What Do the Portfolios of Individual Investors Reveal About the Cross-Section of Equity Returns?

Sebastien Betermier, Laurent E. Calvet, Samuli Knüpfer, and Jens Kvaerner*

December 31, 2021

Abstract

We extract a parsimonious set of equity factors from comprehensive administrative information on the stock holdings of Norwegian individual investors in 1997-2018. A three-factor model featuring the market factor and two investor factors based on stockholders' age and wealth explains common variation in portfolio holdings and prices the cross-section of stock returns. Portfolio tilts toward the investor factors correlate with indebtedness, macroeconomic exposure, gender, education, and investment experience. Our results are consistent with hedging and sentiment jointly driving portfolio decisions and the cross-section of equity premia.

JEL Codes: G11, G12.

Keywords: Asset pricing, factor-based investing, household finance, portfolio allocation.

*Betermier: Desautels Faculty of Management, McGill University, 1001 Sherbrooke St West, Montreal, QC H3A 1G5, Canada; sebastien.betermier@mcgill.ca. Calvet: EDHEC Business School, 16 rue du Quatre-Septembre, 75002 Paris, France, and CEPR; laurent.calvet@edhec.edu. Knüpfer, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway; samuli.knupfer@bi.no. Kvaerner: Tilburg University, Warandelaan 2 5037 AB Tilburg, Netherlands; jkverner@gmail.com. We received helpful comments from Fabio Braggion, David Brown, Joost Driessen, Paul Ehling, Evan Jo, Patrick Konermann, Serhiy Kozak, Hughes Langlois, and Richard Priestley, and seminar participants at Arizona State, BI Norwegian Business School, EDHEC Business School, McGill University, and Tilburg University, the CEPR Sixth European Workshop on Household Finance, the 2021 Northern Finance Association, the 2021 Red Rock Finance Conference, and the 2021 Academic Research Colloquium for Financial Planning and Related Disciplines. We gratefully acknowledge financial support from Finansmarkedsfondet at Research Council of Norway under Project Number No. 309855.

I. Introduction

Standard finance theory states that investor demand is a key determinant of equity risk premia (Merton, 1973; Lintner, 1966; Sharpe, 1964). However, empirical work in asset pricing rarely uses direct observations of demand. Most empirically successful pricing factors rely on firm characteristics (Fama and French, 2015; Hou, Xue, and Zhang, 2015) and a large literature attempts to tie firm factors to arguably noisy measures of aggregate consumption (see Constantinides (2017); Ludvigson (2013); Mehra (2012) and the references therein).

In this paper, we leverage comprehensive administrative information on the stock holdings of individuals and investigate what they reveal about the factors that price the cross-section of equity returns. Our analysis aims to answer four questions. If one sorts stocks by the characteristics of the individual investors who own them, do these characteristics produce factors that price the cross-section of stock returns? Which investor characteristics should matter in theory and which ones matter in the data? How do the investor factors compare with traditional factors constructed from firm characteristics? Last but not least, how do investor characteristics, risks, and biases relate to portfolio tilts toward the pricing factors?

Individual investors are attractive to study because a large body of financial theory predicts important interactions between investors' portfolio decisions and asset returns. Furthermore, the high degree of portfolio heterogeneity among individuals, successfully documented and analyzed in the field of household finance (Campbell, 2006; Gomes et al., 2020; Guiso and Sodini, 2013), holds great promise for understanding the drivers of demand for individual stocks. Here direct stockholdings are likely particularly informative because they are not clouded by the frictions emanating from delegated asset management (He and Xiong, 2013).[1] Betermier, Calvet, and Sodini (2017) accordingly show that the links between individuals' characteristics and portfolio exposures to the value factor are most pronounced in direct stockholdings. Finally, despite individuals' limited ownership fraction of aggregate equity, they appear to matter for stock returns (Blume and Keim, 2012; Barber et al., 2009; Kaniel et al., 2008; Kelley and Tetlock, 2013).

A major impediment to relating individual investor portfolios to asset returns is that available datasets often lack dimensions that are crucial for performing rigorous asset pricing tests, such as a long time series, a large and diverse pool of investors, and detailed investor characteristics. For example, the well-known Barber and Odean (2000, 2001) dataset includes

[1] A parallel literature investigates investor preferences and asset prices through mutual fund flows (Barber et al., 2016; Berk and van Binsbergen, 2016; Guercio and Tkac, 2002).

five years of transactions by retail investors trading through a particular discount broker. By comparison, studies of the cross-section of equity returns frequently use at least twenty years of data.[2] Our paper resolves this challenge by using comprehensive administrative data on all Norwegian direct stockholders in 1997-2018.

The main results of the paper can be summarized as follows. We show theoretically that portfolios of stocks sorted by the age or wealth of their individual investors should produce powerful pricing factors. Using our Norwegian data, we verify empirically that a three-factor model consisting of a mature-minus-young factor, a high wealth-minus-low wealth factor, and the market factor performs well in explaining common variation in portfolio holdings and pricing the cross-section of stock returns, both in and out of sample. The tight connection between investor factors and investor portfolio decisions allows us to shed light on the underlying mechanisms at play. We document that investor wealth, indebtedness, macroeconomic exposure, age, gender, education, and investment experience explain investor portfolio tilts toward the new factors. Our findings support the view that hedging motives and sentiment jointly drive investor factor tilts.

Our detailed contributions are the following. We first study the theoretical link between the cross-sections of investor portfolios and stock returns. We show that conditional on a portfolio factor structure, market clearing implies that the returns on the portfolio factors generate pricing factors that explain the cross-section of returns. As a result, pricing factors can be recovered from a sufficiently heterogeneous set of investor portfolios.

We next use financial theory to endogenize the portfolio factor structure. We demonstrate that two investor characteristics - age and wealth – are likely to drive the cross-section of investor portfolios and therefore the cross-section of equity returns. We derive this result in two complementary settings: an ICAPM model (Merton, 1973) that combines time-varying investment opportunities and labor income risk, and a model with sentiment in the spirit of Fedyk et al. (2013) and Sandroni (2000). These models predict that an investor's portfolio should be closer to the tangency portfolio for more mature or wealthier investors. These results hold irrespective of the details of the model, such as the nature and number of state risks. Mature and wealthy investors should therefore earn higher CAPM alphas than young and less wealthy investors.

[2]As Merton (1980) explains, the high level of volatility in stock returns makes statistical inference on average returns challenging in small samples. As a back-of-the-envelope calculation, consider an asset whose abnormal return has a sample monthly average of 1% and a volatility of 4%. Given 5 years of monthly data, the probability of correctly rejecting the null hypothesis that the average abnormal return (alpha) is 0% is only equal to 47%. Given 20 years of monthly data, the probability goes up to 97%.

We construct investor-based equity factors from a comprehensive administrative panel containing the stock holdings and socioeconomic characteristics of Norwegian individual investors in 1997-2018. This complete ownership record covers more than 400 stocks listed on the Oslo Stock Exchange (OSE) and is remarkable for its large cross-section of investor portfolios (about 365,000 investors a year) and long time series (21 years).

We uncover a strong factor structure in the portfolio holdings of individual investors. These commonalities across investor portfolios are clearly visible when we aggregate portfolio holdings according to investors' age, wealth, and other socioeconomic characteristics. A principal component analysis of these portfolios reveals that three principal components explain 85% of the cross-section of investor portfolios. Consistent with theory, portfolio tilts related to investor age and wealth drive the bulk of this common variation. We find that a three-factor model that combines the market portfolio with two long-short portfolios of investors sorted by their age and wealth explains 73% of the cross-section of investor portfolios. In contrast, the market portfolio alone only explains 28%, versus 85% for the top three principal components.

Given the important theoretical and empirical role of age and wealth, we assess the performance of an empirical pricing model consisting of the age, wealth, and market factors. We follow the standard approach of forming long-short portfolios, but use, for the first time, characteristics based on each stock's individual investor base. A stock's age characteristic is the average age of its retail owners in a particular year, weighted by the number of shares that they hold at the beginning of the year. Similarly, the wealth characteristic is the average net worth of the stock's investors, where net worth is defined as the value of financial and non-financial assets net of liabilities. The age and wealth characteristics display substantial heterogeneity across stocks and over time. We define an investor factor as a portfolio that is long stocks in the top 30% of the stock's investor characteristic and short stocks in the bottom 30%.

Consistent with our theoretical predictions, the age and wealth factors generate average returns that are strictly positive and economically significant. Their monthly CAPM alphas are 1.06% (t-value of 2.6) and 0.98% (t-value of 2.8), respectively, which correspond to yearly alphas of about 12%. These alphas cannot be explained by the most commonly used firm-based factors, such as the investment, profitability, and momentum factors. These results confirm that the stocks held by more mature and wealthier investors deliver significantly higher abnormal returns than the stocks owned by other investors.

Although age and wealth explain the bulk of common variation in investor portfolios, it is possible that other investor characteristics also contain information relevant for pricing. We investigate this possibility by constructing pricing factors from additional investor characteristics such as gender and occupational sector. We find that the age and wealth factors explain the return on all long short portfolios constructed from other stockholders' characteristics, including gender, occupations, and education. It thus appears the age and wealth factors do a good job in capturing all the pricing information contained in investor portfolio holdings.

Our three-factor investor-based model is also a strong performer out of sample. We demonstrate this property by implementing out-of-sample bootstrap tests similar to Fama and French (2018). We randomly select in-sample periods, construct tangency portfolios from the factors' in-sample return moments, and evaluate performance in the remaining out-of-sample-periods. Our investor-based model produces an average out-of-sample annualized Sharpe ratio of 0.66. As a comparison, the Sharpe ratio of the Norwegian market is 0.32. Moreover, the 0.66 out-of-sample Sharpe ratio of our three-factor model exceeds the 0.5 average Sharpe ratio of the six-factor model including the market and the size, value, investment, profitability, and momentum factors. Using fixed factor weights as advocated by DeMiguel et al. (2009) delivers similar results.

We next study how investors adjust their exposures to the age and wealth factors over the life-cycle and across the wealth distribution. To avoid any mechanical correlations arising from an investor featuring both in the construction of the investor factors and in the measurement of exposures to the same factors, we partition our sample of investors into two randomly chosen groups. We define the age and wealth factors using one group and measure the factor tilts of investors in the other group. The factor tilts of investors in the second group vary with age and wealth as one would expect. The results hold even among investors in their first year of direct stock market participation, which indicates that the migration in portfolio tilts is not primarily driven by factor loadings of firms that evolve over time. Instead, investors progressively adjust their stockholdings and therefore their factor tilts over the life cycle.

To understand the drivers of this life cycle migration, we regress the age and wealth factor tilts on a set of investor characteristics. We find that the effects of age and wealth on the factor tilts are robust to the inclusion of controls. Investors with high income beta to GDP growth and high debt-to-income ratio also tilt away from these factors, which is consistent with hedging demands. Additionally, investors prone to sentiment, such as men or investors with little stock market experience, no business education, or no professional experience in finance, also tilt their portfolio away from the age and wealth factors. Echoing the recent survey results in Choi and Robertson (2020) and Giglio et al. (2020), our results suggest that hedging and sentiment jointly drive factor tilts. Our findings are also in line with Kozak, Nagel, and Santosh (2018), who show that hedging and sentiment channels can generate pricing factors that are observationally equivalent to each other.

We gain additional insights into the nature of investor factor tilts by analyzing the characteristics of firms that make up the age and wealth factor portfolios. Relative to other investors, mature and wealthy investors tend to hold stocks with large market capitalizations, high book-to-market ratios, high profitability, low investment, and low CAPM betas. These tilts are similar to those of U.S. institutions reported in Koijen and Yogo (2019). We also document large differences in firm characteristics that prior literature typically associates with sentiment (Baker and Wurgler, 2006; Stambaugh and Yuan, 2017). Young and less wealthy investors are more likely to hold volatile stocks with high share turnover and low institutional ownership. These are the stocks about which investors disagree the most and in which arbitrage can be limited.

Our paper contributes to the extensive literature on the cross-section of equity premia (Cochrane, 2011; Harvey, Liu, and Zhu, 2015) by extracting rich pricing information from the stock holdings of individual investors. Constructing equity factors from the investor portfolios benefits from tying equity pricing directly to investor risks, preferences, and biases. By contrast, firm characteristics may be more informative about firm production decisions that micro-found production-based asset pricing models. Thus, both types of factors can provide useful complementary information about the sources of equity premia, as we show in the data. In fact, both categories of factors are expected to theoretically price the cross-section of stock returns since asset prices are determined by both investor and firm characteristics in general equilibrium (see e.g., Betermier, Calvet, and Jo, 2020).

Our results are particularly relevant for the growing research on the interaction of investor portfolio holdings and asset prices. The contributions of Koijen and Yogo (2019) and Koijen, Richmond, and Yogo (2020b) identify the types of institutional investors that have the strongest price impact in equity markets. Other studies use institutional holdings to examine the allocation of interest rate risk (Hoffmann et al., 2018), currency risk (Maggiori et al., 2020), and the transmission of monetary policy (Carpenter et al., 2015; Koijen et al., 2020a). We adopt a different focus by examining the direct stockholdings of individual investors. We do not study the price impact of their trades but show instead that their stock portfolios contain useful information about the cross-section of equity premia.

Our paper also builds on recent advances in the household finance literature. Household portfolios are known to exhibit a high degree of heterogeneity, which has motivated a wealth of empirical and theoretical explanations. The contemporaneous paper by Balasubramaniam et al. (2020) documents a factor structure in the stock portfolios of Indian investors. We document new facts about the factor structure using a broad set of investor characteristics and show the factor structure of holdings has important implications for equity pricing.[3]

Finally, we contribute to the literature at the intersection of household finance and macroeconomics documenting how heterogeneity in household portfolio returns impacts wealth inequality. Our result that wealthy investors earn higher average returns in equity markets is consistent with the findings in Bach, Calvet, and Sodini (2020) and Fagereng et al. (2020). We show that the return differential can be explained by heterogeneous exposures to a common pricing factor, which is informative about the sources of differences in performance across investors.

The rest of the paper unfolds as follows. Section II develops the theoretical framework linking the cross-section of investor portfolios to equity factors. Section III presents the data and empirical evidence on the factor structure in portfolio holdings. Section IV constructs investor pricing factors and assesses their ability to price the cross-section of stock returns. Section V studies the drivers of investor portfolio tilts toward the new factors. Section VI concludes. An online appendix provides proofs and additional empirical results.

II. Theoretical Linkages Between Investor Portfolios and Pricing Factors

We begin by presenting a framework that maps the cross-section of investor portfolio tilts into the cross-section of stock returns. We then show that, for hedging and behavioral reasons, investor age and wealth are key drivers of portfolio heterogeneity and equity premia.

