A One-factor Model for Expected Night-minus-day Stock Returns*

Zhongjin (Gene) Lu[†] and Zhongling (Danny) Qin[‡]

Abstract

We propose a one-factor model that summarizes the cross-sectional variation in expected night-minus-day (NMD) stock returns. Our constructed pricing factor – the NMD returns on a high Sharpe ratio NMD trading strategy – has substantial exposure to the dominant common risks in the NMD return space, consistent with the absence of near-arbitrage opportunities. Furthermore, the factor premium is related to retail trading demand at the market open and the required returns from liquidity provision. Our findings are consistent with an economic equilibrium in which liquidity providers require compensation for accommodating sentiment-driven demand.

Keywords: Night-minus-day returns, liquidity provision, limits-to-arbitrage, retail trading
JEL Classifications: G12, G14, G23

*We thank Jules van Binsbergen, Vincent Bogousslavsky, Bob Hodrick, Steven Malliaris, Bradley Paye, Paul Tetlock, Baolian Wang, Pengfei Ye, Dexin Zhou, and seminar participants at Dongbei University of Finance and Economics, Virginia Tech, Auburn University, and the University of Georgia for helpful discussions.

†Assistant Professor of Finance, Terry College of Business, University of Georgia; corresponding author, author email: zlu15@uga.edu

‡Assistant Professor of Finance, Harbert College of Business, Auburn University, author email: zzq0018@auburn.edu

Stock return predictability is a subject of intense interest for finance practitioners and academics. In a recent literature, researchers document more than a dozen characteristics that predict the cross-section of night-minus-day (NMD) stock returns, defined as the difference between a stock's overnight and intraday returns (e.g., Berkman, Koch, Tuttle, and Zhang (2012), Lou, Polk, and Skouras (2019), and Hendershott, Livdan, and Rösch (2020)). Moreover, the predictive power of these stock characteristics for NMD returns cannot be explained by standard factor models (e.g., Fama and French (2015); Hou et al. (2015); Stambaugh and Yuan (2017); Daniel et al. (2020).[1] As a first step towards understanding the pricing of these expected NMD returns, we propose a parsimonious one-factor model that summarizesthe cross-sectional variation in expected NMD returns. We then use our proposed NMD pricing factor to examine the economic sources of these NMD return predictabilities.

To achieve our first objective of developing a factor model for explaining NMD returns[2], we start by spanning the NMD return space using long-short zero-investment portfolios sorted on the 17 stock characteristics examined by Lou, Polk, and Skouras (2019) (LPS).[3] We then utilize two approaches to construct the NMD pricing factor. First, we implement the Bayesian estimator of the stochastic discount factor (SDF) proposed by Kozak, Nagel, and Santosh (2020), which is shown to guard against the in-sample overfitting associated with the conventional SDF estimator and performs well in out-of-sample pricing tests. We hereafter refer to this specification of the pricing factor as NMD SDF. Second, in the spirit of Stambaugh and Yuan (2017), we combine the predictive information for NMD returns from all 17 LPS signals into a composite characteristic-based score (CS) using the regression approach used in Lewellen (2015). Following the long-time standard approach of constructing characteristic-based factors, we construct the long-short portfolio sorted on CS (hereafter, NMD CS) and use it an alternative specification of the pricing factor.[4] The

[1]See Table 1.

[2]Since our goal is to understand deviations from the capital asset pricing model (CAPM) in the cross-section of expected NMD returns, unless otherwise stated, all returns hereafter are market-adjusted returns. See Section 1 for empirical details.

[3]Lou, Polk, and Skouras (2019) is the most comprehensive study of the cross-section of NMD returns so far in the literature. Other work include Hendershott, Livdan, and Rösch (2020) which extensively study the relation between the NMD returns and the market beta, Berkman, Koch, Tuttle, and Zhang (2012) which look into the relation between the NMD returns and the squared stock returns, and Bogousslavsky (2021) which documents the relation between the NMD returns and the anomaly characteristics used in Stambaugh and Yuan (2017).

[4]The long-short portfolio sorted on CS can also be interpreted as an approximation of the SDF if all stocks equally contribute to the NMD return covariance matrix. See Section 2 for more detailed discussions.

virtue of NMD CS is that it is constructed using only prevailing information. These two specifications represent two distinct methodologies in a large and growing literature on factor modeling: the first specification imposes the asset pricing restriction linking expected NMD returns to NMD return covariance, whereas the second specification does not. Our approach of using these two specifications is to ensure that our key takeaways are robust to reasonable variations in methodology.

Next, we examine the ability of our proposed NMD pricing factor for explaining the cross-sectional variation in the average NMD returns of the 17 LPS portfolios. Consistent with the notion that the SDF is the maximum sharpe ratio portfolio in the payoff space, we find that NMD SDF and NMD CS have Sharpe ratios of 5.0 and 4.4, respectively, which is higher than that of any of the LPS portfolios. Moreover, the average absolute alpha and the corresponding average absolute t-statistic under our one-factor model based on NMD SDF (NMD CS) is 76.6% and 80.0% (60.6% and 67.2%) smaller than the corresponding quantities under CAPM. The one-factor model based on NMD SDF (NMD CS) explains 94% (84%) of the cross-sectional variation in average NMD returns. In contrast, the corresponding cross-sectional R^2 under the standard factor models used in the anomaly literature range from 6% to 11%. Panel A of Figure 1 illustrates that the average NMD returns align with the predictions of our one-factor model better than those of the standard factor models.

To further validate the pricing performance of our proposed NMD pricing factor, we conduct an out-of-sample (OOS) test by augmenting the test asset set with the long-short portfolios sorted on 80 anomaly characteristics in Green, Hand, and Zhang (2017) that have not been examined by LPS. These 80 additional long-short portfolios introduce independent cross-sectional and time-series return variations relative to the LPS portfolios, resulting in a more stringent pricing test (Lewellen, Nagel, and Shanken (2010)). Using the same one-factor model, we find that NMD SDF (NMD CS) similarly captures much of the cross-sectional variation in average NMD returns in this augmented set of 97 long-short portfolios, achieving a cross-sectional R^2 of 79% (78%). Panel B of Figure 1 illustrates that our proposed one-factor model substantially outperforms standard factor models in pricing this augmented set of NMD returns.

After demonstrating the good empirical performance of our proposed NMD pricing factor, our second objective is to use it to understand the economic sources of NMD return predictabilities. Our first set of tests focus on the relation between the NMD pricing factor and the common risk factors in the NMD return space. These tests are motivated by the insight in Kozak, Nagel, and Santosh (2018) that given substantial cross-sectional variation in expected returns, either the return space has a weak factor structure or the SDF should load heavily on the first few PCs. Since this prediction is predicated on the absence of near-arbitrage opportunities, if we find the opposite in the data, i.e., there is a strong factor structure but small SDF loadings on the common risk factors, then either near-arbitrage opportunities exist temporarily in the sample or NMD return predictabilities are a statistical fluke.

We characterize the factor structure of the NMD return space by performing a principal component analysis (PCA) of the NMD returns of the 17 LPS portfolios. We find that the first five principal components (PCs) have above-average explanatory power for the total variance of the NMD return space. The first five PCs account for 26%, 22%, 13%, 7% and 6% of the common variation in NMD returns, respectively, and the first three PCs explain a majority (61%) of the total NMD return variance. Consistent with the prediction in Kozak, Nagel, and Santosh (2018), we find that both NMD SDF and NMD CS have significant and economically large exposures to the first two PCs in the NMD return space. These first two PCs together explain 64.6% and 64.2% of the variation in NMD SDF and NMD CS, respectively, and a factor model based on the first two PCs explains 76% of the cross-sectional variation in the average NMD returns of 17 LPS portfolios. Therefore, to earn the premium of the NMD pricing factor (i.e., a proxy for the mean-variance efficient NMD portfolio), one has to take on the common risk in the NMD return space. Given that there is no reason to expect a data-mined NMD pricing factor (with a spuriously high in-sample Sharpe ratio) to have significant exposures to common risk factors, our results suggest that data mining or publication bias is unlikely to completely account for NMD return predictabilities.

Next, we conduct direct tests to link the NMD pricing factor to sentiment-driven demand and the required compensation for arbitrageurs to accommodate such demands.[5] Related to the former, existing studies such as Berkman, Koch, Tuttle, and Zhang (2012), Lou, Polk, and Skouras (2019), and Hendershott, Livdan, and Rösch (2020) emphasize the role of sentiment-driven demand in driving NMD return predictabilities.[6] Complementing these studies, we provide new evidence on the timing and identity of the sentiment-driven demand.

[5]As highlighted by Kozak, Nagel, and Santosh (2018), the existence of a common factor model can reflect both these economic sources. However, our finding that a NMD pricing factor loads heavily on common risk factors cannot be interpreted as evidence in support of rational asset pricing models relative to sentiment-driven asset pricing models.

[6]Berkman et al. (2012) attribute their findings to retail trading, Hendershott, Livdan, and Rösch (2020) conjecture that risk-loving speculators play a role, and Lou, Polk, and Skouras (2019) attribute theirs to a tug-of-war between opposing trading demands of different clienteles, which can include trading demands from both sentiment investors and rational arbitrageurs.

We find that stock prices at the market open is a more important driver of the NMD return deviations from CAPM than prices at the market close. Using a comprehensive dataset on retail order imbalances, constructed following the methodology in Boehmer, Jones, Zhang, and Zhang (2020), we also find that retail investors, instead of non-retail investors, trade aggressively in the same direction as these deviations from CAPM at the market open.

Related to the role of required compensation for arbitrageurs, we also put forth a novel hypothesis linking NMD return predictabilities to the required returns from liquidity provision. Our liquidity provision hypothesis is motivated by two key observations. First, exploiting NMD return predictabilities requires high-turnover trading strategies, which are not profitable if the implementation involves paying the bid-ask spreads. Therefore, the most plausible arbitrageur of NMD return predictabilities is the market-making sector, which as a whole does not pay the bid-ask spread. Second, since we find that retail investors' trading demand gives rise to NMD return predictabilities, market makers are in prime situations to exploit NMD return predictabilities by accommodating the trading demand from these retail investors.[7] For these two reasons, we posit that the conditional expected returns of the NMD pricing factor is closely related to the required returns from liquidity provision.

To test our liquidity provision hypothesis, we use the empirical proxy for the required returns from liquidity provision proposed by Nagel (2012). This liquidity provision factor is the return of a short-term reversal strategy sorted on 1 to 5-day lagged daily returns. We find that both NMD SDF and NMD CS load positively and statistically significantly on the liquidity provision factor, with a R^2 of 34.5% and 27.2%, respectively.[8] Furthermore, the liquidity provision factor explains more than 80% of the average returns of the NMD pricing factor. Our results therefore indicate that when the required returns from liquidity provision is high, the NMD pricing factor premium is also high and vice versa. Finally, when we use both the liquidity provision factor and our order imbalance proxy for sentiment-driven demand together to explain the NMD pricing factor, we find both have independent explanatory power.

Our paper builds on the existing literature that documents a larger number of predictors for the NMD returns (e.g., Berkman et al. (2012), Lou, Polk, and Skouras (2019), Hendershott, Livdan, and Rösch (2020)), and Bogousslavsky (2021).[9] We contribute to this

[7]As we discuss in Section 4.2, market makers have the necessary investments in both high-speed market access and the acquisition of uninformed order flows to do so.

[8]In a placebo test, we further find that the NMD pricing factor is uncorrelated with the reversal factor provided by Ken French, which is sorted on past 21-day returns and thus is not a good proxy for the returns from liquidity provision. See Section 4 for additional discussion.

[9]Heston et al. (2010); Bogousslavsky (2021) also document predictability in the cross-section of half-hour

literature by being the first to propose a parsimonious factor model for explaining the cross-section of expected NMD returns, which is a key step towards a common explanation for NMD return predictabilities. We also show that the NMD return space spanned by the 17 LPS portfolios has a strong factor structure and that our proposed NMD pricing factor is tied to the common risk in NMD returns. As a result, NMD return predictabilities are unlikely to be a result of the data mining or publication bias. Finally, we document new evidence linking the NMD pricing factor to both sentiment-driven demand and the required returns from liquidity provision, suggesting that NMD return predictabilities are consistent with the economic equilibrium described by Kozak, Nagel, and Santosh (2018).

1. Data and Measurement

1.1. Sample Construction

We start by collecting data from the Center for Research in Security Prices (CRSP) database for all U.S. common stocks listed on the NYSE, AMEX, and NASDAQ stock exchanges. Because we are interested in overnight and intraday stock returns, we require the daily open price from CRSP, which is available starting on June 15, 1992. We then merge the CRSP data with the NYSE Trade and Quote (TAQ) database using the TAQ-CRSP link table provided by Wharton Research Data Services (WRDS). Since TAQ data starts on January 4, 1993, our analyses requiring TAQ data are conducted over the period between 1993 and 2020. See the Appendix Subsections 6.2 and 6.3 for a more detailed description of the TAQ data and the procedure used to merge the TAQ data with the CRSP data.

We impose the following data filters. First, like Hendershott, Livdan, and Rösch (2020), we drop stock days with an intraday return over 1000% or when the open price is missing. Second, to to mitigate the microstructure issues and ensure that results are not driven by small and illiquid stocks, we require the following for a stock to be included in the portfolios formed at the end of month t: (i) the median daily trading volume in month t is greater than 1,000 shares, (ii) the stock has no more than one missing open price from CRSP in month t, and (iii) following Lou, Polk, and Skouras (2019), stocks need to have a price above $5 and a market capitalization above the NYSE bottom quintile at the end of month t. Since all our portfolio returns are value-weighted, our results are robust to removing the low price filter and the microcap filter.

---

returns.

1.2. Measurement of NMD Returns

We follow Lou, Polk, and Skouras (2019) and compute the intraday return ($r_{D,d}$) on day d as,

$$r_{D,d} = \frac{P_d^{\text{close}}}{P_d^{\text{open}}} - 1, \tag{1}$$

and the overnight return ($r_{N,d}$) from the close of day $d-1$ to the open of day d as,

$$r_{N,d} = \frac{1 + r_{\text{close-to-close},d}}{1 + r_{D,d}} - 1, \tag{2}$$

where $r_{\text{close-to-close},d}$ is the close-to-close return on day d by CRSP.[10] The NMD return is then,

$$r_{NMD,d} = r_{N,d} - r_{D,d}. \tag{3}$$

Both P_d^{close} in Eq. (1) and $r_{\text{close-to-close},d}$ in Eq. (2) are sourced from CRSP, following the convention in the literature. Our main specification of P_d^{open} is the open price from CRSP, similar to Hendershott, Livdan, and Rösch (2020), so that researchers who do not have access to TAQ data can still replicate our results. Other studies in the literature have used alternative specifications of P_d^{open} by relying on TAQ data. For example, Berkman et al. (2012) use the first midquote, Bogousslavsky (2021) the midquote at 9:45 am, and Lou, Polk, and Skouras (2019) the volume-weighted average price between 9:30 am and 10:00 am, all from TAQ. The pricing performance of our proposed one-factor model and other key takeaways from this article are robust to using the specifications from Berkman et al. (2012), Bogousslavsky (2021), or Lou, Polk, and Skouras (2019). Our Appendix reports all our results using Lou, Polk, and Skouras (2019)'s specification of P_d^{open} and the results based on the other two specifications are available upon request.

