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*

March 1, 2021

Abstract

We construct a parsimonious set of equity factors by sorting stocks according to the sociodemographic characteristics of the individual investors who own them. The analysis uses administrative data on the stockholdings of Norwegian investors in 1997-2018. Consistent with financial theory, a mature-minus-young factor, a high wealth-minus-low wealth factor, and the market factor price stock returns. Our three factors span size, value, investment, profitability, and momentum, and perform well in out-of-sample bootstrap tests. The tilts of investor portfolios toward the new factors are driven by wealth, indebtedness, macroeconomic exposure, age, gender, education, and investment experience. Our results are consistent with hedging and sentiment jointly driving portfolio decisions and 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 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, Joost Driessen, Paul Ehling, Evan Jo, Patrick Konermann, and Richard Priestley, and seminar participants at BI Business School, McGill University, and Tilburg University. We gratefully acknowledge financial support from Finansmarkedsfondet at Research Council of Norway under Project Number No. 309855.

I. Introduction

A key objective of financial economics is to relate equity risk premia to the drivers of investor demand (Campbell, 2018). This agenda has proven challenging to achieve because several of the most empirically successful pricing factors are constructed from firm characteristics that are not directly related to investor preferences, risks, and biases (Fama and French, 2015; Hou, Xue, and Zhang, 2015). This disconnect has prompted a large literature to search for economic mechanisms tying empirical pricing factors to consumption (see Constantinides (2017); Ludvigson (2013); Mehra (2012) and the references therein). However, another challenge for investor-based asset pricing is that consumption data are noisy and measured at relatively low frequency.

For these reasons, a new literature has started exploiting the rich information contained in investor portfolio holdings, which can be observed without error in real time. In an influential study, Koijen and Yogo (2019) use the holdings of large U.S. institutional investors to draw causal inference about the investors' influence on asset prices.[1] Other studies use institutional holdings data 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).

While the recent literature focuses on the holdings of institutions, the pricing information contained in the portfolios of individual investors has hitherto remained largely unexplored. This question seems theoretically important because a large body of financial theory predicts the interactions between individual portfolio decisions and asset returns. From an empirical standpoint, even though retail investors own directly only a limited fraction of aggregate equity (Blume and Keim, 2012), they represent a large and diverse group that responds elastically to stock prices and may therefore drive stock returns (Barber et al., 2009; Kaniel et al., 2008; Kelley and Tetlock, 2013).

Another key benefit of investigating individual investor portfolios is that they are informative about the preferences and trading behaviors that micro-found investor-based asset pricing models. Specifically, the relationships between the stockholdings of individual investors and socioeconomic characteristics can be measured in panel datasets. An extensive household finance literature has successfully applied this empirical strategy to components of household balance sheets measured at various levels of disaggregation (Campbell, 2006;

[1] See also Koijen et al. (2020b) and Koijen and Yogo (2020) for other applications for the demand system approach in equity markets.

Gomes et al., 2020; Guiso and Sodini, 2013). A similar approach can be applied to uncover the drivers of demand for individual stocks or groups of stocks. For this purpose, direct stockholdings are likely more informative than indirect holdings managed by institutions facing frictions in delegated asset management (He and Xiong, 2013).[2] Betermier, Calvet, and Sodini (2017) have accordingly uncovered robust links between household characteristics and their portfolio exposures to the value factor, which appear most strongly in direct stockholdings. The next step is to go beyond traditional firm-based factors and to define the equity factors themselves from investor portfolio data.

Our paper makes progress in this direction by investigating what the portfolios of individual investors reveal about the factors that price the cross-section of equity returns. Our analysis aims to answer three 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? If so, how do the new 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 new factors?

A major impediment to answering these questions 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 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.[3] Our paper resolves this challenge by using detailed portfolio data on all Norwegian direct stockholders between 1997 and 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 the Norwegian panel, 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 pricing the cross-section of stock returns, both in and out of sample. The tight connection between investor factors and investor portfolio decisions allows

[2]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).

[3]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%.

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-section of investor portfolios and the cross-section of stock returns. Our starting point is the recent empirical finding by Balasubramaniam et al. (2020) that the portfolios of individual investors exhibit a factor structure. 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 younger and less wealthy investors.

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

We follow the standard approach of forming long-short portfolios based on sorting stocks along a stock-level characteristic. The innovation of the paper is to use 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, over the 1997-2018 period, which correspond to yearly alphas of about 12%. The statistical significance of the age and wealth factors is comparable to the significance over the same period of the investment, profitability, and momentum factors, which are some of the best firm-based factors available. 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 when the CAPM is used as benchmark. Furthermore, the age and wealth factors have negative CAPM betas, which is also consistent with theory.

The three-factor model defined by the age, wealth, and market factors provides a powerful specification of the cross section of equity premia, as theory predicts. We show the strength of our three-factor model by implementing spanning tests as in Barillas and Shanken (2016). We verify that the size, value, investment, profitability, and momentum factors are spanned by our three-factor model over the 1997-2018 period. As Barillas and Shanken (2016) explain, this result implies that our investor-based model matches the pricing ability of some of the best-performing traditional factors available. We draw similar conclusions from Gibbons, Ross, and Shanken (1989) tests.

Our three-factor investor-based model is also a strong performer out of sample. We demonstrate this property by implementing Fama and French (2018) out-of-sample bootstrap tests. That is, we split the 264 months of our sample period into 132 adjacent pairs and randomly assign one month of each pair to the in-sample period and the other month to the out-of-sample period. We use the in-sample period to estimate the moments of factors and then compute the Sharpe ratio of the resulting tangency portfolio out of sample across 100,000 random classifications of the months. Our investor-based model produces an average out-of-sample annualized Sharpe ratio of 0.70. As a comparison, the Sharpe ratio of the market is 0.32 in Norway over the sample period. Moreover, the 0.70 out-of-sample Sharpe ratio of our three-factor model matches the values generated by combining by the six-factor model specified by the market, size, value, investment, profitability, and momentum. When we combine investor and firm factors, the out-of-sample Sharpe ratio increases to 0.84. Investor factors therefore expand the mean-variance frontier compared to that obtained using firm factors only. Overall, investor factors perform strongly in equity pricing tests.