[3]Using data on different U.S. institutional types, Büchner (2020) also finds evidence of commonality in investor demand.

A. Linking Pricing Factors to Aggregate Portfolio Tilts

We consider a financial market with a risk-free asset, risky stocks $j \in \{1, \cdots, J\}$, and investors $i \in \{1, \cdots, I\}$. We focus on the equilibrium at a particular point in time and do not use a time subscript in this section for expositional convenience. We denote by R_f the risk-free rate, by $\boldsymbol{R^e}$ the J-dimensional column vector of excess stock returns, and by $\mathbf{1}$ the J-dimensional column vector with all components equal to unity. It is also convenient to define the vector of expected stock returns, $\boldsymbol{\mu} = \mathbb{E}(\boldsymbol{R^e}) + R_f\,\mathbf{1}$, and the variance-covariance matrix of stock returns, $\boldsymbol{\Sigma}$.

The tangency portfolio

$$\boldsymbol{\tau} = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f\,\mathbf{1})}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f\,\mathbf{1})} \tag{1}$$

is the portfolio of stocks with the highest Sharpe ratio. The market portfolio $\boldsymbol{m}$ is the portfolio of the J stocks weighted by market capitalization. The tangency portfolio and the market portfolio have expected returns μ_τ and μ_m and volatilities σ_τ and σ_m, respectively.

Each investor i invests the nominal wealth E^i in stocks. The vector of weights in her equity portfolio is given by $\boldsymbol{\omega^i} \in \mathbb{R}^J$, where $\mathbf{1'}\boldsymbol{\omega^i} = 1$. The investor can also invest in the riskless asset, but her safe investments play a lesser role in the analysis.

Building on the recent empirical findings of Balasubramaniam et al. (2020), we assume that the cross-section of investor portfolios $\boldsymbol{\omega^i}$, $i \in \{1, \ldots, I\}$, has the following factor structure:

$$\boldsymbol{\omega^i} = \boldsymbol{\tau} + \sum_{k=1}^{K} \eta_k^i\, \boldsymbol{d_k} + \boldsymbol{u_i}, \tag{2}$$

where $\boldsymbol{d_k}$ denotes a portfolio factor, η_k^i is the investor's loading on $\boldsymbol{d_k}$, and $\boldsymbol{u_i}$ is an idiosyncratic tilt. In order to guarantee the additivity condition $\mathbf{1'}\boldsymbol{\omega^i} = 1$, we assume that the portfolios $\boldsymbol{d_k}$ and $\boldsymbol{u_i}$ are zero-investment portfolios: $\mathbf{1'}\boldsymbol{d_k} = 0$ and $\mathbf{1'}\boldsymbol{u_i} = 0$. The idiosyncratic tilts add up to zero: $\sum_{i=1}^{I} \boldsymbol{u_i} = \mathbf{0}$.

The portfolio factors $\boldsymbol{d_k}$, $k \in \{1, \ldots, K\}$, describe the common directions along which investor portfolios deviate from the tangency portfolio. For this reason, we refer to them as *deviation portfolios.* As we explain in the next Section, these portfolios can originate from hedging or sentiment motives. The additional portfolios $\boldsymbol{u_i}$ denote idiosyncratic deviations from the tangency portfolio that are unrelated to the deviation portfolios. These tilts reflect may stem from preferences or forms of inertia that are specific to each investor.

Market clearing imposes that the aggregate portfolio of investors coincides with the market portfolio of stocks: $\sum_{i=1}^{I} E^i \boldsymbol{\omega^i} / \sum_{i=1}^{I} E^i = \boldsymbol{m}$. The aggregation of individual stock portfolios (2) implies that

$$\boldsymbol{m} = \boldsymbol{\tau} + \sum_{k=1}^{K} \eta_k^m \, \boldsymbol{d_k}, \tag{3}$$

where $\eta_k^m = \sum_{i=1}^{I} E^i \eta_k^i / \sum_{i=1}^{I} E_i$ is the aggregate tilt toward the deviation portfolio $\boldsymbol{d_k}$. We assume without loss of generality that $\eta_k^m \geq 0$ for every k.[4]

Let $f_0 = \boldsymbol{m' R^e}$ denote the excess return on the market portfolio and for every k, let $f_k = \boldsymbol{d_k'} \boldsymbol{R^e}$ denote the return on the k^{th} deviation portfolio. Market clearing and equations (1) and (3) imply that the vector of factors $\boldsymbol{f} = (f_0, \ldots, f_K)'$ prices the cross-section of stock returns.

Proposition 1. *The average excess return on every stock j satisfies*

$$\mu_j - r_f = \boldsymbol{\beta_j'} \, \mathbb{E}(\boldsymbol{f}) \tag{4}$$

where $\boldsymbol{\beta_j}$ is the vector of linear regression coefficients of stock j's return on the factors.

Equation (3) and Proposition 1 show a direct connection between priced factors and aggregate tilts. In the special case where the aggregate tilts η_k^m are all equal to zero, the market is the only priced factor and the standard CAPM holds. By contrast, if investors exhibit a positive aggregate tilt ($\eta_k^m > 0$), the factor f_k is also priced. This result is a direct consequence of market clearing and therefore holds regardless of whether the tilt $\boldsymbol{d_k}$ is risk-based or sentiment-based.[5]

A deviation portfolio $\boldsymbol{d_k}$ generates CAPM-alpha if $\eta_k^m \neq 0$. Let $b_{m,j}$ denote stock j's univariate beta to the market portfolio, and let $a_j = \mu_j - r_f - b_{m,j}(\mu_m - r_f)$ denote its CAPM alpha. In the Appendix, we show that

$$a_j = -\phi \sum_{k=1}^{K} \eta_k^m \, \sigma_k^2 \, (b_{k,j} - b_{k,m}), \tag{5}$$

[4] Otherwise we replace $\boldsymbol{d_k}$ by -$\boldsymbol{d_k}$ and η_k^i by -η_k^i for every investor i in equation (2).

[5] Fama and French (2007) obtain a similar result in a simple framework that relates asset prices to disagreement and tastes.

where $\phi = (\mu_\tau - r_f)/\sigma_\tau^2$ is a positive constant, σ_k is the volatility of the k^{th} deviation portfolio, and $b_{k,j}$ and $b_{k,m}$ are, respectively, the univariate betas of stock j and the market relative to the $\boldsymbol{d_k}$. The difference $(b_{k,j} - b_{k,m})$ measures the stock's exposure to the deviation portfolio *net* of the market's exposure. If this difference is positive, the stock earns *negative* alpha. The stock is in high demand so it trades at a premium relative to the CAPM.

In addition to having a negative alpha, a stock with high exposure to the deviation portfolio $\boldsymbol{d_k}$ tends to have a high market beta. In the Appendix, we show that a stock's market beta is a weighted average of its beta to the tangency portfolio, $b_{\tau,j}$, and its beta to the deviation portfolios:

$$b_{m,j} = \frac{\sigma_\tau^2}{\sigma_m^2} b_{\tau,j} + \sum_{k=1}^{K} \eta_k^m \, \frac{\sigma_k^2}{\sigma_m^2} \, b_{k,j}. \tag{6}$$

Because a stock with high exposure to the deviation portfolio $\boldsymbol{d_k}$ is in high demand, it represents a large share of the market portfolio and therefore has a high beta. In the next Sections, we will empirically verify these predictions for alpha and beta.

Proposition 1 provides a roadmap for constructing pricing factors from a cross-section of investor portfolios. Consider a set of investor weights $z_{1,1}, \ldots, z_{I,1}$, where $\sum_{i=1}^{I} z_{i,1} = 0$. We construct a zero-investment portfolio of stocks as follows:[6]

$$\boldsymbol{g_1} = \sum_{i=1}^{I} z_{i,1} \, \boldsymbol{\omega^i}. \tag{7}$$

The portfolio $\boldsymbol{g_1}$ has several appealing properties. By (2), its loading on the tangency portfolio is zero. Moreover, assuming that it provides sufficient diversification so that $\sum_{i=1}^{I} z_{i,1} \, \boldsymbol{u_i} \approx 0$, $\boldsymbol{g_1}$ can be expressed as a linear combination of the deviation portfolios:

$$\boldsymbol{g_1} = \sum_{k=1}^{K} \left(\sum_{i=1}^{I} z_i \eta_k^i \right) \boldsymbol{d_k}. \tag{8}$$

If investor portfolios are sufficiently heterogeneous, we can construct K linearly independent portfolios $\boldsymbol{g_1}, \ldots, \boldsymbol{g_K}$ from different sets of investor weights. By (8), these portfolios fully span the deviation portfolios $\boldsymbol{d_1}, \ldots, \boldsymbol{d_K}$.[7] Consequently, the returns on the market portfolio

[6] The property that $\boldsymbol{g_1}$ is a zero-investment portfolio follows from the fact that $\mathbf{1}'\boldsymbol{g_1} = \sum_{i=1}^{I} z_{i,1} \, \mathbf{1}'\boldsymbol{\omega^i} = \sum_{i=1}^{I} z_{i,1} = 0$.

[7] The linear subspace generated by $\boldsymbol{g_1}, \ldots, \boldsymbol{g_K}$ coincides with the linear subspace generated by the devi-

$\boldsymbol{m}$ and the portfolios $\boldsymbol{g_1}, \ldots, \boldsymbol{g_K}$ price the cross-section of stocks.

We make several observations about the empirical strategy. To construct the pricing portfolios $\boldsymbol{g_1}, \ldots, \boldsymbol{g_K}$ from investor portfolio data, it is not necessary to include every single stock market investor. It is neither necessary to use a representative subset of investors. So long as holdings are sufficiently heterogeneous and provide sufficient diversification of idiosyncratic tilts, the empirical strategy highlighted above will be instructive about the pricing factors. This point suggests that the direct portfolio holdings of individual investors may contain valuable information about equity factors even when these investors own a modest fraction of aggregate market capitalization.

In practice, the investor weights $z_{i,1}$ can be chosen as a function of observable investor characteristics that are likely to be correlated to portfolio tilts. For example, if investor age drives deviations from the tangency portfolio, one would assign a positive weight to all investors above a given age threshold and a negative weight to all investors below this threshold. The resulting portfolio $\boldsymbol{g_1}$ is long the portfolios of mature investors and short the portfolios of young investors. In the next Section, we show that two investor characteristics, age and wealth, are prime candidates for constructing investor-based equity factors.

B. *Main Directions of Investor Portfolio Heterogeneity*

To examine which investor characteristics are most likely to produce investor factors, we consider two complementary models of portfolio choice. We first derive a standard rational ICAPM model in the style of Merton (1973) and Breeden (1979) populated by investors with heterogeneous ages and income profiles. Second, we consider a model with sentiment in the spirit of Fedyk et al. (2013) and Sandroni (2000). The models endogenize the factor structure of portfolio tilts and connect them to investor characteristics.

Case 1: Hedging. We consider an overlapping generations economy populated by heterogeneous investors indexed by i. Time is discrete.[8] Every period, investors can invest in a short-term bond with risk-free rate R_f and in stocks with excess returns $R^e_{1,t+1}, \ldots, R^e_{J,t+1}$. The conditional distribution of the return vector $(R_f, R^e_{1,t+1}, \ldots, R^e_{J,t+1})$ at date t is driven by a state vector $\boldsymbol{y_t}$ that follows a first-order Markov process. In applications, the state vector $\boldsymbol{y_t}$ may for instance follow a vector autoregression. Consistent with the original ICAPM

ation portfolios, or more compactly $\text{Span}[\boldsymbol{g_1}, \ldots, \boldsymbol{g_K}] = \text{Span}[\boldsymbol{d_1}, \ldots, \boldsymbol{d_K}]$.

[8]The Appendix develops a continuous-time version of the model.

(Merton, 1973), the distribution of asset returns and the state vector $\boldsymbol{y}_t$ are exogenous to the model.

An investor i is born in period b_i and lives until period $b_i + T$. She receives an initial endowment $W^i_{b_i}$ and labor income $L^i_{b^i}$ in period $t = b_i$. In all subsequent periods, the investor receives the non-financial income L^i_t, which grows at the stochastic rate $g_{t+1} = L^i_{t+1}/L^i_t$. We assume for simplicity that the income growth rates $\{g_{t+1}\}$ are common to all investors, independent through time, and do not depend on past realizations of labor income.

In every period t, the investor selects the portfolio of stocks $\boldsymbol{\alpha^i_t}$ and the consumption level C^i_t that maximize expected utility $\mathbb{E}_{b_i}\left[\sum_{t=b_i}^{b_i+T} \delta^{t-1} u(C_t)\right]$ subject to the budget constraint

$$W^i_{t+1} = L^i_t\, g_{t+1} + (W^i_t - C^i_t)\left(1 + R_f + \sum_{j=1}^{J} \alpha_{j,t} R^e_{j,t+1}\right). \tag{9}$$

The value function $J(t, W^i_t, L^i_t, \boldsymbol{y_t})$ satisfies the Bellman equation

$$J(t, W^i_t, L^i_t, \boldsymbol{y_t}) = \max_{\{\boldsymbol{\alpha_t}, C_t\}} \left[u(C^i_t) + \delta\, \mathbb{E}_t J(t+1, W^i_{t+1}, L^i_{t+1}, \boldsymbol{y}_{t+1})\right] \tag{10}$$

subject to the budget constraint (9). The optimal portfolio of stocks, $\boldsymbol{\omega^i_t} = \boldsymbol{\alpha^i_t}/(\boldsymbol{1'\alpha^i_t})$, is a function of age, wealth, and labor income:

$$\boldsymbol{\omega^i_t} = \boldsymbol{\tau_t} + \boldsymbol{d}(A^i_t, W^i_t, L^i_t, \boldsymbol{y_t}), \tag{11}$$

where $A^i_t = t - b_i$ denote the investor's age at date t. In the Appendix, we derive the relation between $\boldsymbol{\omega^i_t}$ and the value function. We verify that the deviation portfolio is zero for an investor in the last investment period ($A^i_t = T - 1$) with a labor income-to-wealth ratio equal to 0.