1.3. Construction of the Test Assets

Our main set of test assets consists of the long-short portfolios sorted on the 17 anomaly characteristics examined by LPS. We use the NMD returns of these 17 LPS portfolios to span

[10] Eq. (1) assumes that stock and cash dividends, as well as share splits, occur overnight.

the payoff space.[11] In our out-of-sample (OOS) test, we augment these 17 LPS portfolios with the long-short portfolios sorted on additional 80 anomaly characteristics from Green, Hand, and Zhang (2017) that are not examined by LPS.[12]

To construct these long-short portfolios, following LPS, we sort all stocks into decile portfolios based on an ascending ordering of each of these signals at the end of each month t. The long-short zero investment portfolio goes long the top decile portfolio and short the bottom decile portfolio. We then calculate the daily value-weighted overnight and intraday portfolio returns realized in month $t + 1$ with the prior day's market capitalization as the weights. The daily NMD returns of the long-short portfolio is the return on a trading strategy that goes long this portfolio overnight and short it intraday. Since our focus is on the ability of a factor model for explaining the deviations from the CAPM model, we adjust all portfolio returns by subtracting the CAPM beta times the corresponding market excess returns, respectively.[13] In all subsequent analysis, NMD returns refer to market-adjusted NMD returns.

Table 1 reports the time-series average of the unadjusted NMD returns to our long-short portfolios as well as their average risk-adjusted return (i.e. alphas) relative to standard factor models used in the literature. Consistent with the findings in LPS, we observe that these portfolios have average unadjusted NMD returns ranging from -26.0% to 69.6% per year, with 13 of them having statistically significant market-adjusted NMD returns at the 1% level. Furthermore, we find that these large cross-sectional variations in average NMD returns cannot be explained by the risk adjustments under the five-factor model of Fama and French (2016) (FF5) , the q-factor model of Hou, Xue, and Zhang (2015) (HXZ4), the mispricing model of Stambaugh and Yuan (2017) (SY4), or the behavioral model of Daniel, Hirshleifer, and Sun (2020) (DHS3). We now turn to the construction of our proposed NMD pricing factor.

[11] See the Appendix Subsection 6.1 for detailed descriptions on the construction of these signals.

[12] Among these additional characteristics from Green, Hand, and Zhang (2017), we exclude the small number of binary ones that do not work for forming our portfolios. We thank Green, Hand, and Zhang (2017) for providing the SAS code used to generate these characteristics.

[13] We estimate CAPM beta by regressing the daily close-to-close returns of the long-short portfolio on daily market excess returns in the full sample. Our results are robust to using CAPM betas estimated in rolling estimation windows. Following Heston, Korajczyk, and Sadka (2010), we assume the risk-free rate is earned overnight. So, more precisely, the market-adjusted intraday return is the portfolio intraday return minus the CAPM beta times the market intraday return, whereas the market-adjusted overnight return is the portfolio overnight return in excess of the risk-free rate minus the CAPM beta times the market overnight return in excess of the risk-free rate.

2. A One-factor Model for the Cross-section of Expected NMD Returns

2.1. Construction of the NMD Pricing Factor

In this section, we lay out the motivation and the methodology for constructing our proposed NMD pricing factor. Consider an economy with N test assets. Denote the vector of the (market-adjusted) NMD excess returns for these assets by $R = (R_1, ..., R_N)$ and the covariance matrix of these NMD returns by Γ. Define $\mu \equiv E(R)$. In the payoff space formed by any linear combination of NMD returns, the law of one price implies the existence of a stochastic discount factor (SDF) as follows (Hansen and Jagannathan (1991)),

$$\text{SDF} = 1 - b' (R - \mu), \tag{4}$$

where the SDF coefficients are the weights of the mean-variance efficient portfolio (i.e., $b = \Gamma^{-1}\mu$.) A naive approach for estimating b is to plug in the sample counterparts of Γ and μ, but it is well known that this naive estimator performs poorly out-of-sample due to the uncertainty in the estimated return covariance and means.

Therefore, a large and growing literature since the seminal work of Fama and French (1993) has alternatively used three general approaches to construct the pricing factors. One approach is to construct factors based on behavioral and asset pricing theories (Fama and French (2015); Hou et al. (2015); Daniel et al. (2020)). However, so far there is no formal theory for explaining NMD return predictabilities. In fact, one of our goals of developing a parsimonious one-factor model is to inform future research on building such a theory. A second approach imposes the asset pricing restriction that links expected returns to return covariances in order to extract pricing factors from portfolio returns. The Kozak, Nagel, and Santosh (2020) estimator that we use in our first specification of the NMD pricing factor is representative of this approach.[14] A third approach constructs portfolios sorted on return predicting characteristics as factors. Our second specification of the NMD pricing factor is more in line with this approach, which is commonly used in the anomaly literature. Our goal of using these two specifications is to ensure that our key takeaways are robust to reasonable

[14] See Kozak, Nagel, and Santosh (2020) for a discussion of this literature. Our results are also robust to using the approach proposed by Lettau and Pelger (2020).

variations in methodology.

2.1.1. The First Specification of the NMD Pricing Factor

We use the Bayesian SDF estimator proposed by in Kozak, Nagel, and Santosh (2020) (KNS) to construct the first specification of the NMD pricing factor. With an economically motivated prior, the KNS estimator of b in Eq. (4) resembles a ridge regression estimate with a L^2 norm penalty term,

$$\hat{b} = \left(\overline{\Gamma} + \gamma I\right)^{-1} \overline{\mu}, \tag{5}$$

where I is the identity matrix, $\overline{\Gamma}$ and $\overline{\mu}$ are the estimated return covariance matrix and the average returns of the test assets, respectively, and γ is the hyperparameter associated with the L^2 penalty term. As Kozak et al. (2020) explain, this KNS estimator shrinks the SDF coefficients of the naive estimator towards zero, with the shrinkage factor being stronger for the coefficients on the principal components with smaller variance.[15]

Our implementation of the KNS estimator is as follows. We denote the NMD returns for the 17 LPS portfolios by F_t. With a time period of length T, we estimate the sample moments by,

$$\overline{\mu} = \frac{1}{T}\sum_{t=1}^{T} F_t \tag{6}$$

$$\overline{\Gamma} = \frac{1}{T}\sum_{t=1}^{T} \left(F_t - \overline{\mu}\right)\left(F_t - \overline{\mu}\right)' \tag{7}$$

To choose the optimal γ, we follow KNS in using K-fold cross-validation (CV) with $K = 3$. We first equally divide our sample into 3 subsamples and then set a grid of potential values for γ. For a given γ value, we use $K-1$ subsamples to estimate the in-sample moments $\overline{\mu}_{IS}$ and $\overline{\Gamma}_{IS}$, according to Eqs. (6) and (7), and $\hat{b}_{IS} = \left(\overline{\Gamma}_{IS} + \gamma I\right)^{-1} \overline{\mu}_{IS}$. Then, using the withheld subsample, we compute the OOS moments, $\overline{\mu}_{OOS}$ and $\overline{\Gamma}_{OOS}$. Finally, we compute the out-of-sample R^2 as,

[15] Among several alternatives that KNS explore, they state that this estimator is the natural starting point for applications of their approach if sparsity is not required.

$$R^2_{OOS} = 1 - \frac{\left(\overline{\mu_{OOS}} - \overline{\sum}_{OOS}\hat{b}_{IS}\right)' \left(\overline{\mu_{OOS}} - \overline{\sum}_{OOS}\hat{b}_{IS}\right)}{\overline{\mu_{OOS}}'\overline{\mu_{OOS}}}.$$

We withhold each of the K subsamples, treat it as OOS data, and repeat the above procedure K times. The cross-validated R^2 is the average R^2_{OOS} across these K estimates for a given γ. Then, we select the optimal γ that maximizes the cross-validated R^2. With the optimal γ, we compute the SDF coefficient b^* using the full-sample moments according to Eqs. (5) – (7). Finally, we use b^* to construct the one-dollar long and one-dollar short zero investment portfolio as the NMD pricing factor. This pricing factor is then the linear combination of the 17 LPS long-short portfolios with the following weight on each portfolio i,

$$w_i = \frac{b_i^*}{\sum_{i=1}^{17} |b_i^*|}. \tag{8}$$

We use the NMD returns of this long-short portfolio as our first specification of the NMD pricing factor (NMD SDF). We note that the construction of NMD SDF uses full-sample information. KNS also conduct a pure OOS test, in which they withhold part of the data when estimating the optimal b^* and show that the resulting SDF estimator performs well in the withheld sample. Since our sample period is shorter than that in Kozak, Nagel, and Santosh (2020), we instead evaluate OOS performance by expanding the set of test assets in our cross-sectional pricing test. Furthermore, our second specification of the pricing factor requires only prevailing information in its construction.

2.1.2. The Second Specification of the NMD Pricing Factor

To construct our second specification of the pricing factor, we first combine all 17 LPS characteristics into a composite predictive signal of a stock's NMD return. This procedure is similar in spirit to Stambaugh, Yu, and Yuan (2015), who construct a composite mispricing score based on 11 anomaly characteristics. Instead of equally weighting each characteristic, we follow Lewellen (2015) and use a Fama and MacBeth (1973) regression approach to let the data decide on how much weight to put on each individual characteristic. Specifically, at the end of each month t, we run a multivariate cross-sectional regression of the NMD return on lagged LPS characteristics as follows,

$$\text{NMD}_{i,t} = \alpha + \beta_t^{'} \times x_{i,t-1} + \epsilon_{i,t}, \ i = 1, ..., N, \tag{9}$$

where $\text{NMD}_{i,t}$ is the average unadjusted NMD return for firm i in month t and $x_{i,t-1}$ is the vector of firm i's characteristics in month $t-1$. Similar to Freyberger et al. (2020); Kozak et al. (2020), to minimize the influence of extreme values, $x_{i,t}^s$ is the normalized cross-sectional rank of a characteristic $c_{i,t}^s$,

$$x_{i,t}^s = \frac{\text{rank}(c_{i,t}^s) - \overline{\text{rank}(c_{i,t}^s)}}{n_t + 1}, \tag{10}$$

where the scaling factor n_t is the number of stocks in month t.[16] We then construct the composite signal for firm i in month t using 12-month moving averages of β_t (i.e., $\beta_{t,MA} = \frac{1}{12}\sum_{L=0}^{11}\beta_{t-L}$) and $x_{i,t}$:

$$CS_{i,t} = \beta_{t,MA}^{'} \times x_{i,t}. \tag{11}$$

Similar to the construction the 17 LPS portfolios, we then construct a one-dollar long-short zero-investment portfolio sorted on CS. We use the NMD returns on this long-short portfolio as our second specification of the NMD pricing factor (NMD CS). The portfolio weights of NMD CS can be interpreted as a proxy for b in Eq. (4) under the assumption that CS is proportional to μ and all stocks equally contribute to Γ.

2.2. Pricing Performance of Factor Models

We now examine the performance of our proposed NMD pricing factor via its ability to explain the average NMD returns of the 17 LPS portfolios. In Table 2, we first report the summary statistics for NMD SDF and NMD CS. We observe that NMD SDF and NMD CS have average returns of 44.9% and 94.7% per annum (p.a.), respectively, and annualized Sharpe ratios of 5.0 and 4.4, respectively. In comparison, as shown in Figure 2, the Sharpe ratios of the 17 individual LPS portfolios range from 0.1 to 3.4. Therefore, consistent with the notion that our proposed pricing factors are a proxy for the SDF (the maximum Sharpe ratio portfolio), their Sharpe ratios are higher than that of the individual assets that form the payoff space.

Next, we run a time-series regression of the daily NMD returns of the LPS portfolios

[16] We use the average rank if there is a tie in the characteristic value across stocks. If a characteristic value is missing, we use zero to fill it in.

on NMD SDF and NMD CS, respectively. Column (1) of Panel A of Table 3 reports that the cross-sectional mean of the absolute alpha relative to the CAPM, FF5, HXZ4, SY4, and DHS3 factor models are 31.4%, 29.5%, 30.3%, 33.8%, and 29.8% p.a., respectively. In contrast, the cross-sectional mean of the absolute alpha relative to a one-factor model based on NMD SDF (NMD CS) is much smaller, at 7.3% (12.4%) p.a. Similarly, column (2) shows that the cross-sectional mean of the absolute $t-$statistic under the CAPM, FF5, HXZ4, SY4, and DHS3 factor models are all above 6, whereas the cross-sectional mean of the absolute $t-$statistic of the alpha relative to a one-factor model based on NMD SDF (NMD CS) is only 1.26 (2.06).

Following KNS, we use cross-sectional R^2 as the main measure of the relative performance across the competing factor models for explaining cross-sectional variation in average NMD returns. Specifically, we compute it as,

$$R^{2,\text{XS}} = 1 - \frac{\left(\overline{\mu} - \beta'\overline{F}\right)'\left(\overline{\mu} - \beta'\overline{F}\right)}{\overline{\mu}'\overline{\mu}}, \tag{12}$$

where $\overline{\mu}$ is the vector of the sample means of each test asset's NMD return, β is the matrix of the factor loadings from a time series regression of test asset's NMD return on the factor(s), and $\overline{F}$ is the vector of the sample mean of each factor.

Column (3) of Panel A of Table 3 reports that the cross-sectional R^2 under CAPM is zero because all NMD returns are market-adjusted. The cross-sectional R^2s using the standard factor models range from 6% to 11%. In contrast, a one-factor model based on NMD SDF (NMD CS) achieves a substantially higher cross-sectional R^2 of 94% (84%). In Panel A of Figure 1, we plot the average NMD returns against the predicted NMD returns by these factor models. In the case of CAPM, no cross-sectional explanatory power results in a vertical line. Using standard factor models does not change the vertical pattern by much. In comparison, using the predicted returns under a one-factor model based on either NMD SDF or NMD CS result in a much better alignment along the 45 degree line.

2.3. Out-of-sample Test with an Expanded Set of Test Assets

To further validate the pricing performance of our proposed NMD pricing factor, we conduct an OOS test by augmenting the test asset set with the long-short portfolios sorted on 80 anomaly characteristics examined by Green, Hand, and Zhang (2017) that have not been examined by LPS. Since LPS do not pick their set of 17 anomaly characteristics based on the ability to predict NMD returns,[17] we expect some of these additional anomaly characteristics predict NMD returns. Consistent with this point, in unreported results we find that the average NMD return is significantly different from zero at the 5% level for 58 out of the additional 80 long-short portfolios. Therefore, conducting this test not only examines the OOS fit of the factor models but also introduces independent cross-sectional and time-series variation in the NMD returns relative to that of the LPS portfolios, resulting in a more stringent pricing hurdle (Lewellen, Nagel, and Shanken (2010)).