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 driven primarily by portfolio inertia. 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 short 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 operate in tandem and generate pricing factors that are observationally equivalent to each other.

To gain additional insights into the nature of investor factor tilts, we analyze 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 that have a 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 is limited.

Our paper contributes to the extensive literature on the cross-section of equity premia (Cochrane, 2011; Harvey, Liu, and Zhu, 2015) by extracting pricing information from the stock holdings of individual investors. Our results complement the growing research on the interaction of investor holdings and prices. The seminal contributions of Koijen and Yogo (2019) and Koijen, Richmond, and Yogo (2020b) identify the types of institutional investors that have the strongest price impact. 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 rich information about the cross-section of equity premia.

A key benefit of constructing equity factors from the portfolios of individual investors is that they tie directly equity pricing to investor risks, preferences, and biases. By contrast, firm factors are more informative about firm production decisions that micro-found production-based asset pricing models. Thus, both types of factors provide useful complementary information about the sources of equity premia and can therefore be fruitfully used in tandem. 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). In practice, data limitations may limit the statistical ability of a particular class of factors, so that combining investor and firm factors can produce the most accurate results, as some of our empirical results illustrate.

Our paper 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 show the factor structure of holdings has important implications for equity pricing.[4]

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 single pricing factor, which is informative about the sources of differences in performance across investors.

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

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 the construction of the age and wealth factors. Section IV assesses the ability of investor factors 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

In this Section, we present a simple 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.

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 factors. 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.[5]

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 $\boldsymbol{f} = (f_0, \ldots, f_K)'$ price the cross-section of stock returns.

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

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

[5]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).

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.

A priced factor $\boldsymbol{d_k}$ generates CAPM-alpha. 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

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

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 $\boldsymbol{m}$ 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.

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

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 $R^e_{1,t+1}, ..., R^e_{J,t+1}$. The conditional distribution of the return vector $(R_f, R^e_{1,t+1}, ..., R^e_{J,t+1})$ 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 (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

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

function of age, wealth, and labor income:

$$\boldsymbol{\omega_t^i} = \boldsymbol{\tau_t} + \boldsymbol{d}(A_t^i, W_t^i, L_t^i, \boldsymbol{y_t}), \tag{11}$$

where $A_t^i = t - b_i$ denote the investor's age at date t. In the Appendix, we derive the relation between $\boldsymbol{\omega_t^i}$ and the value function. We verify that the deviation portfolio is zero for an investor in the last investment period ($A_t^i = T - 1$) 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: $\boldsymbol{d}(A_t^i, L_t^i/W_t^i, \boldsymbol{y_t})$. We apply a Taylor expansion to the deviation portfolio around the last investment period ($A_t^i = T - 1$) and a labor income-to-wealth ratio equal to 0, and obtain the portfolio factor structure:

$$\boldsymbol{d_t^i} = (T - 1 - A_t^i)\, \boldsymbol{d_{1,t}} + \frac{L_t^i}{W_t^i} \boldsymbol{d_{2,t}},$$

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 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

[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 $\boldsymbol{d}_{2,t} = \partial \boldsymbol{d^*}/\partial v(0, 0, \boldsymbol{\xi_t})$.

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 Construction of Investor Factors

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 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

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

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 442 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 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. *Investor Factors*

Guided by the theoretical discussion in Section II.B, we construct equity factors based on investor age and wealth. We proceed in two steps. First, we follow the standard approach of sorting stocks by a stock-specific characteristic. In our case, a stock's characteristic will be either the weighted average age or the weighted average net worth in its individual investor base. We then construct equity factors as long-short portfolios of the sorted stocks.[13]

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:

$$\text{Age}_{j,t} = \frac{\sum_{i=1}^{I} N_{j,t}^{i} \, A_t^i}{\sum_{i=1}^{I} N_{j,t}^{i}}, \tag{13}$$

[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.

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

[13]In the Appendix, we also construct the age and wealth factors directly from investor portfolios as motivated by (7). The results are similar. The empirical method outlined in this Section has the advantages of being standard and informative about the implied age and wealth characteristics for each firm.

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 1 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 1

Wealth. Throughout the paper, wealth refers to net worth, defined as the sum of the investor's liquid financial wealth, real estate, vehicles, and business assets, net of liabilities.[14] 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.

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. We form 12 groups of investors based on their net worth percentile each year and assign to each investor the corresponding wealth bracket. The groups include the first 9 deciles of the net worth distribution (groups 1-9), the 90$^{\text{th}}$-99$^{\text{th}}$ percentiles (group 10), the 99$^{\text{th}}$-99.9$^{\text{th}}$ percentiles (group 11), and the 99.9$^{\text{th}}$-100$^{\text{th}}$ percentiles (group 12).[15] We denote by $\text{WB}_t^i \in \{1, \dots, 12\}$ the wealth bracket of investor i at date t.

[14]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.

[15]The net worth distribution is based on the entire Norwegian population that are between 18 and 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 has been advocated by for example Fagereng et al. (2017) in their analysis of portfolio choice in Norway.

We define the stock's wealth characteristic,

$$\text{Wealth}_{j,t} = \frac{\sum_{i=1} N^i_{j,t}\, \text{WB}^i_t}{\sum_{i=1} N^i_{j,t}}, \tag{14}$$

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

Table I 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^i_t 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, $\text{Age}_{j,t}$, has a cross-sectional standard deviation of 5 years. The firm's wealth characteristic, $\text{Wealth}_{j,t}$, 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 very heterogeneous across stocks.

Insert Table I

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.