If the utility function is CRRA, $u(C) = C^{1-\gamma}/(1-\gamma)$, the deviation portfolio can be directly expressed in terms of the income-to-wealth ratio: $\boldsymbol{d}(A^i_t, L^i_t/W^i_t, \boldsymbol{y_t})$. We apply a Taylor expansion to the deviation portfolio around the last investment period ($A^i_t = T - 1$) and a labor income-to-wealth ratio equal to 0, and obtain the portfolio factor structure:

$$\boldsymbol{d^i_t} = (T - 1 - A^i_t)\, \boldsymbol{d_{1,t}} + \frac{L^i_t}{W^i_t} \boldsymbol{d_{2,t}},$$

where $\boldsymbol{d_{1,t}}$ and $\boldsymbol{d_{2,t}}$ are deviation portfolios. The investor's time horizon and income-to-wealth ratio drive the magnitude of portfolio deviations from the tangency portfolio. The model predicts that the portfolios of mature and wealthy investors should be closer to the tangency portfolio and therefore earn higher CAPM alphas than the portfolios of younger and less wealthy investors.

This example illustrates that an ICAPM model with heterogeneous investors naturally generates a factor structure of investor portfolios. Furthermore, the dimensionality of the factor structure is solely driven by the dimensionality of the investor characteristics that drive portfolio choice. Quite strikingly, the rank of the factor structure does not depend on the state vector $\boldsymbol{y_t}$. The model can be extended by considering additional forms of heterogeneity, such as a different income process before and after retirement, or heterogeneity in risk aversion, which would produce richer portfolio factor structures.

Case 2: Sentiment. Deviations of investor portfolios from mean-variance efficiency can also originate from sentiment. Investors may choose inefficient stock portfolios because they overreact to recent returns. They may also adjust their portfolios to forms of public information that do not impact the composition of the tangency portfolio, or they may over- or under-estimate the impact of these data on the tangency portfolio.

While the literature on sentiment is extensive (see Hirshleifer (2015) for a survey), many studies emphasize that the strength of sentiment co-varies with two key variables: age and wealth. Age is generally associated with a reduction in the size of inefficiencies. Young investors tend to be prone to fads and invest in bubbly stocks (Greenwood and Nagel, 2009). As they age, they accumulate experience on the outcomes of past decisions, learn from past mistakes, and end up making more efficient decisions (Seru, Shumway, and Stoffman, 2010). The impact of age is also a natural consequence of Bayesian learning (Barberis, 2000; Ehling, Graniero, and Heyerdahl-Larsen, 2018; Skoulakis, 2008).

Wealth is also positively correlated with more efficient behavior (Vissing-Jorgensen, 2003). Sentiment drives portfolio allocation and therefore wealth accumulation. Over the longer run, investors with low levels of sentiment are therefore likely to be wealthier (Sandroni, 2000). This effect is especially strong in general equilibrium in the presence of multiple assets, as Fedyk, Heyerdahl-Larsen, and Walden (2013) show. There is also empirical evidence that wealthier investors hold financial portfolios with higher Sharpe ratios (Calvet, Campbell, and Sodini, 2007).

These considerations motivate the following reduced-form model:

$$\boldsymbol{\omega_t^i} = \boldsymbol{\tau_t} + \boldsymbol{d_t^i},$$

where the deviation is given by

$$\boldsymbol{d_t^i} = \boldsymbol{d}(A_t^i, W_t^i, \boldsymbol{\xi_t})$$

and $\boldsymbol{\xi_t}$ is the common information set driving portfolios. If more mature investors with large amounts of wealth converge to the tangency portfolio, a simple linearization implies that

$$\boldsymbol{d_t^i} = (T - 1 - A_t^i)\, \boldsymbol{d}_{1,t} + \frac{1}{W_t^i}\, \boldsymbol{d}_{2,t} \tag{12}$$

in every period t.[9] Since investor age and wealth capture sentiment, constructing long-short portfolios according to these characteristics will allow us to recover the factor structure.

To sum up our theoretical discussion, investor factors price stock returns and can be constructed from any large and diverse dataset of investor holdings. The selected group of investors does not need to include every single stock market investor as long as the dispersion in holdings is informative about the drivers of portfolio tilts. Investor age and wealth are two prime candidates for constructing investor-based equity factors because they capture a combination of hedging and behavioral effects. In the next Section, we apply this factor extraction methodology to a high-quality dataset of Norwegian retail investors.

III. Data and Factor Structure of Portfolio Holdings

A. Data

Our analysis combines several sources of data on Norway's stock market. We obtain from the Norwegian Central Securities Depository (VPS) the complete record of stock ownership from 1996 to 2017 at Norway's only regulated market for securities trading, the Oslo Stock Exchange (OSE). For each security listed on the exchange, we observe the anonymized personal identification number of its owners and the number of shares that each owner holds annually. Individual investors are classified in the VPS database as investors with

[9]Specifically, consider the function $\boldsymbol{d^*}(\tau, v, \boldsymbol{\xi_t}) \equiv \boldsymbol{d}(T - 1 - \tau, 1/v, \boldsymbol{\xi_t})$. The linearization of d^* around $(0, 0, \boldsymbol{\xi_t})$ implies (12), where $\boldsymbol{d}_{1,t} = \partial \boldsymbol{d^*}/\partial\tau(0, 0, \boldsymbol{\xi_t})$ and $\boldsymbol{d}_{2,t} = \partial \boldsymbol{d^*}/\partial v(0, 0, \boldsymbol{\xi_t})$.

a non-professional investor account. On average, 365,000 individual investors directly hold OSE-listed stocks each year. A stock has a median number of 1,560 individual investors.

We obtain the demographic and financial characteristics of individual investors from Statistics Norway (SSB). The financial information is collected by the Norwegian Tax Administration and includes a complete breakdown of individuals' balance sheets. This information is collected annually for tax purposes, which means that banks and other third parties are legally required to provide this information to the Tax Administration. Using the personal identification numbers, we merge the SSB data with the stock ownership data in order to track the owners and their socioeconomic characteristics for each stock listed in Norway in 1996-2017. We restrict the sample to investors who file a tax return and are at least 18 years old, the minimum age required to open a personal trading account, and have a minimum liquid financial wealth of 10,000 Norwegian kroner (NOK) at the end of the calendar year. For international comparison, 1 Norwegian krone traded at 0.122 U.S. dollar on December 29, 2017.

Monthly ticker prices, market capitalizations, and information about all corporate events are available from the OSE for our 1996-2018 sample period. We complement this information with accounting data collected by the Norwegian School of Economics (NHH) for 1996-2011 and Thomson Reuters Worldscope (TRW) for 2012-2017.[10] The NHH data provides us with broader coverage than TRW in the beginning of the sample. The TRW contains information about the fraction of free-floating shares (item NOSHFF) from 1997 onward. Free-float adjusted market values ensure that our sample is not dominated by a few large companies predominantly controlled by the Norwegian government.[11]

Our analysis is based on OSE-listed stocks that satisfy the following requirements at the end of June of each year. Following the common practice in the literature, we require stocks to have at least 12 months of return history, non-missing common equity as of December 31 of the previous year, and a share price above 1 NOK in the month of portfolio formation. Our universe includes 484 unique stocks in 1997-2018 with an average of 178 firms per year, which is typical for a European stock market. To ensure that our results are not driven

[10] Link: https://www.nhh.no/en/library/databases/

[11] The government owns a substantial fraction of a few large companies: Equinor ASA (67%, energy), Norsk Hydro (34%, energy), Telenor (54%, telecommunications), DnB (34%, banking), Entra (22.4%, real estate), Yara (36%, chemicals) and Kongsberg gruppen (50%, technology). Data on government ownership is available here: https://www.regjeringen.no/no/tema/naringsliv/statlig-eierskap/selskaper---ny/id2604524/?expand=factbox2607470. By its mandate, the Norwegian sovereign wealth fund does not invest in domestic companies.

by outliers, we winsorize all monthly returns at the 99.9% level.[12] The market portfolio is defined as the value-weighted portfolio of all the stocks included in the analysis.

B. Factor Structure of Investor Portfolios

We first study the factor structure contained in the portfolio holdings of individual investors in Norway. To eliminate the idiosyncratic tilts present in individuals' portfolio holdings, we construct diversified portfolios for a large number of investor groups. Equation (2) predicts that, for each group $g \in \{1, ..., G\}$ that is sufficiently diversified, the weight of stock j in the group's aggregate portfolio is a linear combination of $K+1$ factors:

$$\omega^g_{j,t} = \frac{\sum_{i \in g} E^i_{j,t}}{\sum_{i \in g} E^i_t} = \tau_{j,t} + \sum_{k=1}^{K} \eta^g_k \, d_{k,j,t}, \tag{13}$$

where $\eta^g_k = \sum_{i \in G} E^i \eta^i_k / \sum_{i \in G} E_i$ is the aggregate tilt toward the deviation portfolio $\boldsymbol{d_k}$.

We form investor groups from a broad and diverse set of investor characteristics and require that each group contains more than 10,000 investors in a year to ensure sufficient diversification. Motivated by the theoretical discussion in Section II.B, we form 10 groups of investors each year sorted by age. The first group includes all investors below 30, the next eight groups are set in five-year increments, and the last group includes all investors above 70. We also form 12 groups of investors based on their net worth percentile each year. Net worth is defined as the sum of the investor's liquid financial wealth, real estate, vehicles, and business assets, net of liabilities.[13] The groups include the first 9 deciles of the net worth distribution (groups 1-9), the 90^{th}-99^{th} percentiles (group 10), the 99^{th}-99.9^{th} percentiles (group 11), and the 99.9^{th}-100^{th} percentiles (group 12).[14]

In addition to the 22 age and wealth groups, we form another 12 groups of investors each

[12]As a result, all winsorized stock returns are less than 154% per month.

[13]Non-traded assets include private dwellings, holiday houses, boats, vehicles, forestland, farmland, and other real capital, machinery and equipment, house contents and movables, and real assets held abroad. Liquid financial wealth includes stocks, mutual funds, money market funds, and bank account balances. Financial assets are evaluated at market prices. Other assets are evaluated by using assessed tax values for the 1997-2009 period and estimated market values from 2010 onward.

[14]The net worth distribution is based on the entire Norwegian population that are 18-100 years in a given year and have at least 10,000 NOK of liquid financial wealth. It is therefore not limited to the sample of investors with an investment account. Similar filters have been advocated by for example Fagereng et al. (2017) in their analysis of portfolio choice in Norway. We relax the minimum requirement of 10,000 investors for the top wealth group in order to capture the investment preferences of high-wealth investors.

year based on their permanent real income,[15] two groups of investors based on their gender, three groups based on their highest level of education (high school or lower, bachelor's degree, master's degree or higher), nine groups based on their field of study (e.g. management, science...), 19 groups based on their occupation, and 26 groups based on their region. The details of these groups are provided in the Appendix. Our method yields a total of 93 diversified portfolios each year. Because the coverage of occupational information is of lower quality prior to 2002, we use 2002-2017 for this analysis.

We assess the number of common factors present in the 93 aggregate portfolios by conducting a principal component analysis. For each year t, we construct a $J_t \times G$ matrix of portfolio weights, $\boldsymbol{\Omega_t}$, where J_t is the number of stocks in year t. Each column in $\boldsymbol{\Omega_t}$, which we denote by $\boldsymbol{\omega_t^g}$, adds up to one and represents the portfolio decisions of a particular group g. The principal component decomposition of $\boldsymbol{\Omega_t}$ gives:

$$\boldsymbol{\Omega_t} = \boldsymbol{F_t}\,\boldsymbol{\Lambda_t}, \tag{14}$$

where $\boldsymbol{F_t}$ is a $J_t \times G$ of orthogonal factors and $\boldsymbol{\Lambda_t}$ is $G \times G$ matrix of factor loadings.

Figure 1 reports the average cumulative proportion of the cross-sectional variance of aggregate portfolio holdings explained by the principal components (PCs), where the average is calculated across 2002-2017. The figure reveals a strong factor structure: a small number of PCs is sufficient to explain the cross-section of investor portfolios. The first PC explains 74% of the average cross-sectional variation whereas adding two more PCs explains 85%.

[Insert Figure 1 here]

C. *Linking the Factor Structure to the Market, Age, and Wealth Portfolios*

We next examine the extent to which a three-factor model consisting of the market portfolio and the age and wealth portfolio tilts captures the common factors in investor portfolios.

Building on the proposed factor construction approach given by (7), we construct two zero-weight portfolios of investors sorted by age and wealth. The first portfolio, denoted by $\boldsymbol{g_{age,t}}$, takes a long position in the portfolios of investors who are 60 and older (age groups 8,

[15]The percentiles are defined in a similar way as those of the wealth groups. Permanent income at year t is calculated as the average real earnings over the 5-year period between years $t-6$ to $t-2$. The details of the calculation are given in the Appendix.

9, and 10, with equal weights assigned to each group) and a short position in the portfolios of the youngest investors age 30 and lower (age group 1). The second portfolio, denoted by $\boldsymbol{g_{wealth,t}}$, takes a long position in the portfolios of 10% wealthiest investors (wealth groups 10, 11, and 12 equally-weighted) and a short position in the portfolios of the 30% least wealthy (wealth groups 1, 2, and 3 equally-weighted).

For each of the PC portfolios in the matrix $\boldsymbol{F_t}$, we run a pooled OLS regression of the PC's loading on stock j, $f^n_{j,t}$, on the weights of the market, age, and wealth portfolios:

$$f^n_{j,t} = a^n + \lambda^n_{mkt}\, m_{j,t} + \lambda^n_{age}\, g_{age,j,t} + \lambda^n_{wealth}\, g_{wealth,j,t} + \epsilon^n_{j,t}. \tag{15}$$

This regression is estimated over the full sample period.

Table 1 reports the three-factor model's ability to explain the cross-section of portfolio holdings for several specifications of (15). In the top row, we report the average proportion of the cross-sectional variance of the 93 portfolio holdings explained by each PC from Figure 1. In the next set of rows, we report the cumulative proportion of the cross-sectional variance explained by (i) each PC, (ii) the projection of the PC on the market portfolio as given by an univariate specification of (15), and (iii) the projection of the PC on the market, age, and wealth portfolios as given by (15). These proportions are obtained by summing up the products of (i) the proportion of the cross-sectional variance in investor portfolios explained by each PC, and (ii) the R-square of (15).

[Insert Table I here]

The results in Table 1 confirm that the market portfolio does not accurately summarize the portfolio holdings of individual investors. Alone, the market portfolio only explains 28% of the total variation in portfolio holdings through the top 10 PCs.

However, including the age and wealth portfolios into the estimation of (15) significantly improves the fit. Together with the market, these portfolios explain 71% of the total variation in portfolio holdings through the top 3 PCs. The remaining variation explained by these factors is small – only 2% once we account for PCs 4 to 10. In total, the market, age, and wealth portfolios explain 73% of the total variation in portfolio holdings. In the Appendix, we obtain similar results when we estimate (15) for each of the 93 portfolios instead of the PC portfolios. The evidence altogether shows that the market portfolio combined with two aggregate tilt portfolios - age and wealth - successfully capture the main forms of common variation in portfolio holdings.

