1.551	
	(36.29)	***			(1.76)	*
mkt	0.090		0.011		-0.282	
	(16.19)	***	(3.17)	***	(-0.66)	
smb	0.022		0.010		0.181	
	(8.14)	***	(3.19)	***	(0.83)	
hml	0.012		0.009		-0.274	
	(4.55)	***	(2.51)	**	(-1.01)	
cma	-0.003		0.003		0.000	
	(-1.52)		(2.01)	**	(-0.03)	***
rmw	0.029		0.015		-0.030	
	(9.80)	***	(9.02)	***	(-0.16)	*
mom	-0.021		-0.001		-0.789	
	(-5.05)	***	(-0.68)		(-1.71)	*
IL^L	0.228		0.238		-0.399	
	(44.21)	***	(21.35)	***	(-1.74)	*
IL^H	-0.179		-0.177		1.320	
	(-17.25)	***	(-8.40)	***	(3.87)	***
PB^L	0.036		0.036		0.743	
	(13.58)	***	(6.53)	***	(2.27)	**
PB^H	-0.065		-0.074		-0.677	
	(-12.72)	***	(-6.77)	***	(-2.06)	**
MO^L	-0.053		-0.046		-0.778	
	(-17.77)	***	(-11.56)	***	(-1.82)	*
MO^H	0.001		0.002		0.982	
	(0.27)		(0.39)		(3.74)	***
VO^L	-0.080		-0.088		-0.014	
	(-21.12)	***	(-12.52)	***	(-0.06)	
VO^H	-0.102		-0.073		-0.583	
	(-20.64)	***	(-8.96)	***	(-0.90)	
BT^L	-0.035		-0.062		-0.165	
	(-12.94)		(-11.05)	***	(0.89)	
BT^H	-0.010		0.011		0.170	
	(-3.24)	***	(1.68)		(0.80)	
SZ^L	-0.269		-0.309		1.440	
	(-86.24)	***	(-58.84)	***	(2.66)	***
SZ^H	0.050		0.071		0.061	
	(4.89)	***	(3.41)	***	(0.20)	
FD^L	0.150		0.161		0.131	
	(33.59)	***	(16.47)	***	(0.53)	
FD^H	-0.073		-0.071		3.383	
	(-12.96)	***	(-6.76)	***	(3.32)	***
IV^L	-0.027		-0.026		0.977	
	(-10.63)	***	(-6.25)	***	(2.52)	**
IV^H	-0.015		-0.016		-0.673	
	(-6.94)	***	(-4.00)	***	(-3.21)	***

Table continued:

	(1)		(2)		(3)	
OP^L	-0.068		-0.075		-0.981	
	(-34.37)	***	(-15.24)	***	(-2.90)	***
OP^H	0.018		0.019		0.756	
	(5.37)	***	(2.73)	***	(3.61)	***

E. Return/Tilt Regression with IV EIV correction

Table 18: In Column 1 we report the results of the panel regression according to Equation (13). In Columns 2 and 3 we report the estimation results of the panel regression according to Equation (14). Coefficients are reported in percent***. Standard errors are estimated according to Driscoll and Kraay (1998). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	(1)		(2)		(3)	
			Dependent variable			
	r_{it}^{ex}		$\hat{r}_{it}^{ex,beta}$		$\hat{r}_{it}^{ex,char}$	
constant	0.046		0.020		-0.018	
	(8.02)	***	(7.87)	***	(-4.87)	***
$\hat{T}_{it}^{beta}$	0.018		0.035		-0.007	
	(-0.20)		(0.88)		(-0.25)	
$\hat{T}_{it}^{char}$	-0.162		-0.009		-0.031	
	(-7.29)	***	(-1.60)		(-2.94)	***
Time-Fixed Effects	Yes		Yes		Yes	
Firm-Fixed Effects	Yes		Yes		Yes	
Within-R^2	17,02%		53.68%		14.68 %	
Observations	373,772		373,772		373,772	