[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 $\text{Wealth}_{j,t}$ 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.

C. *Firm Factors*

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. In the Appendix, we show that these factors are able to price reasonably well the cross-section of stock returns in Norway.

IV. Pricing Performance of Investor Factors

We now assess the pricing performance of the pricing model consisting of the age, wealth, and market factors. Section IV.A investigates the mean return, CAPM alpha, and CAPM beta of the age and wealth factors. Section IV.B evaluates the ability of our three-factor model to price traditional factors constructed from firm characteristics. In Section IV.C, we follow Fama and French (2018) and use bootstrap simulations to compare the out-of-sample performance of investor-based and firm-based factor models.

[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.

A. *Summary Statistics and CAPM Tests*

Table II, 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 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 II

In Panel B of Table II, we show that the average return on investor factors is not explained by their exposures to market portfolio risk. We report CAPM regressions of the age and wealth factors on the market over the sample period. The monthly intercepts 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 II 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.A and the recent equilibrium analysis of Betermier, Calvet, and Jo (2020). Young and less wealthy investors tilt their portfolios toward stocks that provide hedging benefits or are attractive due to irrational exuberance or other forms of sentiment. In equilibrium, these attractive stocks generate negative alphas and have low discount rates, which pushes up their valuations. As a result they represent a large share of the market portfolio and have 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. We refer the reader to Betermier, Calvet, and Jo (2020) for further empirical evidence on this mechanism.

[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 $\boldsymbol{d_k}$.

In Figure 2, 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 2 shows that the long legs of the age and wealth factors outperformed the market throughout the sample.

Panel C of Figure 2 illustrates the cumulative performance of the age and wealth factors. Both factors have high average returns and low volatilities. 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 2 further illustrates the benefits of using both factors by reporting the performance of an equal-weighted portfolio of the age and wealth factors. This portfolio yields significantly higher performance than the market portfolio, while also displaying lower volatility than each factor taken separately.

B. Spanning Regressions

We next evaluate the pricing performance of our three-factor model defined by age, wealth, and the market. The benchmark is again the model containing the six traditional factors discussed in Section III.C. Our testing procedure builds on the work of Barillas and Shanken (2016), who show that a candidate model's ability to price the full cross-section of stock returns better than a benchmark model is fully driven by the candidate model's ability to price the benchmark factors.

Table III, Panel A reports the CAPM alpha of each of the five firm factors over the sample period. Three factors have statistically significant monthly CAPM-alphas: the investment factor CMA_t (alpha of 0.66, t-value of 2.01), the profitability factor RMW_t (alpha of 0.96, t-value of 2.7), and the momentum factor MOM_t (alpha of 1.05, t-value of 2.32). These alphas are comparable in size to those of the age and wealth factors. Size and value have statistically insignificant alphas. From now on, we focus on the three firm factors, investment, profitability, and momentum, with significant CAPM alphas.

In Table III, Panel B, we regress each of these three firm factors on the three factors of our model. Remarkably, the alphas of the firm factors drop by about 40% and become statistically insignificant. This reduction in pricing errors reflects the wealth factor's ability to explain the momentum and investment factors, and the age factor's ability to explain the profitability factor over the 1997-2018 period.

In Table III, Panel C, we test the null hypothesis that the intercepts of all five firm factors are jointly equal to zero. We follow the testing procedure of Gibbons, Ross, and Shanken (1989) and regress multiple combinations of the firm factors on the three factors of our model. We consider five combinations of firm factors: i) the size and value factors from Fama and French (1993), ii) the size, value, and momentum factors from Carhart (1997), iii) the size, value, profitability, and investment factors from Fama and French (2015), iv) momentum and the Fama and French (2015) factors, and v) the three factors (profitability, investment, and momentum) delivering significant CAPM alphas in Panel A. For every combination, we fail to reject the null hypothesis that pricing errors are zero at the 5% significance level.

These results indicate that our three-factor model, derived from the direct portfolio holdings of individual investors, spans the size, value, investment, profitability, and momentum factors over the 1997-2018 period. In the Appendix, we show that these results are robust to several alternative specifications of the age and wealth factors. The investor model's ability to span firm factors is important because these factors have been been shown to successfully price the cross-section of stocks over time and around the world (Fama and French, 2012; Griffin et al., 2003). Our analysis thus suggests that a model based on investor-based factors provides a complementary and parsimonious approach to equity pricing. This finding provides encouraging news to theoretical asset pricing models using investor characteristics as inputs.

C. Out-of-Sample Sharpe Ratios

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 tangency portfolios based on the factors in sample and then by estimating the Sharpe ratios of these portfolios out of sample. One important pitfall is that a factor with an uncharacteristically high mean in sample tends to have a large weight in the estimated tangency portfolio, which will tend to reduce portfolio performance out of sample. We control for in-sample biases arising from optimizing over short sample periods by following the bootstrap evaluation approach of Fama and French (2018).

We begin by splitting the 264 months of the full sample period into 132 adjacent pairs:

(1,2), (3,4), . . . , (263,264). We randomly assign one month from each pair to the in-sample period and the other month to the out-of-sample period. For each factor model, we use in-sample returns to estimate the factor weights that maximize the portfolio's Sharpe ratio. We then compute the portfolio's Sharpe ratio over the out-of-sample period. We repeat this simulation 100,000 times (with replacement) and calculate the average in-sample and out-of-sample Sharpe ratios for each model. Whereas in-sample Sharpe ratios are 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.

Figure 3 reports the average out-of-sample annualized Sharpe ratio for each factor model. The horizontal axis classifies models 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).

Investor factors provide the highest Sharpe ratio. Combining the wealth factor with the market generates a Sharpe ratio of 0.64. No combination of a firm factor and the market performs better. The second best factor is profitability, which generates a Sharpe ratio of 0.62 when combined with the market. The third best factor is age, which has a Sharpe ratio of 0.57 when combined with the market.