IV. Investor Factors

We now assess the performance of the pricing model consisting of the age, wealth, and market factors. Section IV.A describes the construction of the age and wealth pricing factors and the parallel construction of firm factors that serve as benchmarks. Section IV.B investigates the return properties of the age and wealth factors. Section IV.C evaluates whether the age, wealth, and market factors capture all the relevant pricing information contained in investor portfolio holdings. In Section IV.D, we implement bootstrap simulations similar to Fama and French (2018) and compare the out-of-sample performance of investor-based and firm-based factor models.

A. Construction of Investor Factors

We develop a two-step version of the age and wealth factors that draws on the standard methodology for constructing pricing factors. In the first step, we sort stocks by a stock-specific characteristic, which will be either the weighted average age or the weighted average net worth in its individual investor base. In the second step, we construct equity factors as long-short portfolios of the sorted stocks.

This two-step method has a number of advantages for the asset pricing analysis. Statistical power increases because the two-step method puts high weight on stocks in both tails of the characteristic distribution. Comparability to firm factors improves because their construction relies on the two-step method. Finally, the two-step method provides information about the implied age and wealth characteristics for each firm.

Age. The age characteristic $\text{Age}_{j,t}$ of firm j at the end of year t is the weighted average age of the individual investors who own the firm:

$$\text{Age}_{j,t} = \frac{\sum_{i=1}^{I} N_{j,t}^{i} \, A_t^i}{\sum_{i=1}^{I} N_{j,t}^{i}}, \tag{16}$$

where A_t^i is the age of investor i and $N_{j,t}^i$ is the number of shares of stock j held by the investor at the end of year t. The age of each investor i is thus weighted by her share of the firm's equity held by retail investors, $N_{j,t}^i / \sum_{i'} N_{j,t}^{i'}$.

Figure 2 illustrates the evolution of the age characteristic for two well-known companies listed on the OSE from 1997 to 2017: Norsk Hydro (blue curve), a global aluminium company, and Nordic Semiconductor (black curve), a manufacturer of wireless devices. The age characteristic of Norsk Hydro increases from 62 years in 1997 to 67 years in 2017. By comparison, the age characteristic of Nordic Semiconductor is 50 in 1997, 48 in 2004, and 56 in 2017. More generally, the data reveal rich cross-sectional and time-series variation in the age characteristic of firms.

[Insert Figure 2 here]

Wealth. The distribution of investor net worth is fat-tailed and positively skewed, so that a few high net worth investors can heavily influence wealth-weighted averages. For this reason, our measure of a stock's wealth characteristic is based on *brackets* of investors' net worth instead of net worth itself in order to mitigate the impact of outliers. Using the classification of the 12 wealth groups defined in Section III, we denote by $\mathrm{WB}_t^i \in \{1, \ldots, 12\}$ the wealth bracket of investor i at date t. The stock's wealth characteristic is

$$\mathrm{Wealth}_{j,t} = \frac{\sum_{i=1} N_{j,t}^i \, \mathrm{WB}_t^i}{\sum_{i=1} N_{j,t}^i}, \tag{17}$$

as the weighted average of its investors' wealth bracket.[16]

Table II reports summary statistics on Norwegian investors in 2017. The average investor is 55 years old and has a net worth of 6 million NOK (about 670,000 USD). The cross-sectional standard deviation of wealth is 46 million NOK (about 5 million USD), and the wealth bracket WB_t^i defined on a 1-to-12 scale has a standard deviation of 3.

The table also reports summary statistics on stocks. A stock's investor age, $\mathrm{Age}_{j,t}$, has a cross-sectional standard deviation of 5 years. The firm's wealth characteristic, $\mathrm{Wealth}_{j,t}$, has a standard deviation of 1 on the 1-12 scale. The standard deviation of the age characteristic of firms is approximately one third of the standard deviation of age in the investor population. A similar ratio holds for wealth. These estimates confirm that the ownership base is heterogeneous across stocks.

[16] In a study of the low-risk anomaly, Bali et al. (2020) use detailed Swedish data and construct a measure of a stock's rich ownership as the proportion of the stock's shares outstanding that are directly held by individual investors in the top 10% of the wealth distribution. One important difference here is that our measure $\mathrm{Wealth}_{j,t}$ is strictly based on the wealth of investors who directly hold the stock. We do not consider the ownership share of institutional and foreign investors in the calculation. Our wealth characteristic thus allows us to compare the demand for stocks by the high and low wealth investors, irrespective of the stocks' aggregate share that is directly held by individual investors.

[Insert Table II here]

From year t to year $t+1$, we form investor factors based on the stocks' investor characteristic $C_{j,t} \in \{\text{Age}_{j,t}, \text{Wealth}_{j,t}\}$ measured at the end of year t. Specifically, for each year and each characteristic $C_{j,t}$, we sort stocks by $C_{j,t}$ and group them into three portfolios: (i) the low portfolio L containing stocks below the 30$^{\text{th}}$ percentile, (ii) the middle portfolio M containing stocks between the 30$^{\text{th}}$ and the 70$^{\text{th}}$ percentiles, and (iii) the high portfolio H containing stocks above the 70$^{\text{th}}$ percentile. Each portfolio is value-weighted by the stocks' free-float market value. The investor factor is defined as the portfolio that is long H and short L. By this definition, the age factor corresponds to a mature-minus-young portfolio, and the wealth factor to a high wealth-minus-low wealth portfolio.

We use as benchmarks a set of equity factors based on firm characteristics. We henceforth refer to these factors as *firm factors.* Following Fama and French (1992), Fama and French (1993), Fama and French (2015), Hou et al. (2018), Carhart (1997), and Novy-Marx (2013), we form the size factor (SMB_t) based on market capitalization, the value factor (HML_t) based on book-to-market ratio, the profitability factor (RMW_t) based on profit margin, the investment factor (CMA_t) based on investments, and the momentum factor (MOM_t) based on the stocks' geometric return over the previous 12 months where the most recent month is left out. For each factor, we group stocks into value-weighted portfolios based on their corresponding characteristic. The size factor goes long stocks in the top half of the size distribution and short stocks in the bottom half. The other factors go long stocks above the 70$^{\text{th}}$ percentile of the corresponding characteristic and short stocks below the 30$^{\text{th}}$ percentile.[17]

These five firm factors are sensible benchmarks for our analysis because they are known to price with reasonable precision the cross-section of stock returns around the world (Fama and French, 2012; Griffin, Ji, and Martin, 2003). Moreover, these factors are based on standard accounting and stock price information that is available for almost all stocks in our database.

B. *Return Properties of Investor Factors*

Table III, Panel A, reports the average excess returns on portfolios of stocks sorted by the age and wealth of their individual investors in 1997-2018. Average portfolio returns increase

[17]The 30$^{\text{th}}$ and 70$^{\text{th}}$ percentiles ensure that the factors are well diversified. The details of the factor construction are provided in the Appendix.

monotonically with the age and wealth characteristics. As a result, the average monthly return on investor factors is large and statistically significant: 0.96% (t-value = 2.32) for age and 0.89% (t-value = 2.52) for wealth. These monthly values correspond to average returns of 12.15% and 11.22%, respectively, in annual units. By comparison, the average monthly excess return on the market portfolio is 0.56% (t-value= 1.51) and the monthly return on firm factors ranges from -0.13% (t-value= -0.52) for the size factor to 0.85% (t-value= 2.34) for the profitability factor over the same sample period, as we report in the Appendix.

[Insert Table III here]

In Panel B of Table III, we show that the average returns on investor factors are not explained by their exposures to market portfolio risk. CAPM regressions of the age and wealth factors on the market over the sample period yield monthly intercepts that are significantly positive and equal to 1.06 (t-value = 2.58) for the age factor and 0.98% (t-value = 2.80) for the wealth factor. The age and wealth factors thus deliver significant and positive abnormal returns relative to the CAPM.

In addition to exhibiting positive alphas, investor factors both have significantly negative betas. Furthermore, Table III also reveals that the relation between the factors' market beta and average return is also negative.

These findings are in line with the theoretical analysis in Section II. Young and less wealthy investors tilt their portfolios toward stocks that provide hedging benefits or are attractive to sentiment-prone investors. In equilibrium, these attractive stocks generate negative alphas and represent a large share of the market portfolio, which in turn lead to high market betas.[18] By contrast, more mature and affluent investors tilt away from these stocks, thereby holding portfolios with positive alphas and low betas.

In Figure 3, we plot the cumulative log growth of 1 NOK invested in 1997 in either the long and short legs of the age and wealth factor portfolios. We use the market portfolio as the benchmark. Economic recessions are shaded in blue. Panel A shows that the short legs of the age and wealth factors performed well in the late 1990s but underperformed the market over the full sample. Underperformance is most pronounced after the 2008 crisis. By contrast, Panel B of Figure 3 shows that the long legs of the age and wealth factors outperformed the market throughout the sample.

[18]Equations (5) and (6) summarize this logic. These equations predict that a positive tilt toward the deviation portfolio $\boldsymbol{d_k}$ should yield a negative alpha *and* a high beta. In the context of our model, a positive exposure to the age and wealth factors can therefore be interpreted as a tilt away from $\boldsymbol{d_k}$.

[Insert Figure 3 here]

Panel C of Figure 3 illustrates the cumulative performance of the age and wealth factors. Both factors have a high average return and a low volatility. The contemporaneous return correlation between the age and wealth factors is only about 0.2, which highlights the importance of including both factors in the pricing model. Panel D of Figure 3 further illustrates the benefits of using both factors by reporting the performance of an equal-weighted portfolio of the age and wealth factors. This combined factor yields significantly higher performance than the market portfolio, while also displaying lower volatility than each factor taken separately. For this reason, we refer to this factor as the age-wealth investor factor in the following sections.

Table V shows that the strong performance of the age-wealth investor factor cannot be explained by its exposures to the firm factors. We consider the following benchmark models: i) the market i) the market, size and value factors from Fama and French (1993) (Firm-3), ii) the market, size, value, and momentum factors from Carhart (1997) (Firm-4), iii) the market, size, value, profitability, and investment factors from Fama and French (2015) (Firm-5), iv) momentum and the Fama and French (2015) factors (Firm-6).

[Insert Table V here]

For every benchmark model, we reject the null hypothesis that pricing errors are zero at the 1% significance level. The monthly intercept is as high as 1.05 when the market is the sole benchmark factor. It decreases to 0.66 when all 6 firm factors are included but remains highly significant, with a t-value of 2.79. The age-wealth investor factor therefore captures pricing information that is not contained in the traditional firm factors.

C. Pricing Information Contained in Investor Portfolio Holdings

A natural question is whether the age and wealth factors capture all the relevant pricing information contained in investor portfolio holdings. We assess the possibility of other characteristics playing a role by constructing pricing factors from the additional investor characteristics described in Section III.B. Specifically, we calculate for each investor characteristic a stock-specific characteristic $C_{j,t}$ that is the weighted average of its investors' individual characteristics:

$$\mathrm{C}_{j,t} = \frac{\sum_{i=1} N^i_{j,t}\,\mathrm{C}^i_t}{\sum_{i=1} N^i_{j,t}}. \tag{18}$$

We then sort stocks by the characteristic and form the equity factor as a long-short portfolio of the sorted stocks. We consider the following investor characteristics: dummy variables for women, retirees, permanent income brackets, labor income-to-wealth brackets, occupation, and a categorical variable equal to 1, 2, or 3 based on the investor's highest education level.

In Table IV, we regress each of these additional investor factors on the age-wealth factor described in the previous Section.

[Insert Table IV here]

The slope coefficients of the additional investor factors are statistically different from zero. The female, education, permanent income, and retirement factors load positively on the age-wealth factor whereas the labor-to-wealth factor loads negatively. The fact that these factor loadings are different from zero is re-assuring in light of the results in Section III and further confirms the presence of a factor structure in portfolio holdings. In Section V we will analyze further the sign of these loadings.

Importantly, we find that none of the additional investor factors has a significant alpha against the age-wealth factor. This result shows the additional investor-based factors are spanned by the age and wealth factors, which suggests age and wealth capture most of the pricing information contained in investor portfolio holdings.

D. Cross-validation of Investor Factor Models

We next compare the out-of-sample performance of the investor-based and firm-based factor models considered in earlier sections. We do so by constructing maximum-Sharpe-ratio portfolios based on the factors in sample and then by estimating the Sharpe ratios of these portfolios out of sample. Here an important pitfall is that a factor with an unusually high in-sample mean attracts a large weight in the estimated tangency portfolio. This overweighting tends to reduce out-of-sample portfolio performance. We control for this in-sample bias by implementing a bootstrap evaluation approach similar to Fama and French (2018).

Using our sample of $T = 264$ months of factor return observations, we run 100,000 simulations. In each simulation, we draw with replacement a new dataset of size T and refer to this sample as the training sample. The hold-out sample consists of the months not included in the training sample.[19]

We use two approaches to obtain the weights of the maximum-Sharpe-ratio portfolios. In the first approach ("fixed-weight"), we follow the methodology developed in Section IV.C and impose fixed factor weights. The weight on the market is set to 1 and the weight on each long-short factor portfolio is set to $1/2K$, where K is the total number of factor portfolios. In the case of the model with age and wealth factors, the return of the maximum-Sharpe-ratio portfolio in year t is estimated as:

$$\hat{\tau}_t = MKT_t + \frac{1}{4}Age_t + \frac{1}{4}Wealth_t. \tag{19}$$

This tangency portfolio can also be viewed as the sum of the market portfolio and 0.5 the age-wealth investor factor. The use of equal weights for constructing factor portfolios builds on the work of DeMiguel et al. (2009), who show that equal-weighted factor portfolios tend to perform best out of sample.

In the second approach ("MV-weight"), we endogenously obtain factor weights from an in-sample mean-variance optimization. To reduce the level of estimation risk, we follow Kozak et al. (2020) and shrink the covariance matrix of factor returns $\boldsymbol{\Sigma_f}$ as:

$$\hat{\boldsymbol{\Sigma_f}} = \boldsymbol{\Sigma_f} + \gamma \boldsymbol{I} \tag{20}$$

where $\boldsymbol{I}$ is the identity matrix and γ is a shrinkage parameter.[20] This shrinkage methodology adds variance to all factors, which in turn has the effect of reducing portfolio weights the most for factors with low volatility. We obtain the factor weights as:

$$\hat{\tau}_t = \hat{\boldsymbol{\Sigma_f}} \boldsymbol{\mu_f}, \tag{21}$$

where $\boldsymbol{\mu_f}$ is the vector of in-sample average returns.