Panel B of Table 3 reports the pricing results for this cross-section of 97 test assets. We find that the cross-sectional mean of the absolute CAPM alpha is 19.4% p.a. and the corresponding mean absolute $t-$statistic is 4.5. Both numbers are lower in this expanded set compared to that of the 17 LPS portfolios, because 22 out of the additional 80 portfolios do not have statistically significant average NMD returns. Adjusting the NMD returns using standard factor models again helps little to reduce the absolute alphas, with cross-sectional R^2s ranging from 14% to 19%. In contrast, our one-factor model based on NMD SDF (NMD CS) achieves a cross-sectional R^2 of 79% (78%). Panel B of Figure 1 illustrates the large improvement in the pricing performance of our one-factor model relative to the standard factors models in this expanded set of test assets.

3. Relation between the NMD Pricing Factors and the NMD Common Risk Factors

After establishing that the NMD pricing factor summarizes the cross-section of expected NMD returns, we subsequently use it to understand the economic sources underlying NMD return predictabilities. In this section, we explore the relation between the NMD pricing factor and the common risk factors in the NMD return space. This test is motivated by the insight in Kozak, Nagel, and Santosh (2018) that given large cross-sectional spreads in expected return, we should expect that either the return space has a weak factor structure or the SDF loads heavily on the first few dominant principal components (PCs). This prediction is based on the absence of near-arbitrage opportunities argument that, given a strong factor structure, expected returns should align with the dominant sources of return variance as otherwise extremely high Sharpe ratio (near-arbitrage) opportunities would exist.

Therefore, we first characterize the factor structure of the NMD return space. We perform

[17] LPS state that their motivation for examining these 17 anomaly characteristics is that they are related to “popular trading strategies”.

a principal component analysis (PCA) using the NMD returns of the 17 LPS portfolios, thereby retrieving the ordered PCs from the largest eigenvalue to the smallest.[18] In Table 4, we find that the NMD return space exhibits a strong factor structure. The first five PCs explain 26%, 22%, 13%, 7% and 6% of the total variance of the 17 LPS portfolios, respectively, and the first three PCs together explain a majority (61%) of the total variance.

Given this strong factor structure, we can now move to test whether our NMD pricing factor aligns with the first few dominant PCs. Columns (1) through (3) of Table 5 show the results of regressions of the NMD pricing factor on progressively more PCs. We find that NMD SDF loads strongly on the first two PCs, with the first PC explaining 46.8% of its variance and the second PC explaining an additional 17.8 percentage points. Adding the third PC increases the adjusted R^2 by a modest 3.9 percentage points. Column (4) through (6) of Table 5 shows that NMD CS loads even more heavily on the first PC, with the first PC explaining 58.3% of its variance. Adding the second PC increases the adjusted R^2 by another 5.9 percentage, whereas adding the third PC negligible increases the adjusted R^2 by 0.2 percentage points. Furthermore, consistent with this time-series evidence, we also find that a factor model using the first two PCs achieves a cross-sectional R^2 of 76% (See Panels A of Table 3 and Figure 1). Therefore, our time-series and cross-sectional results are consistent with the economic equilibrium described in Kozak, Nagel, and Santosh (2018), in which near-arbitrage opportunities do not exist. In contrast, if the large cross-sectional variation in average NMD returns is due to the data mining or publication bias, there is no reason to expect the NMD pricing factor, which summarizes this cross-sectional variation in average NMD returns, to load heavily on the common risk factors in the NMD return space. Thus, our test results can also be interpreted as being inconsistent with the data mining or publication bias hypotheses.

4. Relation to Sentiment-driven Demand and Required Returns from Liquidity Provision

As highlighted by Kozak et al. (2018), our finding that a NMD pricing factor loads heavily on common risk factors cannot be interpreted as evidence in support of rational asset pricing models relative to sentiment-driven asset pricing models. Therefore, in this section, we explore both perspectives by verifying if the premium of the NMD pricing factor is related

[18]Following KNS, we center but do not scale the returns when performing the PCA.

to sentiment-driven demand and the required compensation for arbitrageurs to accommodate such demands.

4.1. Sentiment-driven Demand

4.1.1. Timing of Sentiment-driven Demand

The notion introduced by LPS of a "tug-of-war" between the demand of opposing clienteles at the market open and close provides a good starting point for understanding how sentiment-driven demand can give rise to the NMD return predictabilities. Specifically, as we show in Figure 3, our composite predictive CS predicts (market-adjusted) NMD returns for up to 60 months into the future, while at the same time CS does not predict (market-adjusted) close-to-close returns.[19] This joint pattern of CS strongly predicting NMD returns and weakly predicting close-to-close returns have two possible tug-of-war explanations, which depend on whether trading demand at the market open or close gives rise to NMD return predictability. These two competing hypotheses posit that the associated sentiment-driven demand occurs repeatedly each day at the market open and/or close.

We design a test to distinguish between these two possibilities. If deviations of the NMD returns from CAPM are due to sentiment-driven demand at the market open, which is attenuated by opposing clientele demand throughout the day, then using prices progressively later in the morning as P_d^{open} in Eq. (1) to compute r_N and r_D should result in weaker NMD return predictability. In parallel, if the NMD return patterns are due to the sentiment-driven demand at the market close, then using prices earlier than 4:00 pm as P_d^{close} in Eq. (1) to compute r_N and r_D should lead to substantially weaker NMD return predictability. Figure 5 of the Appendix illustrates these two competing hypotheses.[20]

To test these two competing hypotheses, we divide the trading day into 13 half-hour trading intervals, with the first one starting at 9:30 am and the last one ending at 4:00 pm. Then, we recompute r_N and r_D using the following four alternative specifications of P_d^{open}: the first midquote after 9:30 am, 10:00 am, 10:30 am, and 11:00 am; and the following four alternative specifications of P_d^{close}: the last midquote before 4:00 pm, 3:30 pm, 3:00 pm, and 2:30 pm.[21]

[19]Our results are similar to those documented in Figure 2 of LPS, with the difference being 1) we use our composite signal whereas they use 2 out of the 17 LPS characteristics and 2) we show that our composite signal does not predict close-to-close returns.

[20]We assume the price is equal to the fundamental value at the market close (open) to simplify the exposition, but such an assumption is not necessary in our test.

[21]For all alternative specifications of P_d^{open} and P_d^{close}, we drop days with late openings or early closings and

Panel A of Table 6 reports the average returns of NMD SDF and NMD CS based on these alternative specifications of P_d^{open} or P_d^{close}. Columns "9:30 am" to "11:00 am" show a consistent pattern across the two specifications of the NMD pricing factor. That is, the average returns of the NMD pricing factor decreases substantially when we use midquotes progressively later in the morning as P_d^{open} to compute r_D and r_N. For example, the average NMD returns of NMD SDF (NMD CS) is 48.7%, 16.6%, 10.4%, and 7.5% p.a. (106.6%, 36.5%, 20.9% and 17.8% p.a.) when using the first midquote after 9:30 am, 10:00 am, 10:30 am, and 11:00 am as P_d^{open}, respectively. This declining pattern is consistent with the hypothesis that the NMD return predictabilities are due to sentiment-driven demand at the market open, which is then attenuated by opposing clientele demand throughout the day.

In sharp contrast, columns "2:30 pm" to "3:30 pm" show that using midquotes earlier in the afternoon, as opposed to the 4:00 pm midquote, as P_d^{close} to compute r_D and r_N has relatively minor effects on the resulting average NMD returns of the NMD pricing factor. The average NMD return of the NMD SDF (NMD CS) is 46.0%, 47.7%, 50.2%, and 48.8% p.a. (106.1%, 108.6%, 113.7%, and 109.9% p.a.) when using the last midquote before 2:30 pm, 3:00 pm, 3:30 pm, and 4:00 pm as P_d^{close} , respectively. Thus, these results are inconsistent with the hypothesis that NMD return predictabilities arise from sentiment-driven demand at the market close. Combining the market open and close results, our analysis indicate that sentiment-driven demand underlying NMD return predictabilities is more likely to occur at the market open than at the market close.

4.1.2. Identity of Sentiment-driven Demand

We next test whether the sentiment-driven demand at the market open that gives rise to NMD return predictabilities come from retail or institutional trading. Accurately deciphering the traders' identities on both sides of a transaction is an empirical challenge due to data limitations. We tackle this challenge by using the new method proposed in Boehmer, Jones, Zhang, and Zhang (2020) (BJZZ) to identify retail order flows. The BJZZ algorithm identifies retail purchase and sales by noting that trades at non-midpoints with a subpenny price improvement after the implementation of Regulation National Market System (Reg. NMS) are almost always marketable retail orders. These orders are recorded in TAQ with exchange code "D" and the buy/sell direction of the trade can be identified by the magnitude of the subpenny price improvements. We apply the BJZZ algorithm within the post-October

we drop the first (last) midquote if it happens after (before) midday. We use the WCT dataset from WRDS to compute the midquote as the average of the national best bid and offer. See the description of the WCT dataset here https://wrds-www.wharton.upenn.edu/pages/support/data-overview/wrds-overview-taq/

2007 period when Reg. NMS is fully implemented.[22] We use the retail marketable order imbalances identified by the BJZZ algorithm as the proxy for the retail order imbalance. We then use the buy-minus-sell total order imbalances identified by the Lee and Ready (1991) (LR) algorithm as the proxy for the total order imbalance. The difference between the total order imbalance and the retail order imbalance is then our proxy for the non-retail order imbalance.

We use the following procedure to measure the order imbalances associated with the NMD pricing factor. We first define order imbalances at the stock level for each of the 13 half-hour trading intervals,

$$\text{OI}_{i,d,\tau} = \frac{D_{\text{buy},i,d,\tau} - D_{\text{sell},i,d,\tau}}{\text{MCAP}_{i,d-1}}$$
$$\text{ROI}_{i,d,\tau} = \frac{D^{\text{retail}}_{\text{buy},i,d,\tau} - D^{\text{retail}}_{\text{sell},i,d,\tau}}{\text{MCAP}_{i,d-1}},$$

where $D_{\text{buy},i,d,\tau}$ and $D_{\text{sell},i,d,\tau}$ ($D^{\text{retail}}_{\text{buy},i,d,\tau}$ and $D^{\text{retail}}_{\text{sell},i,d,\tau}$) are the dollar values of the buy and sell orders identified by the LR (BJZZ) algorithm for stock i, day d, and the 30-minute interval τ. We scale them by the lagged market value of the stock, $\text{MCAP}_{i,d-1}$, to account for the size effect. Denote the portfolio weights of the long-short portfolio of NMD SDF (NMD CS) on stock i and day d by $w_{i,d}$. We then compute the order imbalances underlying the NMD pricing factor as,

$$\text{OI}_{d,\tau} = \frac{\sum_{i=1}^{n_{d,\tau}} w_{i,d} \times \text{OI}_{i,d,\tau} \times I\left(\text{OI}_{i,d,\tau}\right)}{\sum_{i=1}^{n_{d,\tau}} w_{i,d} \times I\left(\text{OI}_{i,d,\tau}\right)}$$
$$\text{ROI}_{d,\tau} = \frac{\sum_{i=1}^{n_{d,\tau}} w_{i,d} \times \text{ROI}_{i,d,\tau} \times I\left(\text{ROI}_{i,d,\tau}\right)}{\sum_{i=1}^{n_{d,\tau}} w_{i,d} \times I\left(\text{ROI}_{i,d,\tau}\right)},$$

where $I(\cdot)$ is an indicator function that equals one when $\text{OI}_{i,d,\tau}$ ($\text{ROI}_{i,d,\tau}$) is non-missing and zero otherwise. We annualize these order imbalance numbers by multiplying them by 13 times 252, so that they are in percentage points (relative to market capitalization).

Panel B of Table 6 reports the order imbalances for NMD SDF and NMD CS, respectively. Since our analysis in the previous subsection suggests that the sentiment-driven demand driving the NMD return predictabilities occurs at the market open, we focus on order imbalances in the first 30-minute interval. We observe that for both NMD pricing

[22]The BJZZ algorithm results in a substantial improvement in the coverage of retail order flows over other existing approaches. We refer interested readers to the related discussions in their study.

factors, the average total order imbalances (OI) at the market open are positive and highly significant at the 1% level. This is consistent with the explanation that sentiment-driven demand at the market open is relatively higher for stocks in the long end of the NMD pricing factor than stocks in the short end, resulting in positive overnight (negative intraday) returns on the NMD pricing factor.

When we decompose OI into the retail order imbalance (ROI) and the non-retail order imbalance (OI-ROI), we find that the significant positive total imbalance at the market open is almost entirely driven by the retail order imbalance. For NMD SDF, the OI is 1.41% while the ROI is 1.12%, which results in a OI-ROI of 0.29% that is not statistically significant. For NMD SDF, the OI is 2.99% while the ROI is 3.01%, which implies a OI-ROI of -0.02% that is again not statistically significant. Therefore, the sentiment-driven demand associated with the NMD pricing factor premium mostly comes from the retail trading demand.[23] Over the remaining 30-minute trading intervals, we do not find strong evidence that either retail or non-retail investors are trading aggressively in the opposite direction, suggesting that the price impact of the excess trading demand at the market open is instead potentially offset by non-marketable orders in the opposite direction. Overall, our results indicate that sentiment-driven demand at the market open is a plausible source of the premium of the NMD pricing factor and this sentiment-driven demand is from retail but not non-retail investors.

4.2. Liquidity Provision Hypothesis

To link NMD return predictabilities to the required returns for arbitrageurs, we first discuss what types of traders can exploit this type of predictability. First, we note that exploiting NMD return predictability involves two round-trip transactions per day by i) going long on the long-short portfolio of the NMD pricing factor at the market close and then closing this position at the next market open to earn the positive overnight return and ii) going short on the same long-short portfolio at the market open and then closing this position at the market close to benefit from the negative intraday return. If traders implement such a NMD trading strategy via market orders (i.e., buying stocks at the ask price and selling stocks at the bid price), then they would incur the bid-ask spread as a trading cost for each round-trip transaction. Transaction cost estimates from the existing literature, such as Novy-Marx and Velikov (2016), suggest that the trading cost would dwarf the excess returns on the NMD

[23]Our results are related to the retail trading patterns documented in Berkman et al. (2012), who find positive retail trading demand at the market open for stocks with high squared daily returns. They use a proprietary dataset that identifies retail buy and sell orders for a sample of NASDAQ 100 stocks between 1997 and 2001, whereas our data cover 3,000 stocks between 2007 and 2020.

pricing factor.[24]

While all traders can use limit orders or other types of orders to avoid paying the bid-ask spread, successfully executing limit orders ahead of other competing orders requires investments in both high-speed market access and the acquisition of uninformed order flows. As a result, the most likely marginal investor in NMD trading strategies is the market-making sector, who can use their market-making infrastructure to implement these strategies without paying the bulk of the effective bid-ask spread.