Among three-factor models, the combination of wealth, age, and the market again performs best, with a Sharpe ratio of 0.70 in annual units. Moreover, our model performs nearly as well as the pricing model that contains the market and all five firm factors (Firm-5), which generates an average Sharpe ratio of 0.73.

A final insight from Figure 3 is that investor factors expand the mean-variance frontier compared to that obtained using firm factors only. Indeed, combining the three-factor investor model with all five firm factors allows us to obtain a significant increase in the Sharpe ratio, which now reaches 0.84. Altogether, the evidence suggests that our three-factor pricing model performs strongly both in and out of sample.

V. The Cross-Section of Investor Portfolio 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 issue 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 now 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 random groups. The first group contains two-thirds of the investor population and is used to reestimate the age and wealth factors.[19] 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:

$$\omega^i_{H,t} = \sum_{j=1}^{J} \omega^i_{j,t} \mathbb{1}_{j,H,t}, \tag{15}$$

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

$$\omega^i_{f,t} = \omega^i_{H,t} - \omega^i_{L,t}. \tag{16}$$

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.

[19] 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.

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 before age 30 and progressively increases to 0.4 for the oldest group. The panel shows a large and remarkably linear migration in the age factor tilt over the life-cycle.

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 is more likely to originate 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). The portfolio tilts chosen by new entrants closely mimic the tilts of pre-existing investors of the same age. This result confirms that the age ladder is unlikely to be due to portfolio inertia and that 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 III.B. Investors progressively migrate toward the wealth factor as they climb the wealth ladder. 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 investors without direct stock market experience 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 direct stock market participation.

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 $g \in \{1, \ldots, G\}$, 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:

$$\Delta y_{i,t} = a_g + \beta_g^{GDP} \Delta GDP_t + \varepsilon_{i,t}. \tag{17}$$

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 during which 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:

$$\omega_{f,t}^{i} = a + \boldsymbol{\gamma' X_t^i} + \eta_t + \epsilon_t^i, \tag{18}$$

where $\boldsymbol{X_t^i}$ is a vector of characteristics, η_t is a time fixed effect, and ϵ_t^i is the residual error term. The vector $\boldsymbol{X_t^i}$ includes the investor's debt-to-income ratio, non-financial income exposure to

GDP 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 10 age groups defined in Section V.A and the 12 wealth brackets defined in Section III.B, and we use the median brackets as benchmarks. Standard errors are clustered by year and investor.

In Table IV, 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 IV

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, 10 years of additional experience explain a 0.18 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 10 years of stock market experience. These results are consistent with sentiment driving portfolio tilts, since financial market experience can be viewed as a proxy for learning effects and gender as a proxy for overconfidence.

Table V 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 V

Risk-based and sentiment-based characteristics 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 0.09 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. 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.

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 financial markets 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).[20]

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

[20]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.

(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 VI 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.

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. Viewed through the lens of our theoretical framework, which predicts the relations between aggregate portfolio tilts and pricing factors, these results help to understand why a factor model constructed from the direct portfolio holdings of individual investors can perform so well in pricing the cross-section of stocks.

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 retail 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 price the cross-section of stock returns. Our three factors span the size, value, investment, profitability, and momentum factors and perform strongly in out-of-sample tests. 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 individual 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.

Barillas, Francisco, and Jay Shanken, 2016, Which alpha?, *Review of Financial Studies* 30, 1316–1338.

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.

Campbell, John Y., 2018, *Financial Decisions and Markets: A Course in Asset Pricing* (Princeton University Press).

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.

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, 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.

Gibbons, Michael R, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–52.

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.

Koijen, Ralph S. J, and Motohiro Yogo, 2020, Exchange rates and asset prices in a global demand system, Working paper.

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.

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

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.

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
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 10th, 30th, 50th, 70th, 90th, and 99th 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 10th category the 90-99th percentiles, the 11th category the 99-99.9th bracket, and the 12th 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 III.B of the main text.

			Percentiles					
Characteristic	SD	Mean	10th	30th	50th	70th	90th	99th
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 II
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 III
Spanning Regressions

The table reports OLS spanning regressions of firm factors on the market factor (Panel A) and the market, age, and wealth factors (Panel B) in 1997-2018. We report the intercept (alpha) and slope coefficients of the regressions along with their respective t-values and the regression R^2. The set of factors in Panel B is limited to the factors that have statistically significant alphas in Panel A. Panel C reports p-values for the Gibbons et al. (1989) test that all intercepts are equal to zero.

Panel A: CAPM Regressions of Firm Factors

	Alpha	t(Alpha)	Mkt	t(Mkt)	R^2
Factor:					
SMB	-0.02	-0.09	-0.20	-5.0	0.08
HML	0.16	0.51	-0.12	-2.2	0.02
MOM	1.05	2.32	-0.21	-2.8	0.02
RMW	0.96	2.70	-0.20	-3.4	0.04
CMA	0.66	2.01	-0.22	-4.0	0.05

Panel B: Regressions of Firm Factors on the Age, Wealth, and Market Factors

	Alpha	t(Alpha)	Mkt	t(Mkt)	Age	t(Age)	Wealth	t(Wealth)	R^2
Factor:									
MOM	0.45	1.05	-0.11	-1.5	0.164	2.6	0.439	5.8	0.17
RMW	0.39	1.28	-0.11	-2.1	0.448	9.8	0.095	1.8	0.31
CMA	0.54	1.63	-0.2	-3.5	0.102	2.1	0.008	0.1	0.06

Panel C: GRS Tests of Firm Factor Regressions on the Age, Wealth, and Market Factors

	GRS Statistic	p-value
Set of Factors:		
SMB, HML	1.93	0.14
SMB, HML, MOM	1.67	0.17
SMB, HML, RMW, CMA	2.03	0.09
SMB, HML, RMW, CMA, MOM	1.69	0.14
RMW, CMA, MOM	1.49	0.21