For both approaches, we compute the portfolio's Sharpe ratio over the out-of-sample period. We repeat the simulation 100,000 times and calculate the average in-sample and out-of-sample Sharpe ratios for each factor model. Whereas in-sample Sharpe ratios are

[19] With large T, the expected fraction of the sample with unique observation is $1-\exp(-1) \approx 63.2\%$. The hold-out sample therefore contains approximately 97 months ($\exp(-1) \times 264$).

[20] The details are based on Kozak et al. (2020). The parameter γ, which dictates the level of shrinkage, is equal to $\gamma = tr[\boldsymbol{\Sigma_f}]/E[SR^2] \times T$. We select a value of 0.5 for the average Sharpe ratio and verify that other values of γ do not impact our results.

subject to the upward bias described above, out-of-sample Sharpe ratios are unaffected by it because monthly returns are close to being serially uncorrelated.

Table VI reports the average out-of-sample annualized Sharpe ratio for different factor models and methodologies. Factor models are grouped according to the number of factors included along with the market portfolio. On its own, the market portfolio has a Sharpe ratio of 0.32, which is consistent with typical estimates of market Sharpe ratios (Doeswijk et al., 2020).

[Insert Table VI here]

Among two-factor models, investor factors provide the highest Sharpe ratio. For fixed-weight factor portfolios in Column 1, the combination of the age and market factors generates a Sharpe ratio of 0.58. No combination of a firm factor and the market performs better. The second best factor is wealth, which generates a Sharpe ratio of 0.57 when combined with the market. The third best factor is profitability, which has a Sharpe ratio of 0.56 when combined with the market. We obtain similar results with the MV-weight factors in Column 2. Age and wealth generate out-of-sample Sharpe ratios of 0.51 and 0.54 in combination with the market, while the highest performing firm factor, profitability, only generates a Sharpe ratio of 0.49 with the market.

Among three-factor models, the combination of wealth, age, and the market again performs best, with a Sharpe ratio of 0.61 in annual units for the fixed-weight portfolio and 0.66 for the MV-weight portfolio. The highest performing three-firm-factor is the combination of market, profitability, and momentum, which generates a Sharpe ratio of 0.59 (fixed-weight) and 0.55 (MV-weight).

Adding more firm factors to the market does not change the results. The Firm-6 portfolio, which is made up of the market and five firm factors, generates an average Sharpe ratio of 0.41 (fixed-weight) and 0.5 (MV-weight). These performance results are weaker than those generated by the investor factor model consisting of the market, age, and wealth factors.

A final insight from Table VI is that investor-based models are less subject to the upward in-sample bias than firm-based models. In Column 3, we report the ratio of out-of-sample to in-sample average Sharpe ratio for each model resulting from the optimization in (21). The ratio indicates the proportion of the in-sample average Sharpe ratio that is "retained" in the out-of-sample period. Investor factor models have the highest ratio. For example, the three-factor model consisting of the market, age, and wealth factors has a ratio of 73%, whereas the highest performing three-firm-factor model has a ratio of 68%.

V. The Cross-Section of Investor Factor Tilts

The strong pricing performance of the age and wealth factors raises the question of their economic origins. Are investor deviations from the tangency portfolio driven by hedging motives, sentiment, or a combination of both channels? In this section, we investigate this question by studying how the portfolio tilts of individual investors relate to their socioeconomic characteristics. Section V.A documents how investors adjust their portfolio tilts toward investor factors as they migrate through the wealth distribution over the life-cycle. Section V.B shows that socioeconomic characteristics other than age and wealth also drive investor portfolio tilts. In Section V.C, we build a bridge between investor-based and firm-based factors by documenting the characteristics of the firms that make up investor factor portfolios, which is informative about the economic drivers of these factors.

A. *How Do Investor Portfolio Tilts Vary with Age and Wealth?*

We first document how investors adjust their portfolio tilts toward the age and wealth factors as they age and become more affluent. To break any mechanical link between portfolio tilts and investor characteristics, we partition investors into two randomly chosen groups. The first group contains two-thirds of the investor population and is used to reestimate the age and wealth factors.[21] The second group is used to study the links between portfolio tilts and characteristics.

We calculate the portfolio tilts of an investor as follows. Consider a factor with long leg H and short leg L at time t. The proportion of investor i's stock portfolio invested in equities contained in the long leg is:

$$\omega^i_{H,t} = \sum_{j=1}^{J} \omega^i_{j,t} \mathbb{1}_{j,H,t}, \tag{22}$$

where $\omega^i_{j,t}$ is the weight of stock j in investor i's stock portfolio and $\mathbb{1}_{j,H,t}$ is an indicator variable equal to unity if stock j belongs to the long leg H at time t. A similar definition

[21] In the Appendix, we verify that investor factors constructed from a subset of the investor population contain similar pricing information as the full-sample factors, albeit with lower accuracy.

provides the portfolio share invested in short leg stocks, $\omega^i_{L,t}$. We define the investor's *portfolio tilt toward the factor* by

$$\omega^i_{f,t} = \omega^i_{H,t} - \omega^i_{L,t}. \tag{23}$$

The tilt is bounded between -1 and 1. It is equal to -1 if the investor only selects stocks in the short leg, 0 if the investor allocates equal amounts of capital to the long and short legs, and +1 if the investor only selects stocks in the long leg. This definition provides a convenient and direct measure of an investor's tilt toward a factor based only on portfolio holdings at a given date t. We will refer to $\omega^i_{age,t}$ and $\omega^i_{wealth,t}$ as the investor's age and wealth factor tilts.

In Panel A of Figure 4, we plot the average portfolio tilt toward the age factor for 10 groups of investors sorted by age. The first group includes all investors below 30, the next eight groups are set in five-year increments, and the last group includes all investors above 70. Means are equally-weighted and estimated over the full 1997-2018 sample. The age tilt is less than 0.1 below age 30 and progressively increases to 0.4 for the oldest group. The panel shows a substantial and remarkably linear migration in the age factor tilt over the life-cycle.

[Insert Figure 4 here]

The "age ladder" illustrated in Panel A of Figure 4 relates to the findings in Betermier, Calvet, and Sodini (2017), who report a progressive life-cycle migration toward the value factor among Swedish households. This earlier paper shows that the linearity between the value tilt and age more likely originates from life-cycle variation in age and other characteristics than from combinations of time and cohort fixed effects. The reason is that, in order to generate such a linear structure, cohort and year fixed effects would have to offset each other exactly. The same logic applies to the age factor tilt.

In Panel B of Figure 4, we plot the average age factor tilt of investors who are new to direct stock market investing (black line).[22] The portfolio tilts chosen by new entrants closely mimic the tilts of seasoned investors of the same age. This result confirms that the age ladder unlikely emanates from portfolio inertia and migration of firm characteristics across factors. Instead, investors progressively adjust their age tilts over the life cycle.

In Panel C of Figure 4, we obtain similar results for the average tilt toward the wealth factor for 12 groups of investors sorted by net worth. The groups are described in Section IV.A. Investors progressively migrate toward the wealth factor as they climb the wealth ladder.

[22]Each point estimate contains at least 1,000 investors. Groups that do not satisfy this requirement are dropped.

This migration is again economically significant. The wealth factor tilt is as low as -0.12 for investors in the bottom 10 percentiles (first bracket) and reaches 0.03 for investors in the top 0.1 percentile (12^{th} bracket). The difference is most pronounced among the wealthiest investors.

Panel D of Figure 4 shows that new stock market entrants choose wealth factor tilts similar to those of equally wealthy pre-existing investors. Altogether, these results confirm that the factor tilts of investors vary with age and wealth as one would expect, even among investors in their first year of stock market investing.

B. Which Other Investor Characteristics Drive Portfolio Tilts?

We next use regression analysis to examine which investor characteristics predict portfolio tilts toward the age and wealth factors. Besides age and wealth, we consider a number of socioeconomic characteristics that have been shown to explain portfolio decisions in household finance research.

The first set of characteristics captures risk exposures (see e.g., Cocco, Gomes, and Maenhout, 2005; Gomes and Michaelides, 2005; Heaton and Lucas, 1997, 2000; Viceira, 2001). We measure indebtedness by the debt-to-income ratio, which captures the investor's ability to withstand economic shocks (Campbell, 2006; Iacovello, 2008). We compute the sensitivity of her non-financial income to macroeconomic risk as in Guvenen et al. (2017). To do so, we form 220 groups of investors sorted by employment sector, retirement status, and labor income percentile. For each group $g \in \{1, \dots, G\}$, we run a panel regression of the annual income growth of investor i in year t, denoted by $\Delta y_{i,t}$, on real GDP growth in the same year:

$$\Delta y_{i,t} = a_g + \beta_g^{GDP} \Delta GDP_t + \varepsilon_{i,t}. \tag{24}$$

The regression yields a slope coefficient β_g^{GDP} for each group. We assign β_g^{GDP} to all individuals in the group and use it as a proxy for their exposure to macroeconomic risk. The exact definition of the groups and estimation details are provided in the Appendix.

The second set of characteristics proxy for behavioral traits that may also affect an investor's portfolio tilts toward the age and wealth factors. The impact of stock market experience on portfolio choice has been documented in a number of empirical studies and field experiments (List, 2003; Seru, Shumway, and Stoffman, 2010). Experience is defined as the number of years the investor has held stocks. We also include a set of dummy variables corresponding to graduate education, business education, finance sector occupation, and gender. Previous research has shown that biases such as overconfidence are more prevalent among men than women (Barber and Odean, 2001) and less educated investors (Calvet, Campbell, and Sodini, 2009).

We run panel regressions of the age and wealth factor tilts on investor characteristics:

$$\omega^i_{f,t} = a + \boldsymbol{\gamma' X^i_t} + \eta_t + \epsilon^i_t, \tag{25}$$

where $\boldsymbol{X^i_t}$ is a vector of characteristics, η_t is a time fixed effect, and ϵ^i_t is the residual error term. The vector $\boldsymbol{X^i_t}$ includes the investor's debt-to-income ratio, non-financial income exposure to macroeconomic risk, gender, stock market experience, education variables, and a finance occupation dummy, as well as indicator variables for age and wealth brackets. For each factor, we consider both the ten age groups defined in Section V.A and the 12 wealth brackets defined in Section IV.A, and we use the median brackets as benchmarks. Standard errors are clustered by year and investor.

In Table VII, we report the regression results for the age factor tilt. Several characteristics explain the tilt. Age remains statistically significant for most groups after controlling for the additional characteristics. Young investors, as represented by the first five age groups, tilt away from the age factor, whereas mature investors have positive tilts.

[Insert Table VII here]

Both measures of risk exposure are negatively related to the age factor tilt. The effect of income beta is particularly strong. A 0.5 difference in income beta, which approximately corresponds to the difference between working in public administration and working in the tourism industry for an individual with median income, is associated with a 0.045 reduction in the age factor tilt.

Graduate education, business education, finance sector occupation, and stock market experience are all associated with a higher age factor tilt. In terms of economic magnitude, ten years of additional experience explain a 0.14 increase in the age factor tilt. Female investors also have a greater age factor tilt than male investors. Gender has approximately the same effect on the age factor tilt as ten years of stock market experience. These results are consistent with sentiment driving portfolio tilts.

Table VIII presents remarkably similar results for the wealth factor tilt. As with age, the explanatory power of the wealth dummy variables is robust to the inclusion of other characteristics. Less affluent investors have a negative wealth tilt, whereas more affluent investors have a significantly positive tilt.

[Insert Table VIII here]

Characteristics based on risk exposure and sentiment also explain variation in the wealth factor tilt. A 0.5 increase in the income beta is associated with a 0.025 reduction in the wealth factor tilt. Ten years of stock market experience explain a 0.07 increase in the portfolio tilt. Being female and having a graduate degree or business education also predict a higher wealth factor tilt.

Taken together, these results suggest that hedging motives and sentiment jointly drive the cross-sectional variation in investor factor tilts. On the one hand, investors with low risk exposures are in a better financial position to tilt toward the age and wealth factors than investors with high risk exposures, which is consistent with hedging demands. Moreover, investors progressively migrate toward these factors as they become more mature and wealthier, which is consistent with the predictions of our theoretical framework in Section II.B.[23]

On the other hand, the positive relations between factor tilts and measures of financial sophistication also suggest the presence of a parallel behavioral channel. Investors who are younger, less wealthy, and more prone to sentiment systematically tilt away from the age and wealth factors. These investors include men and individuals with lower educational attainment, shorter stock market experience, no business education, and no professional experience in finance. This complementary explanation is consistent with empirical evidence on correlated sentiment trades in the portfolios of retail investors (Barber et al., 2009; Kumar and Lee, 2006).[24]

[23]In the Appendix, we go one step further and verify that the wealth factor correlates with a factor constructed from investors' wealth-to-income ratio, as the ICAPM predicts.

[24]The results are also consistent with Korniotis and Kumar (2011), who find that older U.S. retail investors are better diversified, trade less frequently, invest in lower-fee funds, and exhibit weaker behavioral biases than younger investors. One difference in their study is that they find older U.S. retail investors generally performed worse than younger investors between 1991 and 1996. One possible explanation for this result is the specific period used in their analysis. Our evidence about the high performance of older investors is based on 21 years of monthly return data.

C. *Firm Characteristics of Investor-Based Factors*

To gain additional insight into the nature of investor factor tilts, we analyze the characteristics of firms that make up the age and wealth factor portfolios. This analysis provides a bridge between firm-based and investor-based factors.

We consider the following firm characteristics: size, book-to-market ratio, profitability, investment, return volatility, the proportion of equity held by institutional investors, and share turnover, defined as the number of shares traded in a year divided by the number of free-float shares outstanding at the beginning of the year. For each factor, we consider four portfolios, corresponding to the stocks in the bottom 30% of the investor characteristic (short leg L), the middle 30%-70% bracket (M), the top 30% (long leg H), and the long-short portfolio H-L defining the factor portfolio.

Table IX reports the median characteristic of each portfolio, where the median is taken in the pooled cross-section. The table highlights clear differences in the properties of stocks in the long and short legs of investor factor portfolios. Stocks held by young and less wealthy investors have significantly higher volatility, higher share turnover, and lower institutional ownership than stocks held by mature and wealthy investors. These results are consistent with prior work arguing that these types of stocks are more difficult to arbitrage and therefore more sensitive to changes in sentiment (Stambaugh and Yuan, 2017). Together with the regression results from the previous section, these findings further suggest the presence of sentiment motives driving factor tilts.