F. Latent Factors Results with IV EIV correction

Table 19: In this table we report the empirical estimation results for the Fama and MacBeth (1973) regressions according to Equation (5): $r_{it}^{ex} = \gamma_{0t} + \sum_{h=1}^{H} \gamma_{h,t} \hat{\beta}_{i,h,t-1} + \sum_{j=1}^{J} \gamma_{j,t} \phi_{i,j,t-1} + \epsilon_{i,t} \ \forall \ t = 1, ..., T$. The dependent variable is the monthly excess market return of individual stocks. The explanatory variables are the estimated historical factor betas and the characteristic labels/indicator dummies. The factor betas are estimated with respect to the five Fama and French (2015) factors augmented with the Carhart (1997) momentum factor. Additionally, we also use the four latent risk factors provided by Kelly et al. (2019) (F1, F2, F3, F4). As described in Section (B.2) we use disjunct sets of odd and even month betas to serve as regressors and instrumental variables in the monthly cross-sectional regressions. In contrast to the results of Table **??** we use factor betas that are estimated from monthly (not daily) returns over a 36-month rolling window. The reported estimates in the table are the time-series averages of the monthly slope-coefficient estimates. The results are reported in percent. The sample period is January 1975 to July 2014. For comparison reasons column *FF Six Factor 2016* provides results for the period January 1975 to June 2016. * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	Dependent variable r_{it}^{ex}						
	(1) **6-Factor-2016**		(2) **6-Factor-2014**			(3) **Latent-4-Factor**	
const	0.774		0.758		*const*	0.746	
	(3.28)	***	(3.11)	***		(3.05)	***
mkt	0.000		0.019		$F1$	0.274	
	(0.00)		(0.31)			(2.36)	**
smb	0.020		0.030		$F2$	0.072	
	(0.61)		(0.89)			(0.70)	
hml	-0.046		-0.045		$F3$	0.059	
	(-1.10)		(-1.03)			(1.48)	
cma	0.045		0.047		$F4$	0.042	
	(1.66)		(1.69)	*		(1.15)	
rmw	0.050		0.060				
	(1.57)		(1.82)	*			
mom	-0.042		-0.047				
	(-1.03)		(-1.13)				
IL^L	-0.091		-0.085		IL^L	-0.083	
	(-1.36)		(-1.24)			(-1.22)	
IL^H	0.290		0.282		IL^H	0.289	
	(3.04)	***	(2.83)	***		(2.90)	***
PB^L	0.218		0.237		PB^L	0.233	
	(3.46)	***	(3.72)	***		(3.63)	***
PB^H	-0.286		-0.313		PB^H	-0.308	
	(-4.08)	***	(-4.38)	***		(-4.30)	***
MO^L	-0.162		-0.144		MO^L	-0.130	
	(-1.65)		(-1.44)			(-1.30)	
MO^H	0.309		0.329		MO^H	0.316	
	(4.44)	***	(4.56)	***		(4.48)	***
VO^L	-0.022		-0.064		VO^L	0.064	
	(-0.260)		(-0.75)			(-0.74)	
VO^H	-0.456		-0.442		VO^H	-0.445	
	(-2.85)	***	(-2.76)	***		(-2.81)	***
BT^L	-0.088		-0.082		BT^L	-0.065	

Table 16 continued:

	(1)		(2)			(3)	
	(-1.57)		(-1.43)			(-1.13)	
BT^H	-0.045		-0.045		BT^H	-0.062	
	(-0.58)		(-0.57)			(-0.78)	
SZ^L	0.458		0.510		SZ^L	0.499	
	(4.40)	***	(4.80)	***		(4.79)	***
SZ^H	-0.084		-0.086		SZ^H	-0.087	
	(-1.14)		(-1.13)			(-1.14)	
FD^L	0.098		0.094		FD^L	0.099	
	(1.64)		(1.49)			(1.60)	
FD^H	0.741		0.807		FD^H	0.807	
	(3.57)	***	(3.79)	***		(3.78)	***
IV^L	0.309		0.273		IV^L	0.262	
	(2.76)	***	(2.38)	***		(2.24)	**
IV^H	-0.183		-0.181		IV^H	-0.183	
	(-2.62)	***	(-2.52)	***		(-2.56)	**
OP^L	-0.251		-0.205		OP^L	-0.198	
	(-3.44)	***	(-2.76)	***		(-2.70)	***
OP^H	0.382		0.410		OP^H	0.404	
	(7.76)	***	(8.18)	***		(8.03)	***
Observations	**498**		**475**			**475**	

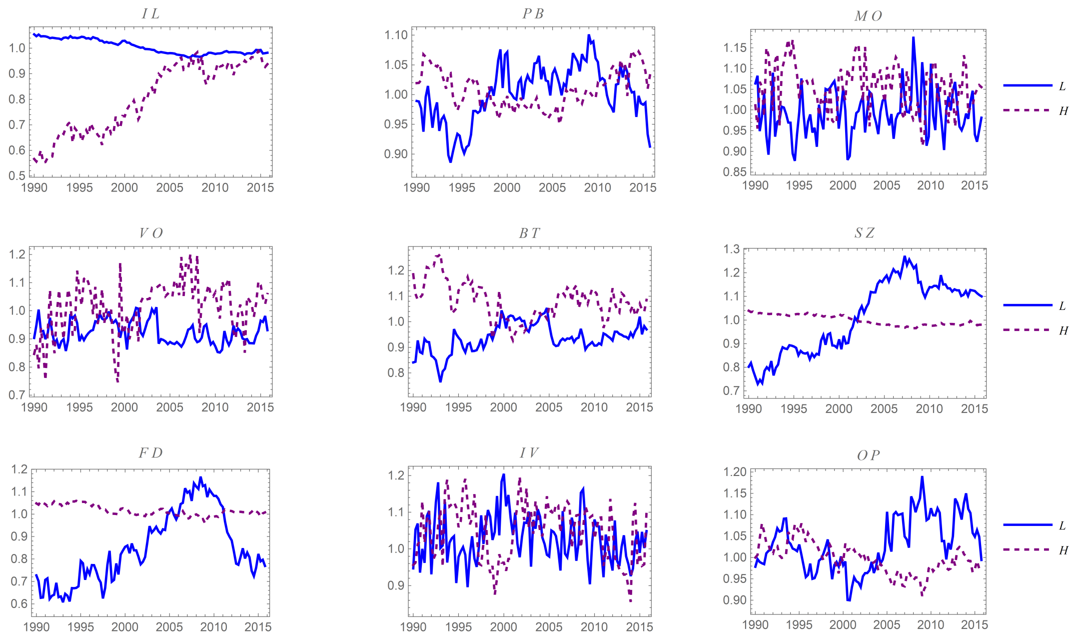
G. Panel Regression Results

Table 20: In this table we report the results of the Hoechle et al. (2020) panel regression model from Equation (16) including firm fixed effects. The dependent variable is the excess stock return expressed in percent. The explanatory variables include the characteristic dummies as introduced in Section (B.3) and the five Fama-French factors market, SMB, HML, CMA, RMW augmented with the Carhart (1997) momentum factor. In column "Jensen-α" we report the marginal contribution of each characteristic to the total Jensen-α of stock i. In the columns β_{MKT}, β_{SMB}, β_{HML}, β_{CMA}, β_{RMW} and β_{MOM} we report the contributions to the corresponding risk-factor loadings for each stock characteristic. In row "Hausman-test" we report the results of the Haumsan(1978)-type specification test as introduced by Hoechle et al. (2020). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%;