Table IV
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	-0.094***	-0.090***	-0.064***	-0.126***	-0.091***
	(0.011)	(0.011)	(0.011)	(0.014)	(0.014)
Debt-to-income ratio			-0.010***		-0.010***
			(0.003)		(0.002)
Experience, Education, and Gender:					
Stock market experience	0.017***	0.017***	0.018***	0.017***	0.018***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Male dummy			-0.139***		-0.137***
			(0.010)		(0.009)
Master's degree dummy				0.018*	0.024**
				(0.009)	(0.009)
Business education dummy				0.028***	0.031***
				(0.008)	(0.008)
Finance occupation dummy				0.125*	0.098
				(0.062)	(0.061)

(Continued)

Table IV - *Continued*

	Dependent Variable: Age Factor Tilt				
	(1)	(2)	(3)	(4)	(5)
Age Group Dummies:					
< 30	-0.046***	-0.039**	-0.028*	-0.037**	-0.027
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
30-34	-0.085***	-0.076***	-0.068***	-0.077***	-0.070***
	(0.013)	(0.011)	(0.012)	(0.012)	(0.013)
35-39	-0.073***	-0.066***	-0.060***	-0.067***	-0.063***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.009)
40-44	-0.052***	-0.048***	-0.045***	-0.048***	-0.046***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
45-49	-0.030***	-0.028***	-0.026***	-0.028***	-0.027***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
55-59	0.025***	0.024***	0.023***	0.024***	0.024***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
60-64	0.016**	0.015*	0.018**	0.011**	0.015***
	(0.007)	(0.007)	(0.008)	(0.004)	(0.004)
65-69	0.001	0.001	0.014	-0.018**	0.001
	(0.012)	(0.012)	(0.013)	(0.007)	(0.007)
≥ 70	0.034**	0.035**	0.051***	0.002	0.023**
	(0.015)	(0.014)	(0.016)	(0.010)	(0.010)
Year FE:	Yes	Yes	Yes	Yes	Yes
Wealth Bracket FE:	No	Yes	Yes	Yes	Yes
Number of observations	985,475	985,475	985,475	921,778	921,778
Adjusted R^2	0.071	0.072	0.082	0.072	0.083

Table V
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***	-0.050***	-0.043***	-0.076***	-0.067***
	(0.008)	(0.008)	(0.007)	(0.012)	(0.014)
Debt-to-income ratio			0.002		0.002
			(0.002)		(0.002)
Experience, Education, and Gender:					
Stock market experience	0.009***	0.009***	0.009***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Male dummy			-0.041***		-0.037**
			(0.012)		(0.014)
Master's degree dummy				-0.006	-0.004
				(0.009)	(0.010)
Finance education dummy				0.032***	0.033***
				(0.008)	(0.008)
Finance occupation dummy				0.075*	0.066
				(0.039)	(0.040)

(Continued)

Table V - *Continued*

	Dependent Variable: Wealth Factor Tilt				
	(1)	(2)	(3)	(4)	(5)
Wealth Percentile Dummies					
Bottom 10%	-0.041***	-0.040***	-0.039***	-0.042***	-0.040***
	(0.009)	(0.008)	(0.010)	(0.008)	(0.009)
10-20	-0.024***	-0.025***	-0.022***	-0.026***	-0.024***
	(0.008)	(0.007)	(0.006)	(0.007)	(0.005)
20-30	-0.018**	-0.018**	-0.016***	-0.019***	-0.017***
	(0.007)	(0.007)	(0.005)	(0.006)	(0.005)
30-40	-0.009**	-0.010***	-0.009***	-0.011***	-0.009***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
40-50	-0.00005	-0.0003	-0.0004	-0.00004	-0.00004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
60-70	-0.003	-0.003	-0.003	-0.004	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
70-80	-0.006	-0.006	-0.005	-0.005	-0.004
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
80-90	-0.001	-0.001	0.003	-0.00005	0.003
	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
90-99	0.028***	0.028***	0.033***	0.027***	0.032***
	(0.007)	(0.007)	(0.006)	(0.007)	(0.006)
99-99.9	0.091***	0.091***	0.100***	0.088***	0.094***
	(0.014)	(0.014)	(0.012)	(0.015)	(0.013)
Top 1%	0.149***	0.149***	0.157***	0.134***	0.141***
	(0.017)	(0.017)	(0.016)	(0.017)	(0.016)
Year FE:	Yes	Yes	Yes	Yes	Yes
Age Group FE:	No	Yes	Yes	Yes	Yes
Number of observations	985,475	985,475	985,475	921,778	921,778
Adjusted R^2	0.055	0.056	0.057	0.055	0.056

Table VI
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	7.00	9.00	16.00	9.00	8.00	10.00	13.00	5.00
Institutional ownership share (%)	5.23	6.35	4.32	-0.91	2.96	6.22	6.42	3.46
Turnover	5.2	2.2	0.4	-4.8	7.1	1.7	0.6	-6.6
Volatility (%)	23.9	13.8	8.4	-15.5	24.6	12.5	9.0	-15.6
CAPM beta	0.94	0.84	0.66	-0.28	0.88	0.83	0.67	-0.22
Size (NOK million)	508	1118	2128	1620	379	1368	1490	1111
BE/ME	0.55	0.66	0.89	0.34	0.72	0.70	0.67	-0.05
Profitability (%)	4.8	6.8	7.7	2.9	2.7	7.2	8.7	6.0
Investment growth (%)	8.7	6.6	6.5	-2.2	4.3	6.7	9.9	4.7