[Insert Table IX here]

Additionally, we find that stocks held by mature and affluent investors tend to have higher market capitalizations, higher profitability, lower investment, and lower CAPM betas than stocks held by the young and the less wealthy. These links are important for several reasons. First, they support Koijen and Yogo (2019)'s modeling assumption that investor portfolio holdings are related to firm characteristics. Second, they reveal that mature and wealthy investors tend to invest in the same stocks as institutional investors, which Koijen and Yogo (2019) study in the U.S. context. This finding suggests that the observed dispersion in the direct stock holdings of individual investors contains valuable information about portfolio tilts outside the retail sector.

VI. Conclusion

This paper constructs a parsimonious set of equity factors from the cross-section of individual investor portfolio holdings. We show theoretically that portfolios of stocks sorted by the age or wealth of their individual investors should produce powerful pricing factors. Using the complete stockholdings of Norwegian individual investors, we verify empirically that a three-factor model consisting of a mature-minus-young factor, a high wealth-minus-low wealth factor, and the market factor explains the bulk of common variation in portfolio holdings and prices the cross-section of stock returns. We also uncover a rich set of links between investor characteristics and portfolio tilts toward the age and wealth factors.

The analysis of investor factors opens new opportunities for equity pricing research. The tight connection between investor factors and the cross-section of portfolio holdings makes it possible to connect equity risk premia to the drivers of investor demand. Our finding that hedging motives and sentiment operate in tandem suggests that there might be interdependencies between both channels, as Kozak, Nagel, and Santosh (2018) explain.

Another interesting question is whether investor-based factors price other asset classes. This question seems important because limitations on firm accounting data may limit the statistical ability of traditional firm factors to price alternative asset classes such as private equity. Information on the characteristics of investors who own these assets provides an alternative avenue for pricing them. We leave these questions for future research.

REFERENCES

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, 2020, Rich pickings? Risk, return, and skill in household wealth, *American Economic Review* 110, 2703–2747.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Balasubramaniam, Vimal, John Y. Campbell, Tarun Ramadorai, and Benjamin Ranish, 2020, Who owns what? A factor model for direct stockholding, Working paper, Harvard University.

Bali, Turan G., A. Doruk Gunyadin, Thomas Jansson, and Yigitcan Karabulut, 2020, Do the rich gamble in the stock market? Low risk anomalies and wealthy households, Working paper, Georgetown University.

Barber, Brad, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *Review of Financial Studies* 29, 2600–2642.

Barber, Brad M., and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics* 116, 261–292.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Systematic noise, *Journal of Financial Markets* 12, 547–569.

Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *Journal of Finance* 55, 225–264.

Berk, Jonathan B., and Jules H. van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Betermier, Sebastien, Laurent E. Calvet, and Evan Jo, 2020, A supply and demand approach for equity pricing, Working paper, EDHEC Business School and McGill University.

Betermier, Sebastien, Laurent E. Calvet, and Paolo Sodini, 2017, Who are the value and growth investors?, *Journal of Finance* 72, 5–46.

Blume, Marshall E., and Donald B. Keim, 2012, Institutional investors and stock market liquidity: Trends and relationships, Working paper, The Wharton School.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

Büchner, Matthias, 2020, What drives asset holdings? Commonality in investor demand, Working paper.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, 2007, Down or out: Assessing the welfare costs of household investment mistakes, *Journal of Political Economy* 115, 707–747.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, 2009, Fight or flight? Portfolio rebalancing by individual investors, *Quarterly Journal of Economics* 124, 301–348.

Campbell, John Y., 2006, Household finance, *Journal of Finance* 61, 1553–1604.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Carpenter, Seth, Selva Demiralp, Jane Ihrig, and Elizabeth Klee, 2015, Analyzing Federal Reserve asset purchases: From whom does the Fed buy?, *Journal of Banking & Finance* 52, 230–244.

Choi, James, and Adriana Robertson, 2020, What matters to individual investors? Evidence from the horse's mouth, *Journal of Finance* 75, 1965–2020.

Cocco, Joao F., Francisco J. Gomes, and Pascal J. Maenhout, 2005, Consumption and portfolio choice over the life cycle, *Review of Financial Studies* 18, 491–533.

Cochrane, John H., 2011, Discount rates, *Journal of Finance* 66, 1047–1108.

Constantinides, Georges M., 2017, Asset pricing: Models and empirical evidence, *Journal of Political Economy* 125, 1782–1788.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009, Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.

Doeswijk, Ronald, Trevin Lam, and Laurens Swinkels, 2020, Historical returns of the market portfolio, *Review of Asset Pricing Studies* 10, 521–567.

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen, 2018, Asset prices and portfolio choice with learning from experience, *Review of Economic Studies* 85, 1752–1780.

Fagereng, Andreas, Charles Gottlieb, and Luigi Guiso, 2017, Asset market participation and portfolio choice over the life-cycle, *Journal of Finance* 72, 705–750.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, 2020, Heterogeneity and persistence in returns to wealth, *Econometrica* 88, 115–170.

Fama, Eugene F, and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–466.

Fama, Eugene F, and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 43, 3–56.

Fama, Eugene F., and Kenneth R. French, 2007, Disagreements, tastes, and asset prices, *Journal of Financial Economics* 83, 667–689.

Fama, Eugene F., and Kenneth R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

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

Fama, Eugene F., and Kenneth R. French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234–252.

Fedyk, Yurii, Christian Heyerdahl-Larsen, and Johan Walden, 2013, Market selection and welfare in a multi-asset economy, *Review of Finance* 17, 1179–1237.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2020, Five facts about beliefs and portfolios, Working paper, Stanford University.

Gomes, Francisco, Michael Haliassos, and Tarun Ramadorai, 2020, Household finance, *IMFS Working Paper Series No. 138* .

Gomes, Francisco, and Alexander Michaelides, 2005, Optimal life-cycle asset allocation: Understanding the empirical evidence, *Journal of Finance* 60, 869–904.

Greenwood, Robin, and Stefan Nagel, 2009, Inexperienced investors and bubbles, *Journal of Financial Economics* 93, 239–258.

Griffin, John M., Xiuqing Ji, and J. Spencer Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515–2547.

Guercio, Diane Del, and Paula A. Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds, *Journal of Financial and Quantitative Analysis* 37, 523–557.

Guiso, Luigi, and Paolo Sodini, 2013, Household finance: An emerging field, in Georges M. Constantinides, Milton Harris, and Rene M. Stulz, eds., *Handbook of the Economics of Finance*, volume 2.

Guvenen, Fatih, Sam Schulhofer-Wohl, Jae Song, and Motohiro Yogo, 2017, Worker betas: Five facts about systematic earnings risk, *American Economic Review* 107, 398–403.

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2015, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

He, Zhiguo, and Wei Xiong, 2013, Delegated asset management, investment mandates, and capital immobility, *Journal of Financial Economics* 107, 239–258.

Heaton, John, and Deborah Lucas, 1997, Market frictions, savings behavior, and portfolio choice, *Macroeconomic Dynamics* 1, 76–101.

Heaton, John, and Deborah J. Lucas, 2000, Portfolio choice and asset prices: The importance of entrepeneurial risk, *Journal of Finance* 55, 1163—1198.

Hirshleifer, David, 2015, Behavioral finance, *Annual Review of Financial Economics* 7, 133–159.

Hoffmann, Peter, Sam Langfield, Federico Pierobon, and Guillaume Vuillemey, 2018, Who bears interest rate risk?, *Review of Financial Studies* 32.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2018, Which factors?, *Review of Finance* 23, 1–35.

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

Iacovello, Matteo, 2008, Household debt and income inequality, 1963-2003, *Journal of Money, Credit, and Banking* 40, 929–965.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *Journal of Finance* 63, 273–310.

Kelley, Eric K., and Paul C. Tetlock, 2013, How wise are crowds? insights from retail orders and stock returns, *Journal of Finance* 68, 1229–1265.

Koijen, Ralph, Francois Koulischer, Benoit Nguyen, and Motohiro Yogo, 2020a, Inspecting the mechanism of quantitative easing in the euro area, *Journal of Financial Economics* forthcoming.

Koijen, Ralph S. J, Robert J. Richmond, and Motohiro Yogo, 2020b, Which investors matter for equity valuations and expected returns?, Working paper.

Koijen, Ralph S. J, and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Korniotis, George M, and Alok Kumar, 2011, Do older investors make better investment decisions?, *The Review of Economics and Statistics* 93, 244–265.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

Kumar, Alok, and Charles Lee, 2006, Retail investor sentiment and return comovements, *Journal of Finance* 61, 2451–2486.

Lintner, John, 1966, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

List, John A., 2003, Does Market Experience Eliminate Market Anomalies?*, *Quarterly Journal of Economics* 118, 41–71.

Ludvigson, Sydney C., 2013, Advances in consumption-based asset pricing: Empirical tests, in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds., *Handbook of the Economics of Finance* (Elsevier).

Maggiori, Matteo, Brent Neiman, and Jesse Schreger, 2020, International currencies and capital allocation, *Journal of Political Economy* 128, 239–258.

Mehra, Rajnish, 2012, Consumption-based asset pricing models, *Annual Review of Financial Economics* 4, 385–409.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.

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

Sandroni, Alvaro, 2000, Do markets favor agents able to make accurate predictions?, *Econometrica* 68, 1303–1341.

Seru, Amit, Tyler Shumway, and Noah Stoffman, 2010, Learning by trading, *Review of Financial Studies* 23, 705–739.

Sharpe, William F., 1964, Capital asset prices: a theory of market equilibrium under conditions of risk, *The Journal of Finance* 19, 425–442.

Skoulakis, Georgios, 2008, Dynamic portfolio choice with Bayesian learning, Working paper.

Stambaugh, Robert, and Yu Yuan, 2017, Mispricing factors, *Review of Financial Studies* 30, 1270–1315.

Viceira, Luis M., 2001, Optimal portfolio choice for long-horizon investors with nontradable labor income, *Journal of Finance* 433–470.

Vissing-Jorgensen, Anne, 2003, Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions, *NBER Macroeconomics Annual* 18, 139–194.

Table I
Principal Component Analysis of Investor Portfolio Holdings

This table reports the results of a principal components analysis done on 93 aggregate portfolios of individual investors grouped by socioeconomic and geographic indicators. The first row reports the proportion of the cross-sectional variance of portfolio holdings explained by each of the top 10 principal components (PCs). The cumulative proportion is calculated each year in 2002-2017 and then averaged across years. The second set of rows reports the cumulative proportion of the cross-sectional variance explained by (i) each PC, (ii) a projection of the PC on the market portfolio, and (iii) a projection of the PC on the market, age, and wealth portfolios. For (ii) and (iii), the variance proportions are calculated by taking summing up the products of the variance of each PC and the regression model's R-square. The age portfolio takes a long position in the portfolios of investors who are 60 and older and a short position in the portfolios of the youngest investors age 30 and lower. The wealth portfolio takes a long position in the portfolios of 10% wealthiest investors and a short position in the portfolios of the 30% least wealthy.

	Top Principal Components									
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Variance explained	0.74	0.07	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.01
Cumulative variance										
PC	0.74	0.81	0.85	0.88	0.90	0.92	0.93	0.95	0.95	0.96
Market	0.27	0.27	0.27	0.28	0.28	0.28	0.28	0.28	0.28	0.28
Market, Age, Wealth	0.67	0.71	0.71	0.72	0.73	0.73	0.73	0.73	0.73	0.73

Table II
Summary Statistics on Investor Characteristics

This table presents summary statistics on the age and wealth characteristics of Norwegian individuals investors holding stocks directly in 2017. We report the standard deviation, mean, and the 10^{th}, 30^{th}, 50^{th}, 70^{th}, 90^{th}, and 99^{th} percentiles of each characteristic (i) in the population of participating investors and (ii) in the population of stocks. The age characteristic of a stock is the average age of the stock's individual investor shareholders, weighted by the number of shares that they own at the beginning of the year. Investor wealth is defined as the value of liquid and illiquid assets (liquid financial wealth, real estate, vehicles, business assets) net of liabilities. Wealth is expressed both in million NOK and on a 1 to 12 scale, where the first 9 categories represent the first 9 deciles of the wealth distribution, the 10^{th} category the 90-99^{th} percentiles, the 11^{th} category the 99-99.9^{th} bracket, and the 12^{th} category the top 0.1%. A stock's wealth characteristic is the average wealth of its individual investor shareholders, weighted by the number of shares that they hold at the beginning of the year. The details of variable construction are provided in Section IV.A of the main text.

			Percentiles					
Characteristic	SD	Mean	10^{th}	30^{th}	50^{th}	70^{th}	90^{th}	99^{th}
Age:								
Household-level	16.5	55.1	32.0	46.0	56.0	65.0	76.0	89.0
Stock-level	4.9	57.2	51.4	54.5	56.6	59.6	64.6	67.0
Wealth:								
Investor-level (MNOK)	46.6	6.0	-0.0	1.5	3.1	5.2	10.9	48.1
Investor-level (Rank 1-12)	1.0	8.7	7.4	8.4	8.9	9.3	9.9	10.6
Stock-level (Rank 1-12)	2.9	7.0	2.0	6.0	8.0	9.0	10.0	11.0

Table III
Return Performance of Age and Wealth Factors

This table reports statistics on the return performance of the age and wealth investor factors constructed from the universe of Norwegian stocks in 1997-2018. Panel A reports monthly value-weighted average excess returns for the low-, medium-, and high- portfolios for each investor characteristic. These portfolios correspond to the bottom 30%, mid 40%, and top 30% of stocks sorted by the investor characteristic. The investor factor is defined as high minus low. Panel B and Panel C report, respectively, the intercept and the slope coefficient of times-series OLS regressions of monthly excess portfolio returns on the market factor.

Panel A: Monthly Returns								
	Average Return				t(Average Return)			
	L	M	H	H-L	L	M	H	H-L
Age	0.11	0.89	1.07	0.96	0.20	2.07	2.95	2.32
Wealth	0.12	1.03	1.01	0.89	0.23	2.71	2.40	2.52
Panel B: Monthly CAPM Alphas								
	Alpha				t(Alpha)			
	L	M	H	H-L	L	M	H	H-L
Age	-0.80	0.00	0.26	1.06	-2.28	0.00	2.26	2.58
Wealth	-0.81	0.18	0.17	0.98	-2.67	1.83	0.78	2.80
Panel C: Monthly CAPM Betas								
	Beta				t(Beta)			
	L	M	H	H-L	L	M	H	H-L
Age	1.12	1.09	0.94	-0.18	18.94	37.62	49.63	-2.54
Wealth	1.15	1.01	0.99	-0.17	22.72	59.88	26.88	-2.82

Table IV
Additional Investor Factors

This table reports the intercept, slope, and R-square coefficients of time-series OLS regressions of additional investor factors on the equal-weighted combination of the age and wealth factors. Investor factors are constructed by sorting stocks on a particular characteristic and constructing a portfolio that takes a long (short) position in the top (bottom) 30% of stocks sorted by the characteristic. Investor characteristics include: dummy variables for women, retirees, permanent income brackets, labor income-to-wealth brackets, occupation, and region, and a categorical variable equal to 1, 2, or 3 based on the investor's highest education level.