Characteristic	Jensen-α		β_{mkt}		β_{smb}		β_{hml}		β_{cma}		β_{rmw}		β_{mom}	
constant	-0.078		0.982		1.024		0.277		-0.038		0.131		-0.243	
	(-0.71)		(33.73)	***	(17.67)	***	(5.31)	***	(-0.46)		(1.72)	*	(-6.38)	***
IL^L	-0.787		0.112		-0.163		-0.135		-0.113		-0.089		-0.016	
	(-7.76)	***	(5.62)	***	(-6.46)		(-3.22)	***	(-1.76)	*	(-1-98)	**	(-0.51)	
IL^H	0.491		-0.129		-0.142		0.048		-0.039		0.171		0.054	
	(3.11)	***	(-3.26)	***	(-2.60)	***	(0.53)		(-0.36)		(2.60)	**	(1.49)	
PB^L	0.595		0.003		-0.007		0.231		0.102		0.014		-0.069	
	(6.40)	***	(0.17)		(-0.25)		(6.55)	***	(1.76)	*	(0.34)		(-3.17)	***
PB^H	-0.641		0.004		0.006		-0.314		-0.028		-0.224		0.123	
	(-7.86)	***	(0.24)		(0.20)		(-7.25)	***	(-0.47)		(-6.15)	***	(6.82)	***
MO^L	0.641		0.039		0.063		-0.138		-0.103		-0.202		-0.400	
	(4.54)	***	(1.15)		(1.39)		(-2.95)	***	(-0.91)		(-4.34)	***	(-6.39)	***
MO^H	-0.085		0.048		0.063		-0.065		0.024		-0.107		0.323	
	(-1.24)		(1.99)	**	(1.84)	*	(-2.04)	**	(0.32)		(-1.68)	*	(9.85)	***
VO^L	0.124		-0.248		-0.271		0.036		0.177		0.086		0.019	
	(1.59)		(-13.99)	***	(-8.73)	***	(0.96)		(3.62)	***	(2.26)	**	(1.10)	
VO^H	0.125		0.144		0.334		0.026		-0.005		-0.600		-0.064	
	(0.66)		(3.12)	***	(3.79)	***	(0.26)		(-0.04)		(-8.52)	***	(-1.25)	
BT^L	-0.076		-0.143		-0.144		0.086		0.116		0.024		0.084	
	(-0.80)		(-6.43)	***	(-4.33)	***	(1.89)	*	(2.41)	**	(0.63)		(4.16)	***
BT^H	-0.380		0.174		0.107		-0.210		0.097		-0.221		-0.157	
	(-2.63)	***	(5.80)	***	(2.13)	**	(-3.88)	***	(1.01)		(-3.65)	***	(-4.59)	***
SZ^L	2.100		-0.186		-0.080		-0.097		0.095		-0.222		0.068	
	(12.55)	***	(-6.86)	***	(-1.68)	*	(-1.38)		(0.92)		(-4.20)	***	(1.57)	
SZ^H	-1.040		0.046		-0.341		0.009		0.140		0.104		-0.023	
	(8.97)	***	(2.34)	**	(-10.57)	***	(0.21)		(2.52)	**	(2.66)	***	(-0.88)	
FD^H	1.802		0.011		0.235		0.081		0.096		-0.464		-0.100	
	(8.20)	***	(0.23)		(2.30)	**	(0.63)		(0.56)		(-4.42)	***	(-1.70)	*
FD^L	-0.104		-0.060		-0.184		-0.062		-0.044		0.118		0.121	
	(-1.44)		(-3.58)	***	(-6.06)	***	(-1.80)	*	(-1.08)		(2.96)	***	(7.80)	***
IV^L	0.940		-0.063		0.068		-0.281		0.295		-0.340		-0.040	
	(8.78)	***	(-1.85)	*	(0.98)		(-2.89)	***	(2.32)	**	(-2.80)	***	(-1.07)	
IV^H	-0.019		0.047		0.059		-0.094		-0.266		0.034		-0.010	
	(-0.23)		(2.81)	***	(1.47)		(-2.47)	**	(-3.65)	***	(0.63)		(-0.49)	
OP^L	0.113		-0.032		0.106		-0.048		-0.158		-0.349		-0.047	
	(0.98)		(-0.96)		(1.92)	*	(-0.94)		(-1.67)	*	(-9.26)	***	(-1.05)	
OP^H	0.273		0.034		0.038		0.084		-0.035		0.359		-0.050	
	(4.77)	***	(2.56)	**	(1.94)	*	(3.31)	***	(-0.83)		(11.95)	***	(-3.61)	***
Observations							1.861.233							
Within R^2							11.66 %							
Hausmann-test							F(126,497)=63.91; Prob>F=0.000							