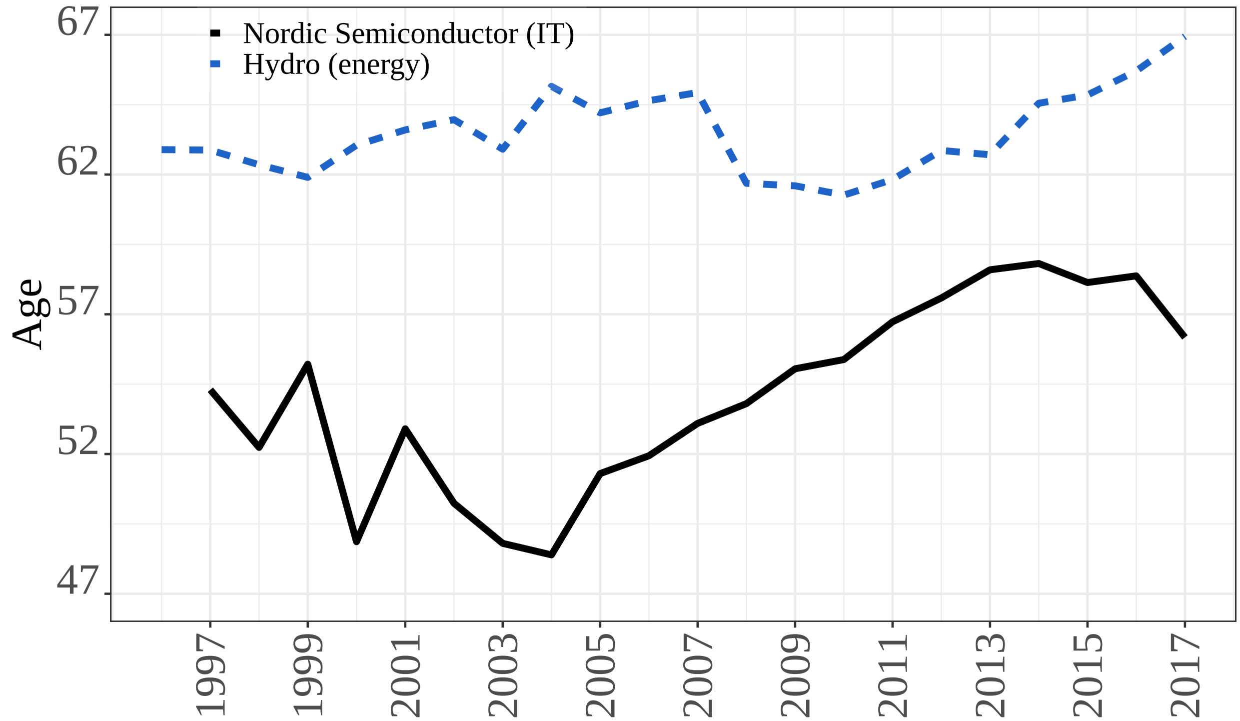
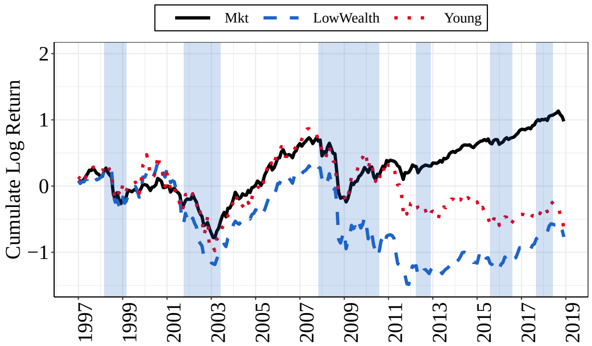
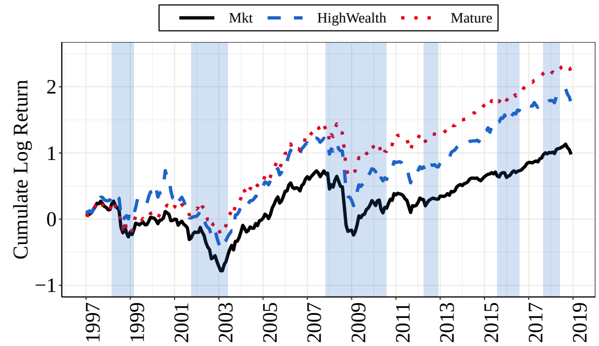


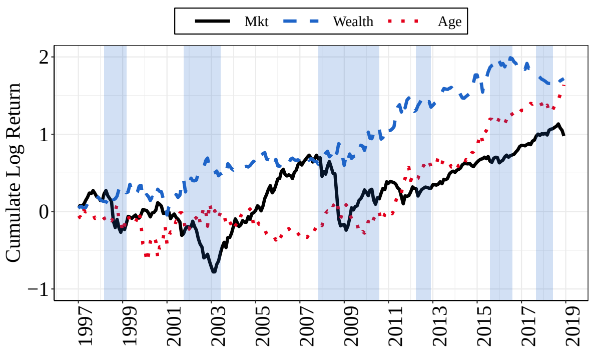
Figure 1 Investor age characteristic for two stocks. This figure plots the age characteristic of Norsk Hydro and Nordic Semiconductor in 1997-2018. Hydro is a fully integrated aluminium company. Nordic Semiconductor is a semiconductor company specializing in wireless technology. For each stock, the age 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.