Factor	α	t(α)	b	t(b)	$R2$
Socioeconomic Characteristics					
Male dummy	0.05	0.15	-0.67	-8.90	0.23
Education level	-0.04	-0.11	0.16	2.26	0.02
Labor-to-wealth	-0.49	-1.49	-0.18	2.69	0.03
Permanent income	-0.51	-1.44	0.19	2.56	0.02
Retiree dummy	0.06	0.22	0.89	15.01	0.46
Occupational Sector					
Resource industries	0.20	0.53	0.09	1.22	0.01
Petroleum	0.06	0.17	-0.11	-1.54	0.01
Consumer manufacturing	0.12	0.33	-0.03	-0.37	0.00
Material manufacturing	-0.25	-0.74	0.08	1.11	0.00
Technological manufacturing	0.02	0.05	-0.33	-4.52	0.08
Public administration	0.15	0.51	0.16	2.82	0.03
Construction	0.24	0.73	-0.48	-7.11	0.17
Trade	-0.06	-0.16	-0.27	-3.55	0.05
Transportation and logistics	-0.26	-0.69	-0.26	-3.50	0.05
Tourism	-0.04	-0.12	-0.06	-0.91	0.00
Media and ICT	0.33	0.98	-0.37	-5.47	0.11
Finance	0.15	0.41	-0.02	-0.27	0.00
Knowledge-based business services	-0.26	-0.76	-0.22	-3.24	0.04
Technological services	-0.18	-0.49	0.04	0.61	0.00
Business support services	0.13	0.31	-0.12	-1.47	0.01
Education	-0.05	-0.13	0.13	1.70	0.01
Health and social services	0.17	0.55	0.16	2.68	0.03
Non-profit and household services	-0.08	-0.24	0.15	2.21	0.02
Real estate activities	0.64	1.60	-0.23	-2.88	0.03

Table V
Spanning Regressions

The table reports OLS spanning regressions of the investor factor consisting of age and wealth (equally-weighted) on different firm factor models in 1997-2018. For each regression, we report the intercept (alpha) and its t-value. All factor models are constructed from the following factors: the market factor (MKT) as the value-weighted portfolio of all the stocks in the universe, the size factor (SMB) based on market capitalization, the value factor (HML) based on industry adjusted book-to-market ratio, the profitability factor (RMW) based on profit margin, the investment factor (CMA) based on total asset growth, and the momentum factor (MOM) based on the geometric return over the previous 12 months where the most recent month is left out. Firm-3 includes MKT, SMB and HML. Firm-4 includes MKT, SMB, HML and MOM. Firm-5 includes MKT, SMB, HML, CMA and RMW. Firm-6 includes MKT, SMB, HML, CMA, RMW and MOM. SR stands for Sharpe Ratio.

		Alpha				
Benchmark factors	SR	Mkt	Firm-3	Firm-4	Firm-5	Firm-6
Test-statistic	0.68	1.05	1.07	0.85	0.76	0.66
$t-$value	9.98	3.63	4.15	3.50	3.10	2.79

Table VI
Out-of-Sample Performance of Factor Models

This table reports out-of-sample Sharpe ratios of mean-variance efficient portfolios of pricing factors. Following Fama and French (2018), we randomly select 100,000 in-sample periods comprised of 264 months, construct tangency portfolios from the factors' in-sample return moments, and evaluate performance in the remaining out-of-sample-periods. Efficient portfolios are constructed by (i) setting weights of 1 on the market and 1/2K for each of the long-short factor portfolios ("fixed-weight"), and (ii) running a mean-variance optimization of in-sample factor returns with shrinkage on their covariance matrix. For each approach, we calculate the Sharpe ratio over the out-of-sample (OS) period and report the average Sharpe ratio across all simulations. The last column scales the average Sharpe ratio of the MV-weight portfolio by its average in-sample (IS) Sharpe ratio. The Firm-4 model includes HML, SMB, RMW, and CMA, and the Firm-5 model includes HML, SMB, RMW, CMA, and MOM.

	Fixed-weight	MV-weight	
	OS Sharpe ratio	OS Sharpe ratio	OS-IS ratio
Mkt, Age	0.58	0.51	0.74
Mkt, Wealth	0.57	0.54	0.75
Mkt, SMB	0.32	0.13	0.48
Mkt, HML	0.34	0.17	0.44
Mkt, MOM	0.55	0.44	0.69
Mkt, CMA	0.15	0.34	0.61
Mkt, RMW	0.56	0.49	0.72
Mkt, Age, Wealth	0.61	0.66	0.73
Mkt, SMB, HML	0.35	0.08	0.24
Mkt, SMB, MOM	0.46	0.39	0.59
Mkt, SMB, CMA	0.24	0.29	0.49
Mkt, SMB, RMW	0.45	0.48	0.63
Mkt, HML, MOM	0.48	0.38	0.55
Mkt, HML, CMA	0.25	0.26	0.42
Mkt, HML, RMW	0.48	0.43	0.59
Mkt, CMA, RMW	0.36	0.52	0.65
Mkt, CMA, MOM	0.37	0.48	0.62
Mkt, RMW, MOM	0.59	0.55	0.68
Firm-4	0.44	0.34	0.48
Firm-5	0.36	0.44	0.50
Firm-6	0.41	0.50	0.52
Firm-6, Age, Wealth	0.48	0.65	0.58

Table VII
Panel Regressions of the Age Factor Tilt on Investor Characteristics

This table reports panel regressions of the age factor tilt on investor characteristics and age dummy variables. The estimation is run on a panel of Norwegian individual investors in 1997-2018. The age factor tilt is calculated annually from the direct stockholdings of investors. Income beta is the slope coefficient from a panel regression of an investor's annual income growth on real GDP growth, where the estimation is conducted within a group of investors in the same employment sector and labor income bracket. The debt-to-income ratio is the ratio of an investor's total debt to labor income. Stock market experience is defined as the number of years of stock market participation. The male dummy, the Master's degree dummy, the business education dummy, and the finance occupation dummy are indicator variables respectively equal to unity if the investor is male, has obtained a Master's degree, has studied business or economics, or works in a finance-related sector. The age dummy variables correspond to 10 groups of investors in five year increments. The median age group (50-55 years) is used as the reference point and the corresponding dummy is removed from the estimation. We include year fixed effects and twelve wealth-bracket fixed effects. Statistical significance is indicated by ***, **, and * for the 0.01, 0.05, and 0.10 levels. Standard errors are clustered at the calendar year and investor levels.

	Dependent Variable: Age Factor Tilt				
	(1)	(2)	(3)	(4)	(5)
Risk Exposures:					
Income beta	-0.093***	-0.089***	-0.064***	-0.122***	-0.088***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Debt-to-income ratio			-0.010***		-0.010***
			(0.003)		(0.002)
Experience, Education, and Gender:					
Stock market experience	0.014***	0.014***	0.015***	0.014***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Male dummy			-0.135***		-0.134***
			(0.007)		(0.006)
Master's degree dummy				0.021**	0.027***
				(0.009)	(0.009)
Business education dummy				0.024***	0.028***
				(0.007)	(0.007)
Finance occupation dummy				0.116**	0.088
				(0.055)	(0.054)

(Continued)

Table VII - *Continued*

	Dependent Variable: Age Factor Tilt				
	(1)	(2)	(3)	(4)	(5)
Age Group Dummies:					
< 30	-0.041***	-0.033**	-0.021*	-0.032**	-0.021*
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
30-34	-0.072***	-0.062***	-0.054***	-0.065***	-0.059***
	(0.012)	(0.011)	(0.011)	(0.012)	(0.012)
35-39	-0.073***	-0.065***	-0.059***	-0.067***	-0.062***
	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
40-44	-0.056***	-0.052***	-0.048***	-0.052***	-0.050***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
45-49	-0.028***	-0.026***	-0.024***	-0.026***	-0.024***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
55-59	0.026***	0.024***	0.023***	0.025***	0.024***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
60-64	0.023***	0.021***	0.025***	0.018***	0.022***
	(0.007)	(0.007)	(0.007)	(0.004)	(0.005)
65-69	0.011	0.011	0.023*	-0.004	0.013
	(0.012)	(0.012)	(0.013)	(0.009)	(0.009)
≥ 70	0.049***	0.050***	0.065***	0.022*	0.042***
	(0.014)	(0.014)	(0.015)	(0.012)	(0.012)
Year FE:	Yes	Yes	Yes	Yes	Yes
Wealth Bracket FE:	No	Yes	Yes	Yes	Yes
Number of observations	943,457	943,457	943,457	880,319	880,319
Adjusted R^2	0.067	0.068	0.078	0.068	0.078

Table VIII
Panel Regressions of the Wealth Factor Tilt on Investor Characteristics

This table reports panel regressions of the wealth factor tilt on investor characteristics and wealth dummy variables. The estimation is run on a panel of Norwegian individual investors in 1997-2018. The wealth factor tilt is calculated annually from the direct stockholdings of investors. Income beta is the slope coefficient from a panel regression of an investor's annual income growth on real GDP growth, where the estimation is conducted within a group of investors in the same employment sector and labor income bracket. The debt-to-income ratio is the ratio of an investor's total debt to labor income. Stock market experience is defined as the number of years of stock market participation. The male dummy, the Master's degree dummy, the business education dummy, and the finance occupation dummy are indicator variables respectively equal to unity if the investor is male, has obtained a Master's degree, has studied business or economics, or works in a finance-related sector. The wealth dummy variables correspond to the first 9 deciles, the 90^{th}-99^{th} percentiles, the 99^{th}-99.9^{th} percentiles, and the top 0.1% of the wealth distribution. The median wealth group (50^{th}-60^{th} percentiles) is used as the reference point and the corresponding dummy is removed from the estimation. We include year fixed effects and ten age-group fixed effects. Statistical significance is indicated by ***, **, and * for the 0.01, 0.05, and 0.10 levels. Standard errors are clustered at the calendar year and investor levels.

	Dependent Variable: Wealth Factor Tilt				
	(1)	(2)	(3)	(4)	(5)
Risk Exposures:					
Income beta	-0.047***	-0.044***	-0.036***	-0.070***	-0.059***
	(0.008)	(0.009)	(0.008)	(0.010)	(0.011)
Debt-to-income ratio		0.001		0.001	
		(0.002)		(0.002)	
Experience, Education, and Gender:					
Stock market experience	0.007***	0.007***	0.007***	0.006***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Male dummy			-0.046***		-0.043***
			(0.012)		(0.014)
Master's degree dummy				-0.002	0.0003
				(0.010)	(0.011)
Finance education dummy				0.032***	0.033***
				(0.007)	(0.008)
Finance occupation dummy				0.077**	0.067*
				(0.031)	(0.033)

(Continued)

Table VIII - *Continued*

	Dependent Variable: Wealth Factor Tilt				
	(1)	(2)	(3)	(4)	(5)
Wealth Percentile Dummies					
Bottom 10%	-0.048***	-0.046***	-0.042***	-0.048***	-0.043***
	(0.006)	(0.005)	(0.010)	(0.005)	(0.010)
10-20	-0.026***	-0.025***	-0.020***	-0.026***	-0.022***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
20-30	-0.018***	-0.018***	-0.014***	-0.019***	-0.016***
	(0.005)	(0.004)	(0.003)	(0.004)	(0.003)
30-40	-0.010***	-0.011***	-0.008***	-0.010***	-0.008**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
40-50	-0.004	-0.004	-0.004	-0.003	-0.002
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
60-70	-0.003	-0.003	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
70-80	-0.004	-0.004	-0.002	-0.003	-0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
80-90	0.004	0.003	0.007**	0.004	0.007**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
90-99	0.030***	0.029***	0.035***	0.028***	0.034***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)
99-99.9	0.084***	0.083***	0.092***	0.078***	0.086***
	(0.011)	(0.011)	(0.009)	(0.013)	(0.010)
Top 1%	0.167***	0.166***	0.175***	0.163***	0.170***
Year FE:	Yes	Yes	Yes	Yes	Yes
Age Group FE:	No	Yes	Yes	Yes	Yes
Number of observations	943,457	943,457	943,457	880,319	880,319
Adjusted R^2	0.047	0.047	0.049	0.048	0.049

Table IX
Firm Characteristics of Investor Factors

This table report the median firm characteristic in the low-, medium-, and high- portfolios that are used to construct the age and wealth factors from the universe of Norwegian stocks in 1997-2018. These portfolios correspond to the bottom 30%, mid 40%, and top 30% of stocks sorted on either the age or wealth investor characteristics each year. For each portfolio, the median firm characteristic is estimated in the panel of firms. Years in sample refer to the number of years the stock is in our panel. The share of institutional ownership is measured in percentage points. Volatility is based on daily returns and is equal to the square root of the realized variance measured over the previous 12 months. Turnover is defined the average daily trading volume multiplied by 30 and divided by the free-float-adjusted market valuation. CAPM beta is estimated from a 60 months rolling-window estimation of the stock excess return on the market factor. Size is the market value of equity reported in million NOK. BE/ME is book value of equity scaled by size. Profitability is the ratio of gross profit (the difference between total revenue and cost of goods sold) to total assets. The investment growth variable refers to the growth rate in total assets.