Moreover, we document in the previous subsection that the NMD pricing factor premium is associated with the excess marketable retail order imbalance at the market open in the same direction. This means that market makers can exploit NMD return predictability by trading against (i.e., providing liquidity to) the retail order imbalances underlying it. Such a scenario is highly plausible because in the U.S. markets, broker-dealer firms often have a first look at (and thus can execute) most marketable equity orders before these orders hit the exchange through either internalization or purchased order flow agreements.[25]

Therefore, we posit that market makers are the marginal arbitrageurs who accommodate the sentiment-driven demand underlying the NMD return predictabilities. As a result, the expected NMD returns should be closely related to the required returns from liquidity provision. When the required returns from liquidity provision vary over time as the risk-bearing capacity and the risk aversion of market makers fluctuate, the expected NMD returns should move in tandem.

To formally test this liquidity provision hypothesis, we follow Nagel (2012) and use the returns on the Lehmann (1990) short-term reversal strategy as a proxy for the required returns from liquidity provision.[26] Nagel (2012) motivates his measure by constructing a model in which investors trade for liquidity and informational reasons as in Kyle (1985), and risk-averse market makers charge a price for providing liquidity as in Grossman and Miller (1988). He theoretically shows a short-term reversal strategy isolates the returns from

[24]Novy-Marx and Velikov (2016) estimate that the round-trip transaction cost for typical value-weighted strategies is in excess of 50 basis points on average based on the effective spread measure in Hasbrouck (2009). This number implies that the trading cost for our NMD strategies would exceed 2% per day or 504% per annum.

[25]For example, Citadel Securities currently executes approximately 47% of all U.S.-listed retail volume. With their market-making infrastructure, Citadel Securities can then implement night-minus-day strategies trading against the retail order imbalances underlying night-minus-day return predictability. Citadel Securities does not pay but rather earns the bid-ask spread, although some of this revenue is offset in order to acquire order flows. See https://www.citadelsecurities.com/products/equities-and-options/.

[26]We replicate the reversal strategy return following the procedure in Nagel (2012) for the 1998 to 2020 sample period.

liquidity provision from the asymmetric information effect because the latter increases the bid-ask spread but does not cause negative serial correlations in returns. Nagel (2012) also carefully discusses why the Lehmann (1990) specification of the short-term reversal strategy is the preferred approximation of the returns earned by a hypothetical representative liquidity supplier. Given these arguments, we refer to Nagel (2012)'s reversal strategy return as the liquidity provision factor (denoted as LIQ) and explore its explanatory power for the premium on the NMD pricing factor.

In the first two columns of Panel A of Table 7, we regress the NMD pricing factor on LIQ. We use the monthly average of daily returns as a proxy for the conditional expected returns. We find that the returns from both NMD SDF and NMD CS load positively on LIQ, with a Newey and West (1987) $t-$statistics of 5.4 and 5.7, respectively. The economic magnitudes of these regression coefficients are large: a one percentage point increase in the return from liquidity provision is associated with a 0.6 (1.2) percentage point increase in NMD SDF (NMD CS). This strong time-series correlation also manifests in substantial regression R^2s of 34.5% and 27.2% for NMD SDF and NMD CS, respectively.

To visualize these time-series correlations, Figure 4 plots the six-month moving averages of the NMD strategy returns against that of LIQ. We observe two consistent patterns across the two specifications of the NMD pricing factor. First, the NMD pricing factor and LIQ covary positively and strongly, with a secular decline over time that is consistent with the profits of both strategies being driven down by increased competition among market makers over time. Second, and more importantly, the NMD pricing factor and LIQ spike simultaneously during the 2000 dot-com crash, the 2009 financial crisis, and the 2020 COVID-19 stock market crash, when liquidity providers have more constrained risk-bearing capacities and higher risk aversions. The COVID-19 stock market crash episode is particularly informative as an out-of-sample test relative to the sample period studied by Nagel (2012). The simultaneous large increases in LIQ and the NMD pricing factor around the COVID-19 stock market crash not only support Nagel's argument that short-term reversal returns proxy for the returns from liquidity provision, but also our hypothesis that NMD return predictabilities are related to the required returns from liquidity provision.

This single-factor model based on LIQ also explains the vast majority of the average returns on the NMD pricing factor. Because both the left- and right-hand-side variables are returns, we can interpret the regression intercepts as alphas, that is, the average returns unexplained by exposure to LIQ. The alphas are more than 80% smaller in magnitude than the premium on the NMD pricing factor, with a $t-$statistic of 1.5 and 1.8 for NMD SDF and NMD CS, respectively. Therefore, these results indicate that the required returns from liquidity provision largely explain NMD return predictabilities.

As a placebo test, we also use a lower frequency reversal factor based on the past 21-day return in Columns (3) and (4). This placebo for the returns from liquidity provision is provided on Ken French's website (hereafter, FF REV). Given that market makers' inventory imbalances have a very short half-life, FF REV is unlikely to accurately capture the returns from liquidity provision. Consistent with this intuition, and in sharp contrast with LIQ, we find that neither specifications of the NMD pricing factor significantly load on FF REV and both regressions have near zero adjusted R^2.

4.3. Sentiment-driven Demand vs. Required Returns from Liquidity Provision

Finally, we explore the relative importance of sentiment-driven demand and the required returns from liquidity provision in driving the time-series variation in the NMD pricing factor. While our evidence on sentiment-driven demand and the required compensation for arbitrageurs are consistent with the economic equilibrium described by Kozak, Nagel, and Santosh (2018), KNS also point out that both are needed for the repeated NMD deviations from CAPM to persist. Therefore, their relative importance in driving the time-series variation in the magnitude of the NMD pricing factor is an empirical question.

Table 8 reports our results for a horserace between proxies for sentiment-driven demand and the required returns from liquidity provision. As before, we use LIQ to measure the required returns from liquidity provision. Then, given our findings in Subsection 4.1, we use the OI underlying the NMD pricing factor in the first 30-minute trading interval as our proxy for the associated sentiment-driven demand. In the period after October 2007 when ROI is available, we also use the first 30-minute ROI underlying the NMD pricing factor as an alternative proxy for related sentiment-driven demand. Columns (1) and (3) of Table 7 report the results using OI as the explanatory variable in the full sample when LIQ is available. Here, we also include a pre-decimalization dummy, which is equal to one before April 2001 to account for the potential effects brought about by the introduction of decimalization. We find that both NMD SDF and NMD CS load significantly and positively on OI, with $t-$statistics of 3.9 and 4.1, respectively. Columns (5) and (7) of Table 7 report the results using ROI in the post-October 2007 period. Similarly, we find that both NMD SDF and NMD CS continue to have positive loadings on ROI, with $t-$statistics of 2.5 and 1.7, respectively. These results alternatively confirm our results from Subsection 4.1 that sentiment-driven demand is positively related to the NMD return predictability.

When we add LIQ to these regressions, we find that the loadings on our proxies for sentiment-driven demand and the required returns from liquidity provision remain positive and statistically significant at the 5% level. In Columns (5) and (7), we find that ROI by itself explains around 4.3% or 1.3% of the time-series variation in NMD SDF or NMD CS, respectively. In contrast, adding LIQ increases the adjusted R^2 by 22.6 or 19.6 percentage points for NMD SDF or NMD CS, respectively. Therefore, our results indicate that sentiment-driven demand and the required returns from liquidity provision have independent explanatory power, with the latter economic fundamental being determining more of the time-series variation.

5. Conclusion

We document that a one-factor model that summarizes the predictive information in a large number of stock characteristics for the cross-section of the night-minus-day (NMD) stock returns. For this model, we construct the NMD pricing factor using both a covariance-based SDF approach as in Kozak, Nagel, and Santosh (2020) and a characteristic-based approach similar to Stambaugh and Yuan (2017). We find that both specifications of the NMD pricing factor can adequately price the average NMD returns of 17 long-short portfolios examined by Lou, Polk, and Skouras (2019) and achieves a cross-sectional R^2 well above those of standard factor models. Out-of-sample, our proposed one-factor model also works well when the test asset set is augmented by 80 long-short portfolios sorted on additional anomaly characteristics.

Using the NMD pricing factor as a parsimonious summary of NMD return predictability, we subsequently explore underlying economic sources. Consistent with the near-absence of arbitrage opportunities, we find the NMD pricing factor has substantial exposures to the dominant common risk factors within the NMD return space. We also provide new complementary evidence that the NMD factor premium is related to sentiment-driven trading demand at the market open, which almost exclusively comes from retail trading demand. Finally, we introduce a novel liquidity provision hypothesis, and find that the NMD factor premium is highly correlated with the required returns from liquidity provision. Overall, our findings are consistent with the pricing equilibrium described in Kozak, Nagel, and Santosh (2018), in which liquidity providers require compensation for accommodating sentiment-driven demand.

6. Data Appendix

6.1. Anomaly characteristics for the LPS portfolios

We provide detailed descriptions of the characteristics used in constructing the 17 LPS portfolios as follows. Annual accounting variables from Compustat are merged with the CRSP information under the assumption that the accounting information is available with a 6-month lag.

- One-month overnight and intraday returns ($r_{N,t}$ and $r_{D,t}$, respectively): the average of the overnight and intraday returns of a stock over month t, respectively, as in Lou, Polk, and Skouras (2019).
- Exponentially-weighted moving averages of the overnight and intraday returns ($r^{\text{ewma}}_{N,t}$ and $r^{\text{ewma}}_{D,t}$, respectively): The exponentially-weighted moving averages of the one-month overnight and intraday returns with a half-life of 12 months. These exponentially-weighted moving averages are then lagged by one month so that they are not mechanically contemporaneously correlated with $r_{N,t}$ and $r_{D,t}$, as in Lou, Polk, and Skouras (2019).
- Size (ME): Natural log of market capitalization at the end of month t.
- Book-to-market ratio (BM): Book value of equity (ceq) divided by end of fiscal year-end market capitalization.
- Momentum (MOM): 11-month cumulative returns ending at month $t-1$, as in Jegadeesh (1990).
- Industry momentum (INDMOM): Equal weighted industry average momentum at the two-digit SIC level.
- SUE: I/B/E/S actual earnings minus median forecasted earnings (if unavailable, the seasonally differenced quarterly earnings before extraordinary items from Compustat quarterly file) divided by fiscal-quarter-end market cap, as in Rendleman, Jones, and Latane (1982).
- ROE: Earnings before extraordinary items divided by one-quarter lagged common shareholders' equity as in Hou, Xue, and Zhang (2015).

- BETA: the CAPM beta in 12-month rolling windows ending at month t by regressing daily excess returns on market excess returns. We require at least 120 valid excess return observations in each rolling window.

- ISSUE: Annual percent change in shares outstanding (csho) as in Pontiff and Woodgate (2008).

- TURNOVER: Average monthly trading volume between month $t-2$ and month t (inclusive) scaled by number of shares outstanding at month t, as in Datar, Naik, and Radcliffe (1998).

- Investment (INV): Annual change in gross property, plant, and equipment (ppegt) + annual change in inventories (invt) all scaled by one-year lagged total assets (at), as in Chen and Zhang (2010).

- IVOL: One-month standard deviation of the residuals from the FF3 model, as in Ang, Hodrick, Xing, and Zhang (2006).

- ACCRUALS: Annual income before extraordinary items (ib) minus operating cash flows (oancf) divided by average total assets (at); if oancf is missing, change in act - change in che - change in lct + change in dlc + change in txp-dp, as in Sloan (1996).

- Reversal (STR): One-month return at month t, as in Jegadeesh and Titman (1993).

6.2. TAQ Data

We collect intraday prices and trades from the NYSE Trade and Quote (TAQ) database available from Wharton Research Data Service (WRDS). The monthly updated TAQ database is a legacy product for trades and quotes between January 1, 1993 and December 31, 2014 that are time-stamped at the frequency of a second. The daily updated TAQ database provides trades and quotes from September 10, 2003 to the present that are time-stamped with millisecond precision between September 10, 2003 and March 31, 2015 and with microsecond precision from April 1, 2015 to present. The daily product contains the official national best bid and ask (NBBO) quotes, while the monthly product does not.

Matching trades to the prevailing NBBO quotes requires significant processing time. In particular, quotes are updated more frequently than the occurrence of the trades, which means one needs to update the NBBO quotes every time quotes are updated. For our research purposes, we follow WRDS guidance and implement the Lee and Ready (1991) algorithm and the Boehmer, Jones, Zhang, and Zhang (2020) algorithm using the WRDS Consolidated Trades (WCT) dataset, which contains all the trades and their matched NBBO quotes for the seconds 0, -1, -2 and -5 relative to their trade time. WRDS provides a monthly WCT dataset based on the TAQ monthly product and a daily WCT dataset based on the TAQ daily product. We primarily use the daily WCT dataset and, when unavailable, the monthly WCT dataset.

We only use trades and quotes between 9:30 am and 4:00 pm inclusive. To account for holidays and trading halts, we exclude trading days with less than 6.5 trading hours per day. When we use the monthly WCT dataset, we impose the following filters to identify regular trades within the regular trading hours following Bollerslev et al. (2016): "cond" in the list of ('@','F','E',' ','@E','@F',' E',' F'), "corr"=0, "price">0 and "size">0. When we use the daily WCT dataset, we use the same filters except for "corr"=0 as this variable is not available in the daily WCT dataset. The sale condition variable "cond" in the monthly WCT dataset is changed to be "Tr_SCond". Importantly, from December 9, 2013 onwards, we change the filter for "Tr_SCond" to be in the list of ('@','F','E',' ','@E','@F',' E',' F','@ I', '@F I', 'F I','I'). We also add '@ I', '@F I', 'F I','I' to the list because from December 9, 2013 onward TAQ uses 'I' to indicate Odd Lot trades.

The Lee and Ready (1991) algorithm requires matching each trade to a bid-ask quote. When we use the monthly WCT dataset, we follow the SAS code provided by WRDS and use the quotes at the end of the previous second.[27] When we use the daily WCT dataset, we follow Holden and Jacobsen (2014) and use the prevailing NBBO matched to trades by WRDS and drop data points for which the market is either locked or crossed.[28] The NBBO of a market is defined as being crossed (locked) if the national best bid is greater than (equal to) the national best offer.

In a few cases, the TAQ data show recording errors such as misplaced decimal points. We thus exclude a data point (including the trade and the matched quote data) if the trade price is 50% smaller than the mid-point quote price or vice versa. We also compute the median trade price within a day. If the median price is 90% smaller than a trade price or vice versa, we exclude that data point.