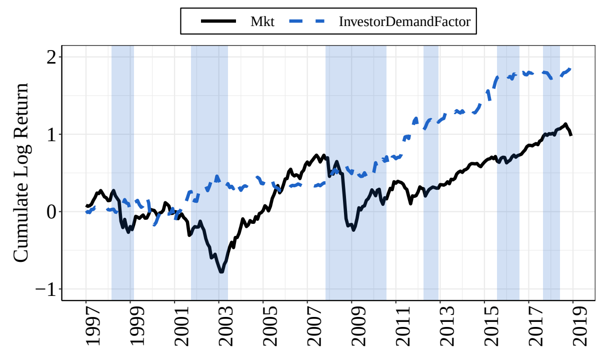
(a) Low wealth and young portfolios



(b) High wealth and mature portfolios



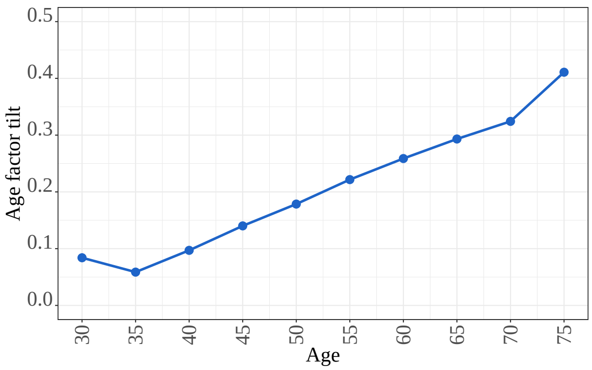
(c) Age and wealth factors



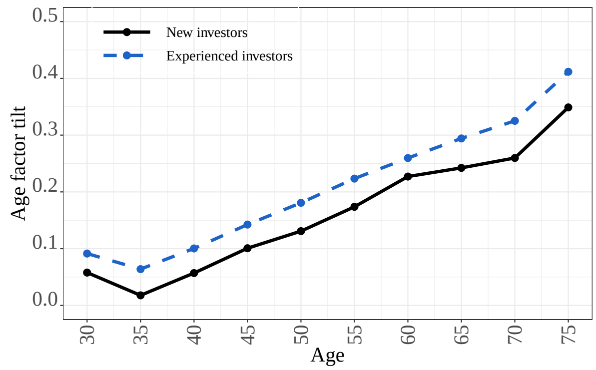
(d) Age and wealth factors combined

Figure 2 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.

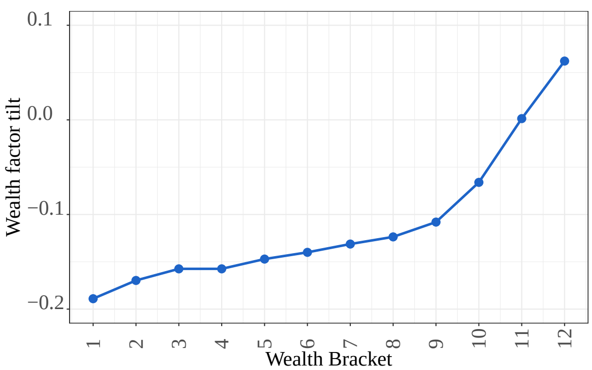
Figure 3 Comparison of models with firm factors and investor factors. This figure plots the average out-of-sample annualized Sharpe ratio of mean-variance efficient portfolios constructed from several sets of factors. The analysis follows the Fama and French (2018) out-of-sample bootstrap methodology. We consider 264 months of Norwegian stock return data from 1997 to 2018. We split the sample period data into 132 adjacent pairs and run 100,000 bootstrap simulations. In each simulation, we randomly assign one month from each pair to the in-sample dataset and the other month to the out-of-sample dataset. We estimate the mean-variance efficient portfolio from sets of factors observed over the in-sample period. We then calculate the Sharpe ratio over the out-of-sample period and report the average Sharpe ratio across all simulations. The factor models are categorized according to the number of factors other than the market that they contain. The Firm-3 factor model includes HML, SMB, and MOM, the Firm-4 model includes HML, SMB, RMW, and CMA, and the Firm-5 model includes HML, SMB, RMW, CMA, and MOM. Models containing investor factors are represented by a black dot.



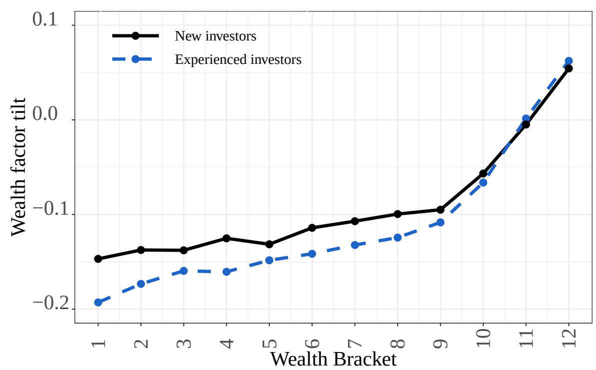
(a) Age tilt: All participating investors



(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 (black) and preexisting participants (blue) 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 (black) and preexisting participants (blue) 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.

Internet Appendix for "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*

March 2021

This Internet Appendix provides the proofs of the main theoretical results, describes the details of data construction, discusses the empirical methodology, and presents additional empirical evidence.

*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.

Contents

I. Appendix to Section 2: Theoretical Linkages Between Investor Portfolios and Pricing Factors

I.A. *Proof of Proposition 1*

In Section 2.1 of the main text, we show that the market portfolio satisfies

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

where $\eta_k^m = \sum_{i=1}^{I} E^i \eta_k^i / \sum_{i=1}^{I} E_i$ is the aggregate portfolio tilt toward the deviation portfolio $\boldsymbol{d_k}$.

We now derive a multi-beta asset pricing equation, which shows that the model has $K+1$ pricing factors: the market $\boldsymbol{m}$ and the K deviation portfolios $\boldsymbol{d_k}$. Let μ_m denote the expected return on the market portfolio and let $\boldsymbol{\mu_d}$ denote the vector of expected returns on the deviation portfolios.