H. Tilt Persistence – Non-Micro Evidence

Figure 8: In this figure, we plot the time series for aggregate characteristics tilts $T_{j,t}^{P}$ from equation (17). Each panel reports the low (high) manifestation of the characteristic as blue solid (purple dashed) line. Sample excludes microcap stocks, i.e. all stocks below the 20% breakpoint. Quarterly frequency. Time period 1990Q1 – 2015Q4.



I. Return/Tilt Regression – Non-Micro Evidence

Table 22: In Column 1 we report the results of the panel regression according to Equation (13), where we *exclude microcap* stocks. In Columns 2 and 3 we report the estimation results of the panel regression according to Equation (14). Coefficients are reported in percent***. Standard errors are estimated according to Driscoll and Kraay (1998). * denotes significant at 10%, ** denotes significant at 5%, *** denotes significant at 1%, whereby p-values are rounded to two digits.

	(1)		(2)		(3)		(4)	
	Dependent variable							
	r_{it}^{ex}		r_{it}^{ex}		$\hat{r}_{it}^{ex,beta}$		$\hat{r}_{it}^{ex,char}$	
constant	0.130		0.037		-0.022		-0.019	
	(7.17)	***	(3.02)	***	(1.81)	*	(-8.10)	***
$\hat{T}_{i,t-1}$	-0.081							
	(-4.80)	***						
$\hat{T}_{i,t-1}^{beta}$			-0.000		0.035		0.027	
			(0.00)		(1.41)		(1.88)	*
$\hat{T}_{i,t-1}^{char}$			-0.046		0.009		-0.008	
			(-2.15)	**	(0.85)		(0.027)	
Time-Fixed Effects	Yes		Yes		Yes		Yes	
Firm-Fixed Effects	Yes		Yes		Yes		Yes	
Within-R^2	17.97%		22,20%		50.89%		11.76 %	
Observations	171,822		154,939		154,939		154,939	

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Back, K., Kapadia, N., Ostdiek, B., 2013. Slopes as factors: Characteristic pure plays. Working Paper. Rice University.

Back, K., Kapadia, N., Ostdiek, B., 2015. Testing factor models on characteristic and covariance pure plays. Working Paper. Rice University.

Banz, R.W., 1981. The relationship between return and the market value of common stocks. Journal of Financial and Quantitative Analysis 14, 421–441.

Bhandar, L.C., 1988. Debt/equity ratios and expected common stock returns: Empirical evidence. The Journal of Finance 43, 507–528.

Brunnermeier, M., Farhi, E., Koijen, R.S., Krishnamurthy, A., Ludvigson, S.C., Lustig, H., Nagel, S., Piazzesi, M., 2021. Review Article: Perspectives on the Future of Asset Pricing. Review of Financial Studies 34, 2126–2160.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. Is the risk of bankruptcy a systematic risk? The Journal of Finance 63, 2899–2939.

Carhart, M.M., 1997. On Persistence in Mutual Fund Performance. The Journal of Finance 52, 57–82.

Chordia, T., Goyal, A., Shanken, J., 2015. Cross-Sectional Asset Pricing with Individual Stocks: Betas versus Characteristics. Working Paper. University of Atlanta and University of Lausanne.

Cochrane, J.H., 2005. Asset Pricing. volume 2. Princeton University Press, Princeton.