[27]See, https://wrds-www.wharton.upenn.edu/pages/support/applications/microstructure-research/lee-and-ready-1991-algorithm/

[28]See, http://www.excelmodeling.com/Holden-and-Jacobsen-Daily-TAQ-and-Monthly-TAQ-Codes-2018-03-16.zip

6.3. Merging TAQ Data with CRSP Data

We use the WRDS-provided link tables to merge both the monthly updated and daily updated TAQ data to CRSP. We use the daily linktable when available and supplement it with the monthly linktable. The monthly and daily linktable provides the TAQ symbol (i.e., SYMBOL for the monthly linktable, and SYM_ROOT + SYM_SUFFIX for the daily linktable), NCUSIP, and CRSP PERMNO for each stock-date. However, the match between TAQ symbol and CRSP PERMNO in the linktable is not unique for a stock-date. Most of these cases are related to multiple share classes. We use the following procedure to identify the correct TAQ symbol and CRSP PERMNO match when a match is not unique for a stock-date.

1. For the daily linktable:

    (a) We first delete preferred stocks (i.e., SYM_SUFFIX in the list of PRA, PRB, PRC) and then compute a match score as the sum of the boolean value of the following conditions:

        i. TAQ SYM_ROOT is equal to CRSP ticker

        ii. TAQ SYM_ROOT and SYM_SUFFIX are equal to CRSP ticker and CRSP shrcls, respectively.

        iii. TAQ NCUSIP is equal to CRSP NCUSIP

    (b) We then choose the match with the highest score. If we still see multiple TAQ symbol and CRSP PERMNO matches per stock day, we use the one match for which TAQ SYM_SUFFIX is empty (i.e., a regular share class).

2. For the monthly linktable:

    (a) We compute a match score as the sum of the boolean value of the following conditions.

        i. TAQ SYMBOL is equal to CRSP ticker

        ii. TAQ SYMBOL are equal to the concatenated CRSP ticker and CRSP shrcls

        iii. TAQ CUSIP is equal to CRSP NCUSIP

    (b) We then choose the match with the highest score. In 0.02% of the cases, we still see multiple TAQ symbol and CRSP PERMNO matches per stock day, due to the recording error. For these stock dates that have duplicated matches, we use the unique TAQ symbol and CRSP PERMNO match for this stock from the nearest date.

3. The cleaned daily linktable and the monthly linktable as described above match a unique TAQ symbol to each CRSP PERMNO-date. We then use these cleaned linktables to merge with CRSP data. For each CRSP PERMNO-date, we use the TAQ symbol from the daily linktable whenever possible, which is then supplemented with the monthly linktable. We find that some CRSP PERMNO-dates do not have TAQ symbol data from either the daily linktable or the monthly linktable. Most of these cases happen around the dates when there are ticker changes. For these CRSP PERMNO-dates that have missing TAQ symbol, we use the TAQ symbol for the same CRSP PERMNO-ticker in the same year (forward filled first and then backward filled) from the daily linktable, and again supplemented with the monthly linktable. After these steps above, we have a unique non-missing TAQ symbol for 99.8% of the PERMNO-date observations in our sample.

6.4. Alternative Open and Close Prices

When we compute r_N and r_D using prices later (earlier) than 9:30 am (4:00 pm) from TAQ database as the alternative specifications of $P_{\text{open}}(P_{\text{close}})$, we follow Nagel (2012) and Bogousslavsky (2021) in using the CRSP adjustment factor (FACPR) and dividend payment (DIVAMT) to adjust for stock splits and dividends. Specifically, we first compute the daily dividend, div_d, as the sum of DIVAMT for each firm. We then compute the close-to-close return based on the P_{close} as follows,

$$r_{\text{close-to-close},d} = \frac{P_{\text{close},d} \times \text{FACPR}_d + div_d}{P_{\text{close},d-1}}.$$

The calculation for $r_{D,d}$ remains the same as that in Eq. (1) of the main paper and we then use Eq. (2) of the main paper to back out the overnight return.

To mitigate the influence of outliers when using the alternative open and close prices, we set the intraday and overnight returns to be missing on a given day if the intraday or overnight return calculated using first/last mid-quotes is more than 25 percentage points different from the corresponding CRSP intraday or overnight return.

References

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. 2006. The cross-section of volatility and expected returns. Journal of Finance, 61(1), 259–299. ISSN 0022-1082.

Berkman, H., Koch, P. D., Tuttle, L., and Zhang, Y. J. 2012. Paying Attention: Overnight Returns and the Hidden Cost of Buying at the Open. The Journal of Financial and Quantitative Analysis, 47(4), 715–741. ISSN 0022-1090. Publisher: Cambridge University Press.

Boehmer, E., Jones, C. M., Zhang, X., and Zhang, X. Tracking Retail Investor Activity. SSRN Scholarly Paper ID 2822105, Social Science Research Network, Rochester, NY, 2020. doi:10.2139/ssrn.2822105.

Bogousslavsky, V. 2021. The cross-section of intraday and overnight returns. Journal of Financial Economics. ISSN 0304-405X. doi:10.1016/j.jfineco.2020.07.020.

Bollerslev, T., Li, S. Z., and Todorov, V. 2016. Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns. Journal of Financial Economics, 120(3), 464–490. ISSN 0304-405X. doi:10.1016/j.jfineco.2016.02.001.

Chen, L. and Zhang, L. 2010. A better three-factor model that explains more anomalies. Journal of Finance, 65(2), 563–595.

Daniel, K., Hirshleifer, D., and Sun, L. 2020. Short- and Long-Horizon Behavioral Factors. The Review of Financial Studies, 33(4), 1673–1736. ISSN 0893-9454. doi:10.1093/rfs/hhz069.

Datar, V. T., Naik, N. Y., and Radcliffe, R. 1998. Liquidity and stock returns: An alternative test. Journal of financial markets, 1(2), 203–219.

Fama, E. F. and French, K. R. 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56. ISSN 0304-405X. doi:10.1016/0304-405X(93)90023-5.

Fama, E. F. and French, K. R. 2015. A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1–22. ISSN 0304405X. doi:10.1016/j.jfineco.2014.10.010.

Fama, E. F. and French, K. R. 2016. Dissecting Anomalies with a Five-Factor Model. Review of Financial Studies, 29(1), 69–103. ISSN 0893-9454, 1465-7368. doi:10.1093/rfs/hhv043.

Fama, E. F. and MacBeth, J. D. 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy, 81(3), 607–636.

Freyberger, J., Neuhierl, A., and Weber, M. 2020. Dissecting Characteristics Nonparametrically. The Review of Financial Studies, 33(5), 2326–2377. ISSN 0893-9454. doi:10.1093/rfs/hhz123. Publisher: Oxford Academic.

Green, J., Hand, J. R. M., and Zhang, X. F. 2017. The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. The Review of Financial Studies, 30(12), 4389–4436. ISSN 0893-9454. doi:10.1093/rfs/hhx019.

Grossman, S. J. and Miller, M. H. 1988. Liquidity and Market Structure. The Journal of Finance, 43(3), 617–633. ISSN 1540-6261. doi:https://doi.org/10.1111/j.1540-6261.1988.tb04594.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1988.tb04594.x.

Hansen, L. P. and Jagannathan, R. 1991. Implications of Security Market Data for Models of Dynamic Economies. Journal of Political Economy, 99(2), 225–262. ISSN 0022-3808. doi:10.2307/2937680.

Hasbrouck, J. 2009. Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data. The Journal of Finance, 64(3), 1445–1477. ISSN 1540-6261. doi:10.1111/j.1540-6261.2009.01469.x.

Hendershott, T., Livdan, D., and Rösch, D. 2020. Asset pricing: A tale of night and day. Journal of Financial Economics, 138(3), 635–662. ISSN 0304405X. doi:10.1016/j.jfineco.2020.06.006.

Heston, S. L., Korajczyk, R. A., and Sadka, R. 2010. Intraday Patterns in the Cross-section of Stock Returns. The Journal of Finance, 65(4), 1369–1407. ISSN 1540-6261. doi:https://doi.org/10.1111/j.1540-6261.2010.01573.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01573.x.

Holden, C. W. and Jacobsen, S. 2014. Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions. The Journal of Finance, 69(4), 1747–1785. ISSN 1540-6261. doi:10.1111/jofi.12127.

Hou, K., Mo, H., Xue, C., and Zhang, L. 2019. Which factors? Review of Finance, 23(1), 1–35.

Hou, K., Xue, C., and Zhang, L. 2015. Digesting anomalies: an investment approach. Review of Financial Studies, 28(3), 650–705. ISSN 0893-9454. doi:10.1093/rfs/hhu068.

Jegadeesh, N. 1990. Evidence of Predictable Behavior of Security Returns. The Journal of Finance, 45(3), 881. ISSN 00221082. doi:10.2307/2328797.

Jegadeesh, N. and Titman, S. 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance, 48(1), 65–91. ISSN 00221082. doi:10.2307/2328882.

Kozak, S., Nagel, S., and Santosh, S. 2018. Interpreting Factor Models: Interpreting Factor Models. The Journal of Finance, 73(3), 1183–1223. ISSN 00221082. doi:10.1111/jofi.12612.

Kozak, S., Nagel, S., and Santosh, S. 2020. Shrinking the cross-section. Journal of Financial Economics, 135(2), 271–292. ISSN 0304-405X. doi:10.1016/j.jfineco.2019.06.008.

Kyle, A. S. 1985. Continuous Auctions and Insider Trading. Econometrica, 53(6), 1315. ISSN 00129682. doi:10.2307/1913210.

Lee, C. M. C. and Ready, M. J. 1991. Inferring Trade Direction from Intraday Data. The Journal of Finance, 46(2), 733–746. ISSN 1540-6261. doi:https://doi.org/10.1111/j.1540-6261.1991.tb02683.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1991.tb02683.x.

Lehmann, B. N. 1990. Fads, Martingales, and Market Efficiency. The Quarterly Journal of Economics, 105(1), 1–28. ISSN 0033-5533. doi:10.2307/2937816. Publisher: Oxford University Press.

Lettau, M. and Pelger, M. 2020. Factors That Fit the Time Series and Cross-Section of Stock Returns. The Review of Financial Studies, 33(5), 2274–2325. ISSN 0893-9454. doi:10.1093/rfs/hhaa020.

Lewellen, J. 2015. The Cross-section of Expected Stock Returns. Critical Finance Review, 4(1), 1–44. ISSN 2164-5744, 2164-5760. doi:10.1561/104.00000024.

Lewellen, J., Nagel, S., and Shanken, J. 2010. A skeptical appraisal of asset pricing tests. Journal of Financial Economics, 96(2), 175–194. ISSN 0304-405X. doi:10.1016/j.jfineco.2009.09.001.

Lou, D., Polk, C., and Skouras, S. 2019. A tug of war: Overnight versus intraday expected returns. Journal of Financial Economics, 134(1), 192–213. ISSN 0304-405X. doi:10.1016/j.jfineco.2019.03.011.

Nagel, S. 2012. Evaporating Liquidity. Review of Financial Studies, 25(7), 2005–2039. ISSN 0893-9454, 1465-7368. doi:10.1093/rfs/hhs066.

Newey, W. K. and West, K. D. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703. ISSN 00129682.

Novy-Marx, R. and Velikov, M. 2016. A Taxonomy of Anomalies and Their Trading Costs. Review of Financial Studies, 29(1), 104–147. ISSN 0893-9454, 1465-7368. doi:10.1093/rfs/hhv063.

Pontiff, J. and Woodgate, A. 2008. Share Issuance and Cross-sectional Returns. The Journal of Finance, 63(2), 921–945.

Rendleman, R. J., Jones, C. P., and Latane, H. A. 1982. Empirical anomalies based on unexpected earnings and the importance of risk adjustments. Journal of Financial economics, 10(3), 269–287.

Sloan, R. G. 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? The Accounting Review, 71(3), 289–315. ISSN 0001-4826.

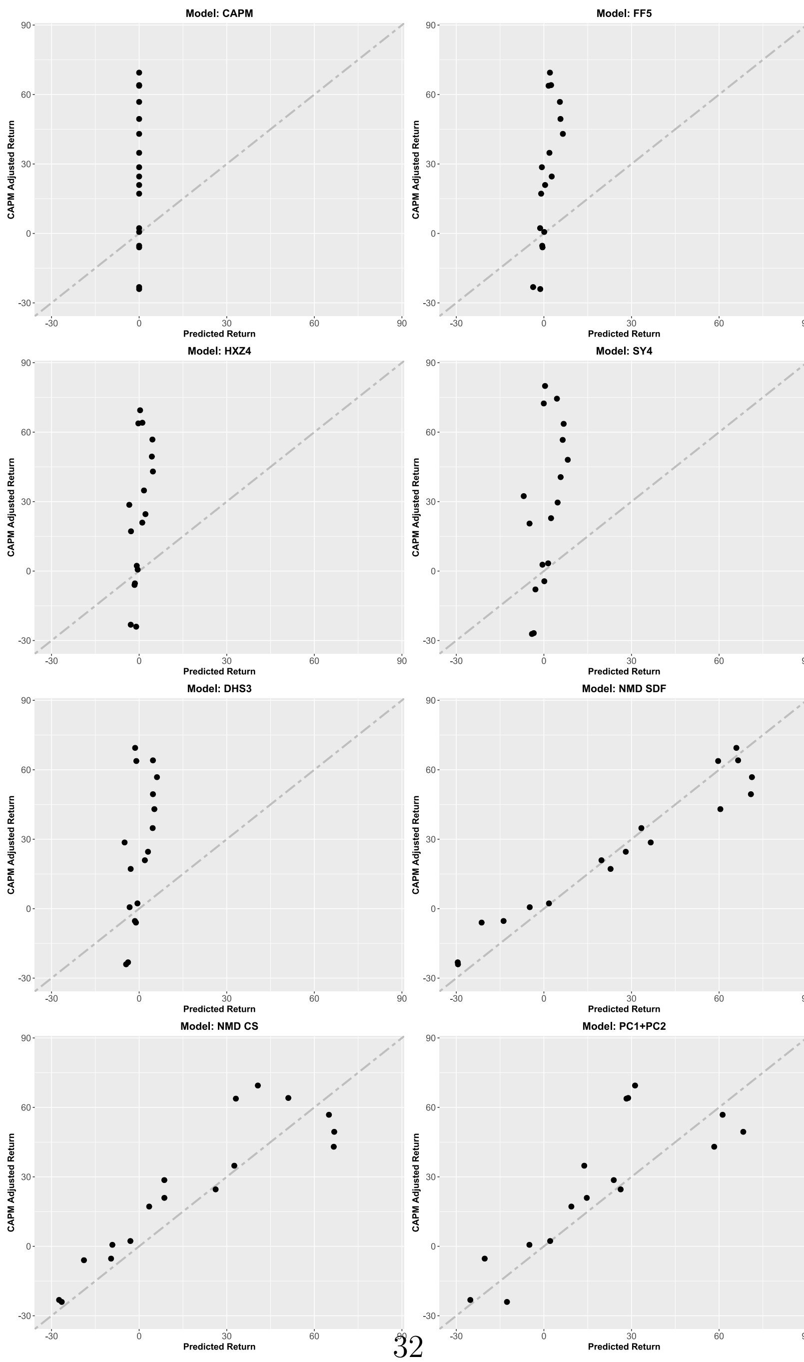
Stambaugh, R. F., Yu, J., and Yuan, Y. 2015. Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle: Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. The Journal of Finance, 70(5), 1903–1948. ISSN 00221082. doi:10.1111/jofi.12286.