	Age-Sorted Portfolios				Wealth-Sorted Portfolios			
	L	M	H	H-L	L	M	H	H-L
Years in sample	8.00	10.00	13.00	5.00	7.00	9.00	16.00	9.00
Institutional ownership share (%)	3.1	6.4	6.4	3.36	5.1	6.4	4.4	-0.67
Turnover	7.23	1.65	0.56	-6.67	5.26	2.17	0.38	-4.88
Volatility	0.25	0.13	0.09	-0.16	0.24	0.14	0.08	-0.16
CAPM Beta	0.88	0.83	0.67	-0.22	0.94	0.84	0.66	-0.28
Size (NOK million)	384	1342	1485	1102	508	1103	2118	1610
BE/ME	0.72	0.70	0.68	-0.05	0.55	0.65	0.89	0.34
Profitability (%)	0.03	0.07	0.09	0.06	0.05	0.07	0.08	0.03
Investment (%)	0.04	0.07	0.09	0.05	0.09	0.07	0.07	-0.02

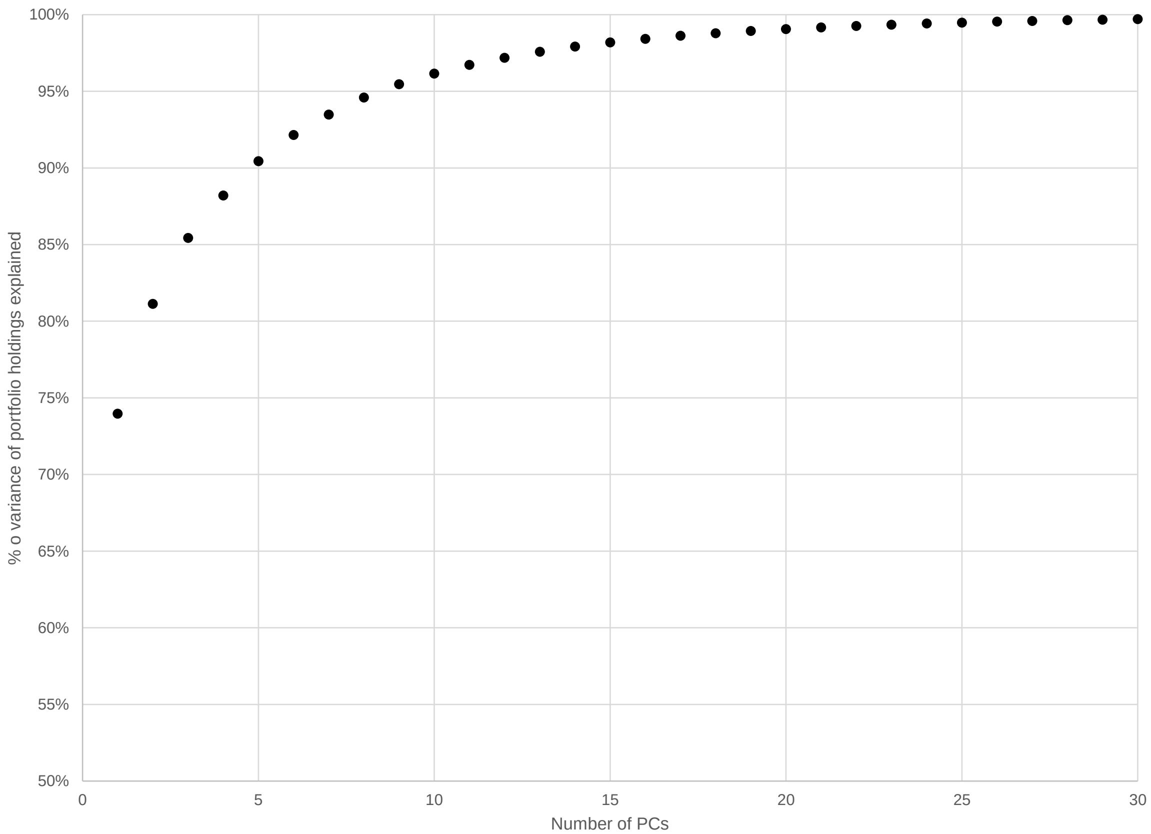


Figure 1 Factor Structure of Investor Portfolio. This figure plots the cumulative proportion of the cross-sectional variance of aggregate portfolio holdings explained by the principal components (PCs). The cumulative proportion is calculated each year in 2002-2017 and then averaged across years. The principal component analysis is based on 93 portfolios consisting of the aggregate holdings of individual investors grouped by age, wealth, and other socioeconomic characteristics.

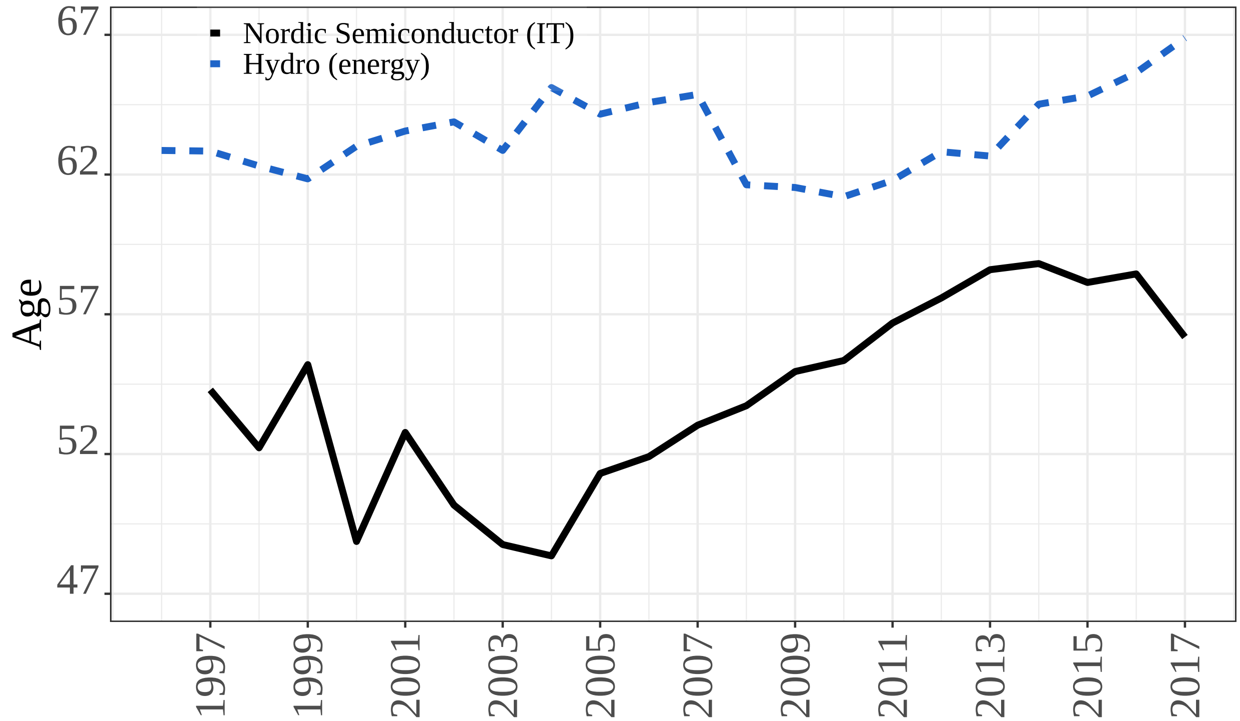
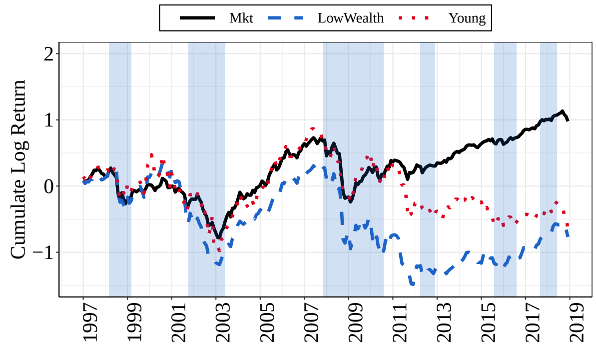
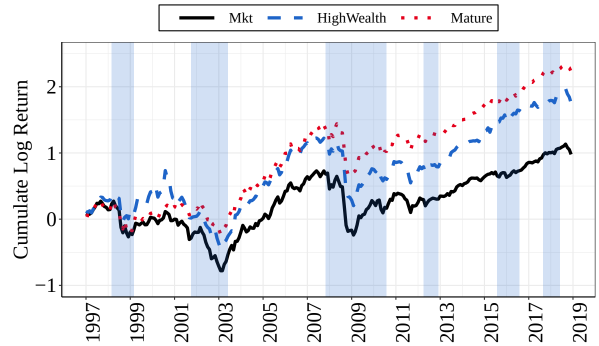


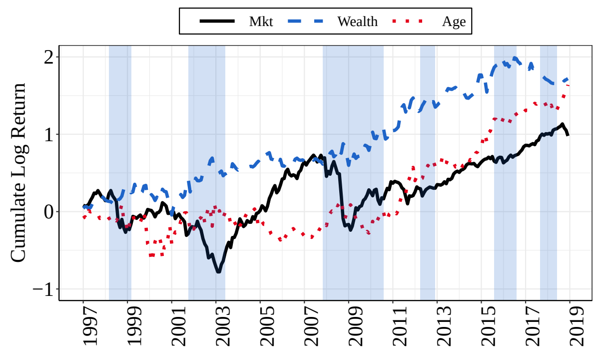
Figure 2 Investor age characteristic for two stocks. This figure plots the age characteristic of Norsk Hydro and Nordic Semiconductor in 1997-2018. Hydro is a fully integrated aluminium company. Nordic Semiconductor is a semiconductor company specializing in wireless technology. For each stock, the age characteristic is calculated as the average age of individual investors who directly own the stock, weighted by the relative number of shares that each investor directly holds at the beginning of the year.



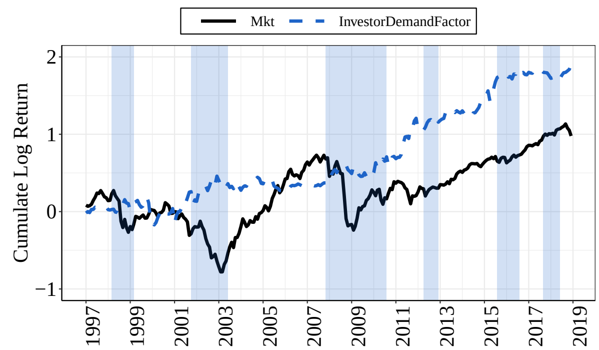
(a) Low wealth and young portfolios



(b) High wealth and mature portfolios

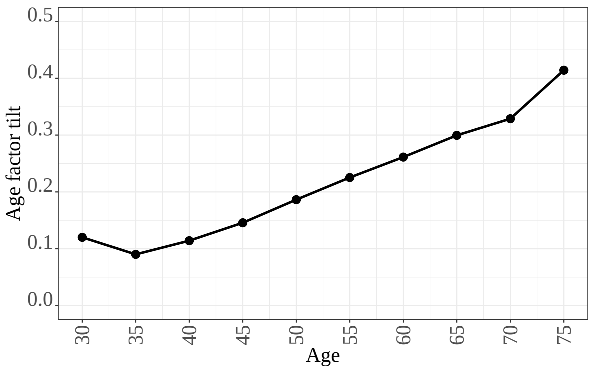


(c) Age and wealth factors

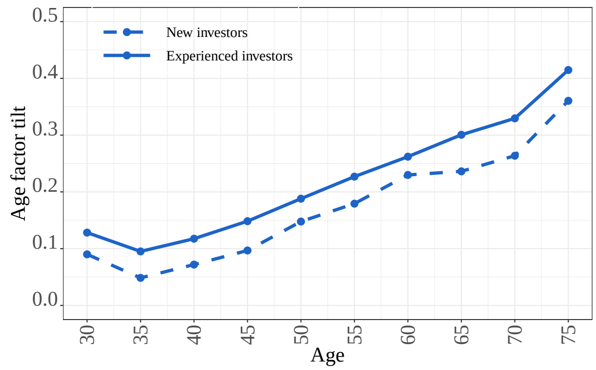


(d) Age and wealth factors combined

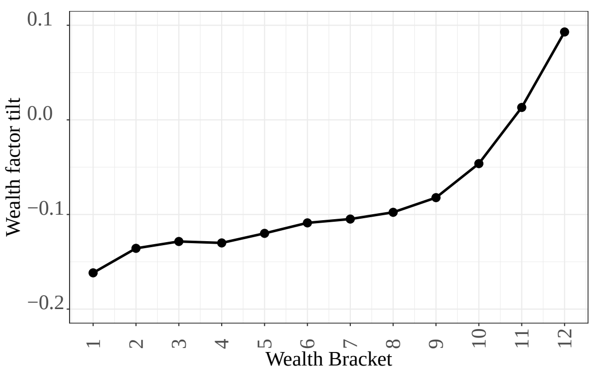
Figure 3 Cumulative return on investor factors. This figure plots the log cumulative return on portfolios of Norwegian stocks sorted by investor age and wealth characteristics in 1997-2018. Panel A plots historical returns on the young portfolio and on the low-wealth portfolio. Panel B plots historical returns on the mature portfolio and on the high-wealth portfolio. Panel C plots the age factor (mature-minus-young) and wealth factor (high wealth-minus-low wealth) portfolios. Panel D plots a factor obtained by the equal-weighted combination of the age and wealth factors. In each panel, the black line represents the performance of the market portfolio. The blue bars indicate economic recessions. The portfolios are constructed as follows. We first sort stocks by the age characteristic of the individual investors who directly own the stocks. We then define the young portfolio as the value-weighted portfolio of stocks in the bottom 30%, and the mature portfolio as the value-weighted portfolio of stocks in the top 30%. We similarly define the high-wealth and low-wealth portfolios by sorting stocks according to the net worth of their investors.



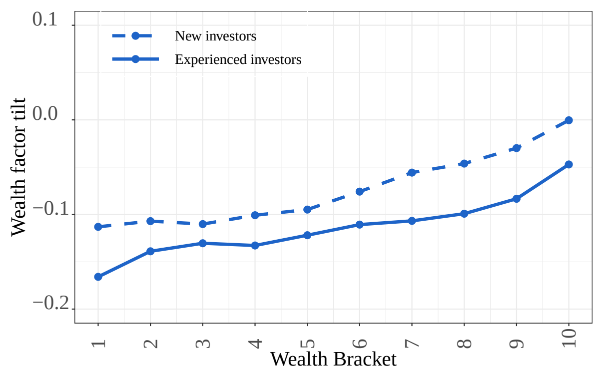
(a) Age tilt: All participating investors



(b) Age tilt: New and existing participants



(c) Wealth tilt: All participating investors



(d) Wealth tilt: New and existing participants

Figure 4 Factor tilts across investor groups. This figure plots the average tilts toward the age and wealth factors of investors in different age and wealth groups. The analysis is based on the panel of Norwegian individual investors who hold stocks directly during the 1997-2018 period. Panel A plots the average age tilt across 10 age groups. Panel B plots the average age tilt of new participants (dotted) and preexisting participants (solid) each year. Panel C plots the average wealth tilt of individual investors across 12 different wealth groups. Panel D plots the average wealth tilt of new participants (dotted) and preexisting participants (solid) each year. Averages are equally-weighted. New participants are investors with less than one year of experience with direct stock investing, while preexisting participants have at least one year of experience.