Cochrane, J.H., 2011. Presidential Address: Discount Rates. The Journal of Finance 66, 1047–1108.

Daniel, K., Titman, S., 1997. Evidence on the Characteristic of Cross Sectional Variation in Stock Return. Journal of Finance 52, 1–33.

Davis, J.L., Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929 to 1997. The Journal of Finance 55, 389–406.

Dimson, E., Mussavian, M., 1999. Three centuries of asset pricing. Journal of Banking & Finance 23, 1745–1769.

Driscoll, J.C., Kraay, A.C., 1998. Consistent covaraince matrix estimation with spatially dependent panel data. Review of Economics and Statistics 80, 549–560.

Edelen, R.M., Ince, O.S., Kadlec, G.B., 2016. Institutional Investors and stock return anomalies. Journal of Financial Economics 119, 472–488.

Fama, E., MacBeth, J.D., 1973. Risk, return and equilibrium: empirical tests . Journal of Political Economy 81, 607–636.

Fama, E.F., 1976. Foundations of Finance. Basic Books, New York.

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

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E.F., French, K.R., 2020. Comparing Cross-Section and Time-Series Factor Models. The Review of Financial Studies 33, 1891–1926.

Feng, G., Giglio, S., Xiu, D., 2020. Taming the Factor Zoo: A Test of New Factors. Journal of Finance LXXV, 1327–1370.

Frazzini, A., Pedersen, L.H., 2014. Betting against beta. Journal of Financial Economics 111, 1–25.

Green, J., Hand, J.R.M., Zhang, X.F., 2017. The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. The Review of Financial Studies , 1–48.

Harvey, C.R., Liu, Y., 2021. Lucky factors. Journal of Financial Economics 141, 413–435.

Harvey, C.R., Liu, Y., Zhu, H., 2016. ...and the Cross-Section of Expected Returns. The Review of Financial Studies 29, 5–68.

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.

Hirshleifer, D., 2020. Presidential address: Social transmission bias in economics and finance. The Journal of Finance 75, 1779–1831.

Hoechle, D., Schmid, M., Zimmermann, H., 2020. Does Unobservable Heterogeneity Matter for Portfolio-Based Asset Pricing Tests? Working Paper. University of Basel, FHNW School of Business Basel and University of St. Gallen.

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., Wang, J., 2019. Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium. Journal of Financial Economics 133, 273–298.

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134, 501–524.

Kim, S., Korajczyk, R.A., Neuhierl, A., 2021. Arbitrage Portfolios. Review of Financial Studies 34, 2813–2856.

Koijen, R.S.J., Yogo, M., 2019. A Demand System Approach to Asset Pricing. Journal of Political Economy 127, 1475–1515.

Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting Factor Models. The Journal of Finance 73, 1183–1223.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1992. The Structure and Performance of the Money Management Industry. Journal of Financial Markets 5, 31–56.

Light, N., Maslov, D., Rytchkov, O., 2017. Aggregation of Information About the Cross Section of Stock Returns: A Latent Variable Approach . Review of Financial Studies 30, 1339–1381.

Pastor, L., Stambaugh, R.F., Taylor, L.A., 2020. Sustainable investing in equilibrium. Journal of Financial Economics , 1–53.

Pukthuanthong, K., Roll, R., Subrahmanyam, A., 2019. A Protocol for Factor Identification. The Review of Financial Studies 32, 1573–1607.

Raponi, V., Robotti, C., Zaffaroni, P., 2020. Testing Beta-Pricing Models Using Large Cross-Sections. Review of Financial Studies 33, 2796–2842.

Ross, S.A., 1976. The Arbitrage Theory of Capital Asset Pricing. Journal of Economic Theory 13, 341–360.

Shiller, R.J., 2017. Narrative economics. American Economic Review 107, 967–1004.

Stattman, D., 1980. Book values and stock returns. The Chicago MBA: A Journal of Selected Papers 4, 25–45.