Stambaugh, R. F. and Yuan, Y. 2017. Mispricing Factors. The Review of Financial Studies, 30(4), 1270–1315. ISSN 0893-9454. doi:10.1093/rfs/hhw107.

Figure 1: Cross-sectional Pricing Performance

This figure plots the average (market-adjusted) NMD return against the predicted NMD returns from five benchmark factor models (CAPM, FF5, HXZ4, SY4, and DHS3), our proposed one-factor models based on NMD SDF and NMD CS, and a two-factor model based on the first two PCs. In Panel A, the test assets include the 17 LPS portfolios. In Panel B, the test assets include the 17 LPS portfolios and 80 long-short portfolios sorted on additional anomaly signals from Green, Hand, and Zhang (2017). The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

Panel A - 17 LPS Portfolios



Panel B - Expanded Test Assets

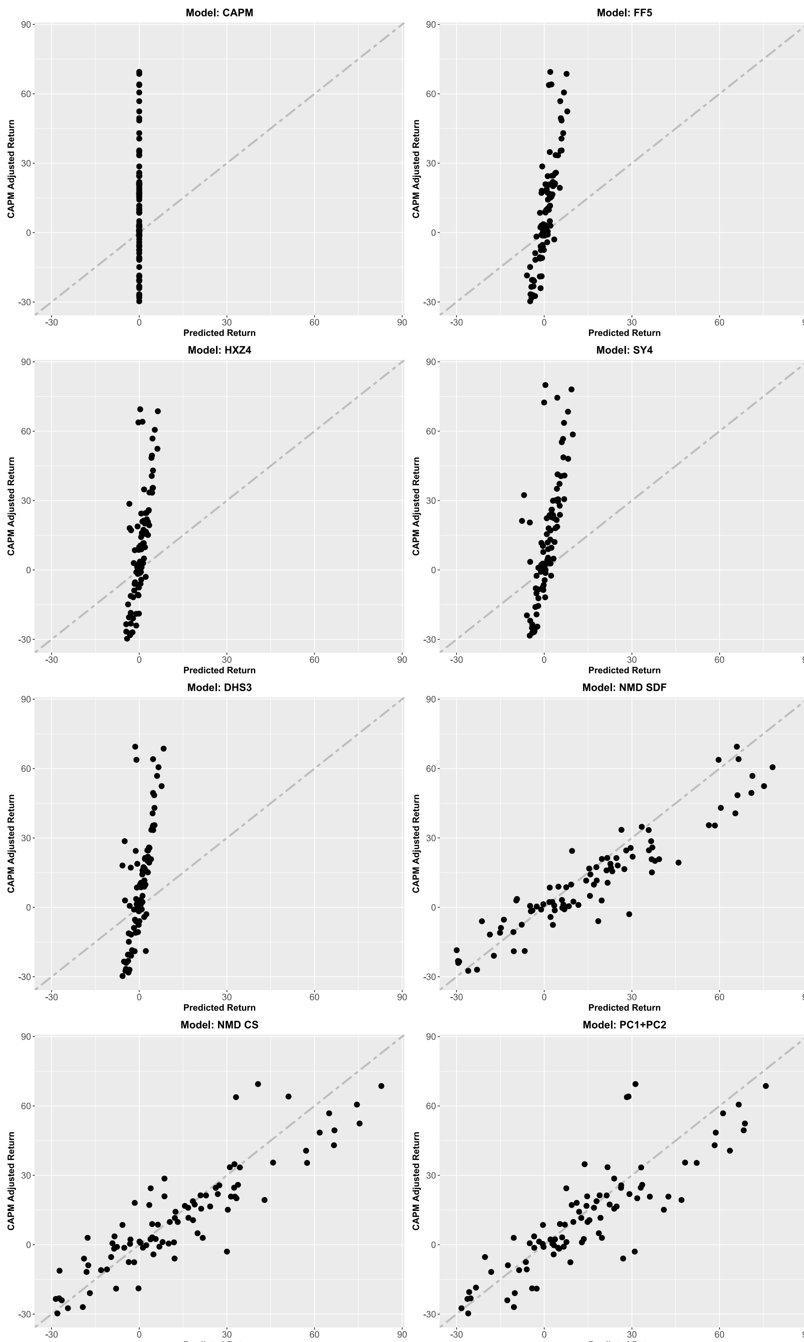


Figure 2: Ordered Sharpe Ratios

This figure plots the annualized Sharpe ratio of NMD SDF and NMD CS, as well as that of the 17 LPS portfolios. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

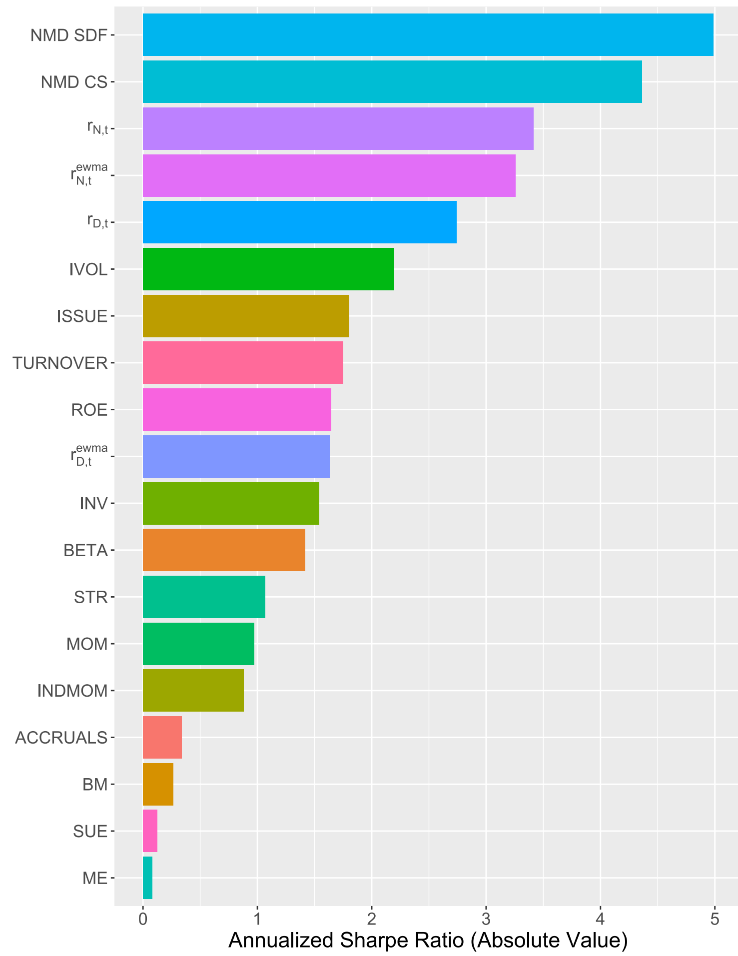


Figure 3: Long-term Predictability of NMD CS

This figure plots the time-series average of the overnight (RN), intraday (RD), night-minus-day (NMD), and close-to-close (RET) returns of the long-short portfolio sorted on CS (the composite predictive signal for NMD returns) 1 to 60 months after the measurement of CS. We report 95% confidence intervals around the mean return based on two times the Newey and West (1987) standard error with 21 lags. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

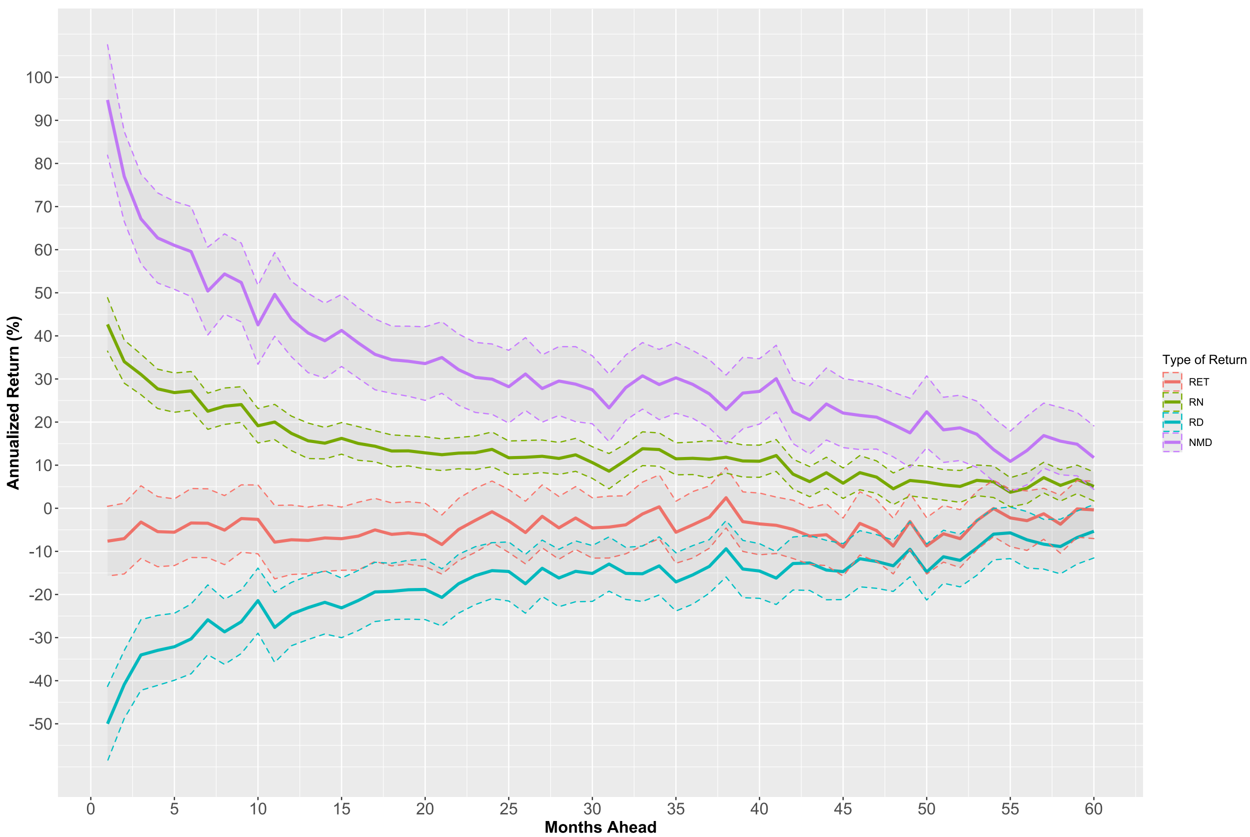
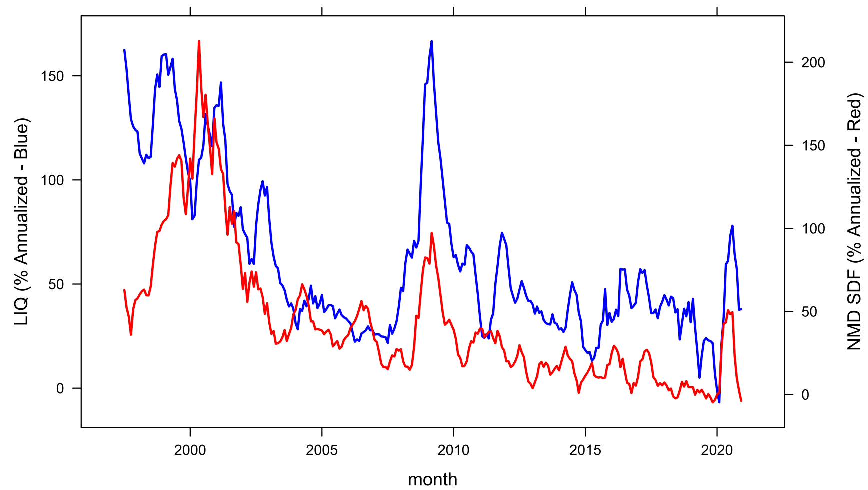


Figure 4: Relation with Liquidity Provision Factor

This figure plots the six-month moving averages of liquidity provision factor (LIQ) proposed by Nagel (2012) against each of NMD SDF and NMD CS. The left axis refers to the scale of LIQ and the right axis refers to the scale of the NMD pricing factor. The sample period is from January 1998 to December 2020.



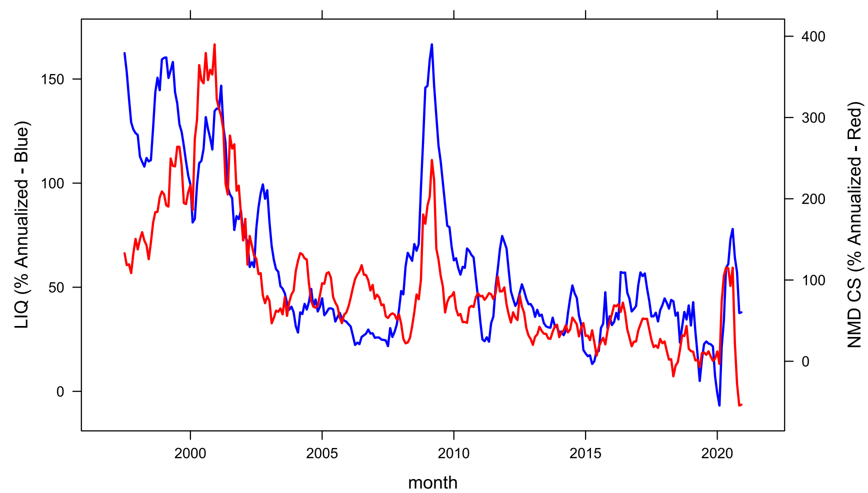


Table 1: NMD Returns and Standard Factor Models

This table presents the time-series average of the unadjusted NMD returns of the 17 LPS portfolios defined in Section 1.3, as well as the corresponding alphas relative to standard factor models. The standard factor models include the CAPM, the five-factor model of Fama and French (2016) (FF5) , the q-factor model of Hou, Mo, Xue, and Zhang (2019) (HXZ4), the mispricing model of Stambaugh and Yuan (2017) (SY4), or the behavioral model of Daniel, Hirshleifer, and Sun (2020) (DHS3). All portfolio returns are annualized to be in percentage points per annum. We report t-statistics computed from Newey and West (1987) standard errors with 21 lags in the parentheses. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

Signal	R	α^{CAPM}	α^{FF5}	α^{HXZ4}	α^{SY4}	α^{DHS3}
$r_{D,t}$	65.87	66.94	64.42	65.87	72.44	62.10
	(10.58)	(10.79)	(10.79)	(10.81)	(10.98)	(10.15)
$r_{N,t}$	69.57	70.40	68.32	70.09	80.36	71.74
	(11.94)	(11.98)	(12.25)	(11.98)	(11.74)	(12.14)
$r_{D,t}^{\text{ewma}}$	36.20	36.93	34.99	35.30	36.64	32.21
	(7.19)	(7.38)	(7.16)	(7.19)	(6.71)	(6.72)
$r_{N,t}^{\text{ewma}}$	64.37	65.21	63.57	65.51	73.64	66.11
	(13.34)	(13.42)	(13.63)	(13.39)	(13.28)	(13.46)
BETA	49.00	53.10	46.39	48.52	48.46	47.46
	(7.92)	(9.04)	(8.22)	(8.42)	(7.40)	(8.16)
IVOL	60.01	62.38	56.79	57.91	61.54	56.04
	(8.98)	(9.63)	(9.76)	(9.76)	(8.74)	(9.09)
BM	-5.10	-5.27	-4.81	-3.68	-4.41	-4.20
	(-1.32)	(-1.36)	(-1.44)	(-1.01)	(-1.05)	(-1.10)
ISSUE	25.73	26.76	24.04	24.63	26.75	23.64
	(8.07)	(8.37)	(8.12)	(8.09)	(7.69)	(7.85)
ACCRUALS	-5.33	-5.59	-5.01	-4.14	-4.77	-4.12
	(-1.64)	(-1.71)	(-1.62)	(-1.26)	(-1.34)	(-1.28)
INV	21.07	21.50	21.08	20.44	20.88	19.52
	(6.68)	(6.85)	(6.96)	(6.61)	(6.18)	(6.28)
ROE	-23.87	-25.24	-21.51	-22.39	-25.10	-21.36
	(-6.38)	(-6.92)	(-6.83)	(-6.85)	(-6.62)	(-6.53)
ME	1.18	-0.09	-0.19	0.41	1.26	3.20
	(0.37)	(-0.03)	(-0.09)	(0.18)	(0.51)	(1.06)
SUE	1.59	1.70	3.02	2.52	2.76	2.31
	(0.55)	(0.59)	(1.10)	(0.89)	(0.85)	(0.81)
MOM	27.03	27.50	28.22	30.87	38.34	32.53
	(4.48)	(4.51)	(4.91)	(5.09)	(5.78)	(5.57)
STR	-26.00	-26.34	-25.05	-25.35	-25.04	-21.77
	(-4.53)	(-4.57)	(-4.50)	(-4.50)	(-3.99)	(-3.79)
TURNOVER	53.43	55.90	50.09	51.64	55.70	50.91
	(9.43)	(10.17)	(9.61)	(9.87)	(9.32)	(9.42)
INDMOM	16.47	16.47	17.43	19.25	24.86	19.40
	(3.92)	(3.91)	(4.26)	(4.48)	(5.26)	(4.48)

Table 2: Summary Statistics

This table presents the mean, standard deviation, and Sharpe ratio for NMD SDF, NMD CS, the first three PCs, the liquidity provision factor of Nagel (2012) (LIQ), and the short-term reversal factor provided by Ken French (FF Rev). All returns are in percentage points per annum. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

	Mean	SD	SR
NMD SDF	44.89	9.01	4.99
NMD CS	94.73	21.72	4.36
PC1	34.16	12.47	2.74
PC2	21.33	11.13	1.92
PC3	10.01	10.03	1.00
LIQ	62.47	10.07	6.21
FF REV	14.02	15.18	0.92

Table 3: Pricing Performance of the Factor Models

This table presents the pricing performance of several factor models for explaining NMD returns. Columns (1) and (2) report the cross-sectional means of the absolute alpha and the absolute $t-$statistic, respectively. Column (3) reports the cross-sectional R^2 calculated using Eq. (12). We include five benchmark factor models: CAPM, FF5, HXZ4, SY4, and DHS3. Our proposed one-factor models are based on NMD SDF and NMD CS. We also include a two-factor model based on the first two PCs. In Panel A, the test assets consist of the 17 LPS portfolios. In Panel B, the test assets include the 17 LPS portfolios and 80 long-short portfolios sorted on additional anomaly characteristics from Green, Hand, and Zhang (2017). We report t-statistics computed from Newey and West (1987) standard errors with 21 lags. All returns are in percentage points per annum. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

Panel A - 17 LPS Portfolios

Model	$\lvert\alpha\rvert$	$\lvert t\rvert$	$R^{2,\text{XS}}$
CAPM	31.42	6.29	0.00
FF5	29.54	6.31	0.11
HXZ4	30.30	6.29	0.06
SY4	33.82	6.38	0.10
DHS3	29.82	6.07	0.09
NMD SDF	7.34	1.26	0.94
NMD CS	12.38	2.06	0.84
PC1+PC2	14.57	3.36	0.76

Panel B - Expanded Test Assets

Model	$\lvert\alpha\rvert$	$\lvert t\rvert$	$R^{2,\text{XS}}$
CAPM	19.43	4.49	0.00
FF5	17.28	4.44	0.19
HXZ4	17.86	4.43	0.14
SY4	20.17	4.44	0.18
DHS3	17.66	4.25	0.17
NMD SDF	9.20	1.86	0.79
NMD CS	9.36	1.91	0.78
PC1+PC2	10.05	2.68	0.72

Table 4: NMD Factor Structure

This table reports on the percentage of the total variance explained by the first 10 PCs extracted from the NMD returns of the 17 LPS portfolios. All returns are in percentage points per annum. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020.

Type		PC1	PC2	PC3	PC4	PC5	PC6 to PC10
R_N	Perc Variance	0.25	0.21	0.13	0.07	0.06	0.17
	Cumulative Perc	0.25	0.46	0.59	0.66	0.72	0.89
R_D	Perc Variance	0.27	0.23	0.13	0.06	0.06	0.15
	Cumulative Perc	0.27	0.50	0.63	0.70	0.75	0.90
NMD	Perc Variance	0.26	0.22	0.13	0.07	0.06	0.16
	Cumulative Perc	0.26	0.48	0.61	0.68	0.73	0.90

Table 5: Relation to PCs

This table reports on time-series regressions of the NMD pricing factor on the first three PCs of the NMD returns of the 17 LPS portfolios. The dependent variable is either NMD SDF or NMD CS. We report t-statistics computed from Newey and West (1987) standard errors with 21 lags in the parentheses. All returns are in percentage points per annum. The portfolio returns are measured between February 1st, 1996 and December 31st, 2020. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	NMD SDF_t			NMD CS_t		
	(1)	(2)	(3)	(4)	(5)	(6)
$PC1_t$	0.494***	0.468***	0.465***	1.331***	1.295***	1.293***
	(11.225)	(15.559)	(15.731)	(21.472)	(24.331)	(24.782)
$PC2_t$		0.343***	0.343***		0.472***	0.473***
		(7.939)	(9.416)		(5.655)	(5.944)
$PC3_t$			0.178***			0.117*
			(5.958)			(1.840)
Constant	28.019***	21.595***	19.895***	49.283***	40.431***	39.315***
	(13.063)	(10.806)	(10.670)	(12.520)	(10.941)	(10.496)
N	6,273	6,273	6,273	6,273	6,273	6,273
Adjusted R^2	0.468	0.646	0.685	0.583	0.642	0.644

Table 6: Relation to Sentiment-driven Demand

Columns "9:30 am" to "11:00 am" of Panel A reports the average returns of NMD SDF and NMD CS when using the first midquote after 9:30 am, 10:00 am, 10:30 am, and 11:00 am to replace the CRSP open price as P_d^{open} to compute the NMD returns, respectively. Columns "2:30 pm" to "4:00 pm" of Panel A reports the average returns of NMD SDF and NMD CS when using the last midquote before 2:30 pm, 3:00 pm, 3:30 pm, and 4:00 pm as P_d^{close} to compute the NMD returns. Panel B presents the time-series average of the order imbalances associated with NMD SDF and NMD CS. Each column presents the ending point of a half-hour trading interval during market trading hours, with the "Open" ("Close") column ending at 10:00 am (4:00 pm). OI is the total order imbalances identified by the Lee and Ready (1991) algorithm and ROI is the retail order imbalances identified by the Boehmer, Jones, Zhang, and Zhang (2020) algorithm. We report t-statistics computed from Newey and West (1987) standard errors with 21 lags in the parentheses. The portfolio returns are in percentage per annum and are measured between February 1st, 1996 and December 31st, 2020. The order imbalances are in annualized percentages of the market capitalization and are measured between November 1st, 2007 and December 31st, 2020.

Panel A - Alternative Specifications of P_d^{open} and P_d^{close}

NMD SDF	09:30 am	10:00 am	10:30 am	11:00 am	02:30 pm	03:00 pm	3:30 pm	4:00 pm
r_D	-24.99	-9.18	-6.12	-4.71	-24.15	-24.85	-26.34	-25.68
	(-11.48)	(-6.36)	(-4.91)	(-4.27)	(-12.46)	(-12.00)	(-11.76)	(-11.49)
r_N	23.68	7.45	4.25	2.82	21.82	22.83	23.84	23.10
	(15.47)	(6.42)	(3.46)	(1.97)	(13.19)	(13.39)	(13.61)	(13.99)
NMD	48.67	16.62	10.37	7.54	45.97	47.68	50.18	48.78
	(14.91)	(9.19)	(6.48)	(4.36)	(15.11)	(14.71)	(14.27)	(14.30)

NMD CS	09:30 am	10:00 am	10:30 am	11:00 am	02:30 pm	03:00 pm	3:30 pm	4:00 pm
r_D	-52.51	-18.05	-10.43	-8.89	-51.18	-52.32	-55.40	-53.59
	(-9.74)	(-4.57)	(-3.19)	(-3.12)	(-10.62)	(-10.38)	(-10.45)	(-10.01)
r_N	54.08	18.45	10.50	8.94	54.93	56.28	58.26	56.35
	(14.96)	(5.94)	(3.13)	(2.32)	(14.35)	(14.55)	(15.21)	(15.44)
NMD	106.60	36.50	20.93	17.83	106.11	108.60	113.66	109.94
	(14.08)	(7.57)	(5.12)	(4.11)	(15.56)	(15.20)	(15.20)	(14.95)

Panel B - Intraday Order Imbalances

NMD SDF	Open	10:30 am	11:00 am	11:30 am	12:00 pm	12:30 pm	1:00 pm	1:30 pm	2:00 pm	2:30 pm	3:00 pm	3:30 pm	Close
OI	1.41	1.03	0.76	0.61	0.60	0.47	0.37	0.41	0.34	0.25	0.64	0.28	0.89
	(3.97)	(5.71)	(2.98)	(3.53)	(4.19)	(4.93)	(3.74)	(4.24)	(3.34)	(2.93)	(4.34)	(2.19)	(4.98)
ROI	1.12	0.62	0.48	0.40	0.36	0.31	0.29	0.25	0.22	0.24	0.24	0.14	-0.07
	(7.12)	(6.42)	(6.09)	(6.51)	(6.96)	(5.36)	(5.78)	(6.09)	(5.72)	(4.80)	(5.70)	(3.20)	(-1.13)
OI-ROI	0.29	0.41	0.28	0.21	0.24	0.16	0.09	0.16	0.12	0.01	0.41	0.14	0.96
	(1.08)	(3.00)	(1.36)	(1.52)	(1.98)	(2.19)	(0.90)	(1.70)	(1.37)	(0.20)	(3.40)	(1.17)	(6.71)

NMD CS	Open	10:30 am	11:00 am	11:30 am	12:00 pm	12:30 pm	1:00 pm	1:30 pm	2:00 pm	2:30 pm	3:00 pm	3:30 pm	Close
OI	2.99	1.64	1.59	1.42	0.77	0.84	0.63	0.45	0.16	0.18	0.86	-0.32	1.14
	(3.24)	(3.07)	(2.44)	(2.97)	(1.89)	(2.29)	(2.16)	(1.11)	(0.44)	(0.73)	(1.65)	(-0.78)	(1.76)
ROI	3.01	1.63	1.19	0.97	0.87	0.72	0.69	0.52	0.46	0.49	0.47	0.20	-0.36
	(8.68)	(8.87)	(8.65)	(7.90)	(9.37)	(6.25)	(7.31)	(5.73)	(5.33)	(4.72)	(5.27)	(1.94)	(-1.70)
OI-ROI	-0.01	0.01	0.39	0.45	-0.11	0.11	-0.06	-0.07	-0.31	-0.31	0.38	-0.52	1.50
	(-0.02)	(0.01)	(0.70)	(1.08)	(-0.30)	(0.36)	(-0.22)	(-0.20)	(-0.95)	(-1.48)	(0.79)	(-1.27)	(2.54)

Table 7: Relation to the Required Returns from Liquidity Provision

This table presents time-series regressions of the monthly average returns of the NMD pricing factor on that of Nagel (2012)'s liquidity provision factor (LIQ) or the short-term reversal factor provided by Ken French (FF REV). We report t-statistics computed from Newey and West (1987) standard errors with 12 lags in the parentheses under the coefficients. All returns are in percentage points per annum. The sample period is from January 1998 to December 2020. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	NMD SDF$_t$	NMD CS$_t$	NMD SDF$_t$	NMD CS$_t$
	(1)	(2)	(3)	(4)
LIQ$_t$	0.603***	1.212***		
	(5.393)	(5.737)		
FF REV$_t$			0.055	−0.068
			(0.965)	(−0.451)
Constant	6.698	17.767*	43.522***	94.112***
	(1.461)	(1.762)	(5.077)	(5.513)
N	287	287	287	287
Adjusted R^2	0.345	0.272	−0.0005	−0.003

Table 8: Sentiment-driven Demand vs. Required Returns from Liquidity Provision

This table presents monthly time-series regressions of the NMD pricing factor on the sentiment-driven demand proxied by the OI (ROI) in the first 30-minute trading interval and/or the required returns from liquidity provision proxied by Nagel (2012)'s liquidity provision factor. We report t-statistics computed from Newey and West (1987) standard errors with 12 lags in the parentheses under the coefficients. The sample period is from January 1998 to December 2020. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

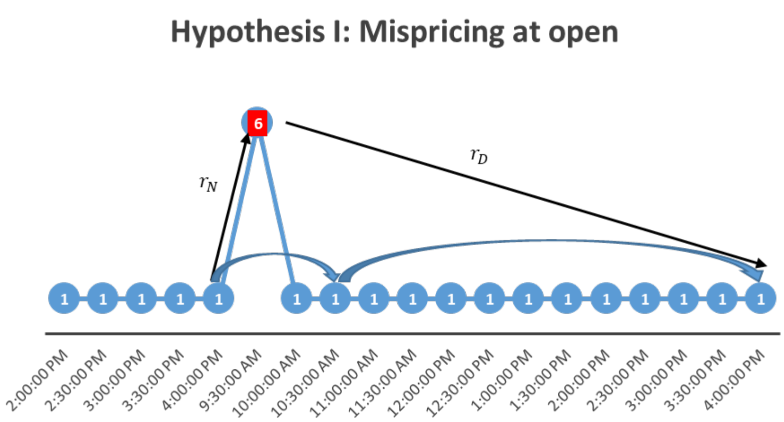
	NMD SDF_t		NMD CS_t		NMD SDF_t		NMD CS_t	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OI SDF_t	1.254***	1.254***						
	(3.915)	(3.606)						
OI CS_t			0.910***	0.905***				
			(4.130)	(3.679)				
ROI SDF_t					8.178**	5.930**		
					(2.493)	(1.989)		
ROI CS_t							5.505*	5.734**
							(1.733)	(2.057)
LIQ_t		0.330***		0.702***		0.345***		0.889***
		(5.193)		(4.094)		(3.758)		(4.182)
Pre-$Decim_t$	59.203***	32.198**	116.154***	58.924*				
	(4.077)	(1.993)	(3.846)	(1.676)				
Constant	23.204***	7.461**	52.613***	19.177*	12.436**	−1.733	32.098**	−11.632
	(5.050)	(1.999)	(5.716)	(1.952)	(2.262)	(−0.486)	(2.169)	(−0.968)
N	287	287	287	287	158	158	158	158
Adjusted R^2	0.457	0.518	0.326	0.379	0.043	0.269	0.013	0.209

7. Appendix

Figure 5: Timing of the Mispricing Underlying NMD Return Predictabilities

Panel A (Panel B) visualizes the hypothesis that the NMD return predictabilities are driven by sentiment-driven demand that occurs repeatedly each day at the market open (at the market close), while the opposing clientele demand offsets such demands during the rest of the day. We assume the price is equal to the fundamental value at the market close (open) in Panel A (B) to simplify the exposition, but such an assumption is not necessary for what we want to show. In Panel A, the fundamental value of the stock is \$1. The price at 9:30 am is overpriced at \$6, and the overpricing is fully corrected by 10:00 am. Using the 9:30 am price as P_{open} to compute returns results in a positive r_N and a negative r_D. In contrast, using the 10:00 am price as P_{open} to compute returns results in a zero r_N and r_D and thus zero night-minus-day return predictability. In Panel B, the fundamental value of the stock is \$6. The price at 4:00 pm is underpriced at \$1, and the underpricing is fully corrected by the market open of the next day. Using the 4:00 pm price as P_{close} to compute returns results in a positive r_N and negative r_D. In contrast, using the 3:30 pm price as P_{close} to compute returns results in a zero r_N and r_D and thus zero night-minus-day return predictability.

Panel A



Panel B

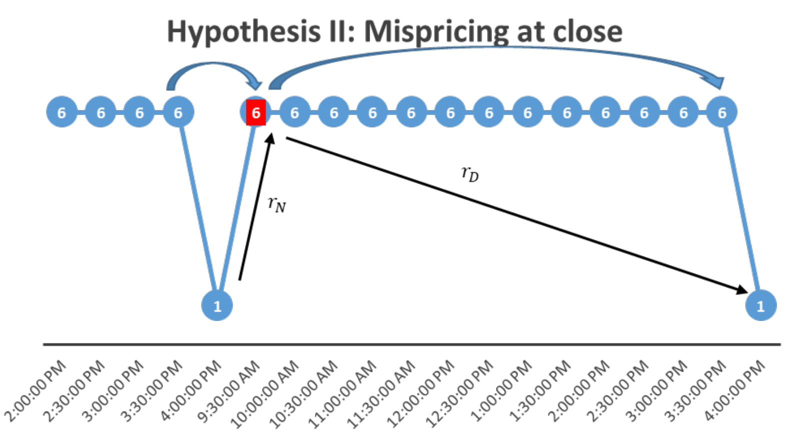


Table 9: VWP Specification: NMD Returns and Standard Factor Models

This table reports the results of Table 1, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

Signal	R	α^{CAPM}	α^{FF5}	α^{HXZ4}	α^{SY4}	α^{DHS3}
$r_{D,t}$	36.64	36.90	35.35	36.15	42.06	34.53
	(7.26)	(7.35)	(7.21)	(7.29)	(7.66)	(6.89)
$r_{N,t}$	42.06	42.63	41.68	42.43	51.09	43.36
	(10.01)	(10.13)	(10.15)	(9.87)	(10.29)	(10.06)
$r_{D,t}^{\text{ewma}}$	16.42	16.69	15.31	15.47	17.10	14.59
	(3.96)	(4.04)	(3.73)	(3.78)	(3.69)	(3.58)
$r_{N,t}^{\text{ewma}}$	42.82	43.34	42.66	43.73	50.08	43.71
	(11.05)	(11.17)	(11.26)	(11.11)	(11.92)	(11.19)
BETA	38.30	40.96	37.18	38.80	38.81	37.85
	(6.91)	(7.57)	(6.96)	(7.07)	(6.19)	(6.88)
IVOL	42.57	43.85	40.67	41.33	46.32	40.36
	(7.75)	(8.12)	(8.00)	(8.04)	(7.75)	(7.62)
BM	-0.19	-0.43	0.36	1.03	1.02	0.38
	(-0.06)	(-0.13)	(0.11)	(0.32)	(0.28)	(0.12)
ISSUE	9.88	10.42	8.83	9.18	12.94	8.60
	(3.59)	(3.77)	(3.40)	(3.43)	(4.32)	(3.23)
ACCRUALS	-7.66	-7.88	-7.58	-6.94	-7.44	-6.82
	(-2.72)	(-2.81)	(-2.74)	(-2.44)	(-2.36)	(-2.42)
INV	11.75	11.78	11.61	11.41	12.44	10.96
	(4.61)	(4.61)	(4.60)	(4.44)	(4.38)	(4.26)
ROE	-15.36	-15.88	-13.80	-14.30	-17.42	-13.39
	(-5.04)	(-5.22)	(-5.02)	(-5.11)	(-5.56)	(-4.76)
ME	10.94	10.38	10.35	10.49	11.97	12.57
	(3.83)	(3.71)	(4.57)	(4.64)	(4.83)	(4.51)
SUE	3.16	3.06	3.82	3.52	3.14	3.72
	(1.29)	(1.23)	(1.55)	(1.41)	(1.14)	(1.48)
MOM	23.26	23.79	24.54	25.46	28.58	26.92
	(4.74)	(4.80)	(5.00)	(5.02)	(5.02)	(5.50)
STR	-13.97	-13.71	-12.59	-12.95	-13.57	-11.51
	(-2.96)	(-2.91)	(-2.77)	(-2.82)	(-2.63)	(-2.44)
TURNOVER	47.32	48.78	45.34	46.32	50.47	45.93
	(9.77)	(10.13)	(9.82)	(9.80)	(9.51)	(9.66)
INDMOM	17.59	17.61	18.11	18.77	23.07	19.66
	(4.62)	(4.58)	(4.75)	(4.71)	(5.17)	(4.93)

Table 10: VWP Specification: Summary Statistics

This table reports the results of Table 2, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

	Mean	SD	SR
NMD SDF	26.48	7.10	3.73
NMD CS	60.48	20.20	2.99
PC1	22.64	11.53	1.96
PC2	17.14	9.82	1.75
PC3	5.13	9.06	0.57
LIQ	62.45	10.07	6.20
FF REV	14.03	15.18	0.92

Table 11: VWP Specification: Pricing Performance of the Factor Models

This table reports the results of Table 3, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

Panel A - 17 LPS Portfolios

Model	$\lvert\alpha\rvert$	$\lvert t\rvert$	$R^{2,\text{XS}}$
CAPM	21.20	5.16	0.00
FF5	20.09	5.10	0.10
HXZ4	20.50	5.10	0.06
SY4	23.76	5.45	0.07
DHS3	20.44	4.97	0.06
NMD SDF	4.71	1.20	0.95
NMD CS	9.90	2.44	0.76
PC1+PC2	9.21	2.82	0.79

Panel B - Expanded Test Assets

Model	$\lvert\alpha\rvert$	$\lvert t\rvert$	$R^{2,\text{XS}}$
CAPM	13.92	3.88	0.00
FF5	12.68	3.76	0.15
HXZ4	12.99	3.76	0.11
SY4	15.31	4.04	0.14
DHS3	12.93	3.66	0.12
NMD SDF	5.95	1.69	0.83
NMD CS	6.57	1.89	0.74
PC1+PC2	7.27	2.45	0.73

Table 12: VWP Specification: NMD Factor Structure

This table reports the results of Table 4, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

Type		PC1	PC2	PC3	PC4	PC5	PC6 to PC10
R_N	Perc Variance	0.25	0.22	0.14	0.06	0.06	0.17
	Cumulative Perc	0.25	0.46	0.61	0.67	0.73	0.89
R_D	Perc Variance	0.28	0.22	0.13	0.07	0.06	0.15
	Cumulative Perc	0.28	0.51	0.63	0.70	0.75	0.90
NMD	Perc Variance	0.27	0.21	0.13	0.06	0.06	0.16
	Cumulative Perc	0.27	0.48	0.61	0.67	0.73	0.89

Table 13: VWP Specification: Relation to PCs

This table reports the results of Table 5, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

	NMD SDF_t			NMD CS_t		
	(1)	(2)	(3)	(4)	(5)	(6)
$PC1_t$	0.368***	0.348***	0.345***	1.317***	1.294***	1.292***
	(10.513)	(12.396)	(12.362)	(26.273)	(24.335)	(24.661)
$PC2_t$		0.380***	0.385***		0.428***	0.431***
		(12.794)	(13.785)		(5.637)	(5.653)
$PC3_t$			0.140***			0.082
			(6.509)			(1.343)
Constant	18.147***	12.084***	11.351***	30.671***	23.851***	23.423***
	(12.614)	(9.575)	(9.642)	(9.898)	(8.354)	(8.058)
N	6,271	6,271	6,271	6,271	6,271	6,271
Adjusted R^2	0.357	0.633	0.665	0.565	0.608	0.610

Table 14: VWP Specification: Relation to Sentiment-driven Demand

This table reports the results of Panel B of Table 6, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

NMD SDF	Open	10:30 am	11:00 am	11:30 am	12:00 pm	12:30 pm	1:00 pm	1:30 pm	2:00 pm	2:30 pm	3:00 pm	3:30 pm	Close
OI	1.33	0.68	0.51	0.38	0.35	0.32	0.20	0.26	0.12	0.04	0.40	-0.06	0.20
	(3.92)	(4.24)	(2.04)	(2.31)	(2.30)	(2.68)	(1.74)	(2.45)	(1.06)	(0.43)	(2.70)	(-0.49)	(1.01)
ROI	1.12	0.58	0.44	0.35	0.32	0.26	0.25	0.21	0.19	0.20	0.19	0.10	-0.12
	(7.87)	(6.65)	(6.38)	(6.53)	(7.14)	(5.18)	(5.86)	(5.83)	(5.18)	(4.39)	(5.55)	(2.33)	(-1.84)
OI-ROI	0.21	0.11	0.07	0.02	0.04	0.06	-0.05	0.04	-0.07	-0.16	0.21	-0.16	0.31
	(0.80)	(0.84)	(0.33)	(0.19)	(0.26)	(0.63)	(-0.54)	(0.44)	(-0.66)	(-2.16)	(1.60)	(-1.43)	(1.95)

NMD CS	Open	10:30 am	11:00 am	11:30 am	12:00 pm	12:30 pm	1:00 pm	1:30 pm	2:00 pm	2:30 pm	3:00 pm	3:30 pm	Close
OI	2.99	1.64	1.59	1.42	0.77	0.84	0.63	0.45	0.16	0.18	0.86	-0.32	1.14
	(3.24)	(3.07)	(2.44)	(2.97)	(1.89)	(2.29)	(2.16)	(1.11)	(0.44)	(0.73)	(1.65)	(-0.78)	(1.76)
ROI	3.01	1.63	1.19	0.97	0.87	0.72	0.69	0.52	0.46	0.49	0.47	0.20	-0.36
	(8.68)	(8.87)	(8.65)	(7.90)	(9.37)	(6.25)	(7.31)	(5.73)	(5.33)	(4.72)	(5.27)	(1.94)	(-1.70)
OI-ROI	-0.01	0.01	0.39	0.45	-0.11	0.11	-0.06	-0.07	-0.31	-0.31	0.38	-0.52	1.50
	(-0.02)	(0.01)	(0.70)	(1.08)	(-0.30)	(0.36)	(-0.22)	(-0.20)	(-0.95)	(-1.48)	(0.79)	(-1.27)	(2.54)

Table 15: VWP Specification: Relation to the Required Returns from Liquidity Provision

This table reports the results of Table 7, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

	NMD SDF_t	NMD CS_t	NMD SDF_t	NMD CS_t
	(1)	(2)	(3)	(4)
LIQ_t	0.342***	0.896***		
	(3.905)	(4.392)		
FF REV_t			0.004	−0.143
			(0.135)	(−1.200)
Constant	4.630	3.370	25.864***	60.990***
	(1.261)	(0.361)	(4.935)	(4.691)
N	287	287	287	287
Adjusted R^2	0.245	0.222	−0.003	0.002

Table 16: VWP Specification: Sentiment-driven Demand vs. Required Returns from Liquidity Provision

This table reports the results of Table 8, using the volume-weighted price in the first 30-minute trading interval to compute NMD returns.

	NMD SDF_t		NMD CS_t		NMD SDF_t		NMD CS_t	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OI SDF_t	0.940***	0.931***						
	(5.361)	(4.773)						
OI CS_t			0.645***	0.643**				
			(2.775)	(2.561)				
ROI SDF_t					5.049**	4.703**		
					(2.379)	(2.292)		
ROI CS_t							5.168**	5.280**
							(2.043)	(2.119)
LIQ_t		0.120***		0.380**		0.101**		0.434**
		(3.064)		(2.589)		(2.297)		(2.449)
Pre-$Decim_t$	35.801***	26.173***	104.189***	73.199***				
	(5.861)	(4.464)	(4.463)	(3.117)				
Constant	10.765***	5.108*	26.472***	8.382	4.449	−0.041	5.480	−15.882
	(5.210)	(1.854)	(4.489)	(0.920)	(1.446)	(−0.012)	(0.599)	(−1.312)
N	287	287	287	287	158	158	158	158
Adjusted R^2	0.465	0.481	0.337	0.359	0.025	0.058	0.021	0.094