By (IA-1) and the definition of the tangency portfolio, the vector of excess returns satisfies

$$\boldsymbol{\mu} - R_f \mathbf{1} = \phi \, \boldsymbol{\Sigma \tau}. \tag{IA-2}$$

The normalizing constant ϕ is the performance measure

$$\phi = \mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f \mathbf{1}) = \frac{\mu_\tau - R_f}{\sigma_\tau^2}, \tag{IA-3}$$

where $\mu_\tau = \boldsymbol{\mu}'\boldsymbol{\tau}$ and $\sigma_\tau^2 = \boldsymbol{\tau}'\boldsymbol{\Sigma\tau}$ denote, respectively, the drift and variance of the tangency portfolio.[1]

We use (IA-1) to substitute out the tangency portfolio,

$$\boldsymbol{\mu} - R_f\mathbf{1} = \phi\,\boldsymbol{\Sigma}\left(\boldsymbol{m} - \sum_{k=1}^{K}\eta_k^m\,\boldsymbol{d_k}\right). \tag{IA-4}$$

Let $\boldsymbol{\Sigma_{jm}} = \boldsymbol{\Sigma m}$ denote the $J \times 1$ covariance vector between the J assets and the market portfolio, $\boldsymbol{\Sigma_{jd}} = [\boldsymbol{\Sigma}\,\boldsymbol{d_1}, \cdots, \boldsymbol{\Sigma}\,\boldsymbol{d_K}]$ the $J \times K$ covariance matrix between the assets and the long-short portfolios, and $\boldsymbol{\Sigma_{j,md}} = [\boldsymbol{\Sigma_{jm}}, \boldsymbol{\Sigma_{jd}}]$. We can then write equation (IA-4) in vector form,

$$\boldsymbol{\mu} - R_f\mathbf{1} = \phi\,\boldsymbol{\Sigma_{j,md}}\begin{pmatrix}1\\ -\boldsymbol{\eta^m}\end{pmatrix}, \tag{IA-5}$$

where $\boldsymbol{\eta^m}$ is the $K \times 1$ vector of all the η_k^m.

From (IA-5), we can express the risk premium on the market $\boldsymbol{m}$ and the portfolios $\boldsymbol{d_k}$ as

$$\begin{pmatrix}\mu_m - R_f\\ \boldsymbol{\mu_d} - R_f\mathbf{1}\end{pmatrix} = \phi\,\boldsymbol{\Sigma_{md,md}}\begin{pmatrix}1\\ -\boldsymbol{\eta^m}\end{pmatrix}, \tag{IA-6}$$

where $\boldsymbol{\Sigma_{md,md}}$ is the $(K+1)\times(K+1)$ covariance matrix of the market and long-short

[1]The result can be derived as follows. The tangency portfolio has excess drift

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

The instantaneous variance of the tangency portfolio, $\sigma_\tau^2 = \boldsymbol{\tau}'\,\boldsymbol{\Sigma\tau}$, therefore satisfies

$$\sigma_\tau^2 = \frac{(\boldsymbol{\mu} - R_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f\mathbf{1})}{[\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f\mathbf{1})]^2} = \frac{\mu_\tau - R_f}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f\mathbf{1})},$$

which implies that (IA-3) holds.

portfolio returns. We infer that

$$\phi \begin{pmatrix} 1 \\ -\boldsymbol{\eta^m} \end{pmatrix} = \boldsymbol{\Sigma^{-1}_{md,md}} \begin{pmatrix} \mu_m - R_f \\ \boldsymbol{\mu_d} - R_f \mathbf{1} \end{pmatrix}. \tag{IA-7}$$

We plug (IA-7) into (IA-5) and obtain

$$\boldsymbol{\mu} - R_f \mathbf{1} = \boldsymbol{\beta'_{j,md}} \begin{pmatrix} \mu_m - R_f \\ \boldsymbol{\mu_d} - R_f \mathbf{1} \end{pmatrix}, \tag{IA-8}$$

where $\boldsymbol{\beta_{j,md}} = \boldsymbol{\Sigma^{-1}_{md,md}} \boldsymbol{\Sigma'_{j,md}}$. We let $\boldsymbol{\beta'_{j,md}} = [\boldsymbol{\beta'_{j,m}}, \boldsymbol{\beta'_{j,d}}]'$.

I.B. CAPM Alpha and Beta

We can re-write (IA-4) as follows:

$$\boldsymbol{\mu} - R_f \mathbf{1} = \phi \left(\sigma_m^2 \boldsymbol{b_m} - \sum_{k=1}^{K} \eta_k^m \sigma_k^2 \boldsymbol{b_k} \right), \tag{IA-9}$$

where $\boldsymbol{b_m}$ is the $J \times 1$ vector of univariate betas to the market factor, and $\boldsymbol{b_k}$ is the vector of univariate betas to the deviation portfolio $\boldsymbol{d_k}$.

The risk premium of the market portfolio is equal to

$$\mu_m - R_f = \phi \left(\sigma_m^2 - \sum_{k=1}^{K} \eta_k^m \sigma_k^2 b_{m,k} \right), \tag{IA-10}$$

where $b_{m,k}$ is the univariate beta of the market portfolio with respect to the deviation portfolio $\boldsymbol{d_k}$.

A stock's CAPM-alpha is defined as $a_j = \mu_j - R_f - b_{m,j}(\mu_m - R_f)$, where $b_{m,j}$ is the stock's univariate beta to the market portfolio. Combining (IA-9) and (IA-10) into the stock's alpha definition yields

$$\boldsymbol{a} = -\phi \sum_{k=1}^{K} \eta_k^m \, \sigma_k^2 \, (\boldsymbol{b_k} - b_{m,k}\mathbf{1}). \tag{IA-11}$$

Likewise, equation (IA-1) implies that the vector of market betas can be decomposed as follows:

$$\boldsymbol{b_m} = \frac{\sigma_\tau^2}{\sigma_m^2}\boldsymbol{b_\tau} + \sum_{k=1}^{K} \eta_k^m \, \frac{\sigma_k^2}{\sigma_m^2} \, \boldsymbol{b_k}. \tag{IA-12}$$

A stock's market beta is the weighted average of its beta to the tangency portfolio and its betas to the deviation portfolios, where the weights are driven by the ratio of portfolio variances to the market variance, σ_τ^2/σ_m^2 and σ_k^2/σ_m^2, and aggregate tilts, η_k^m.

I.C. Individual Portfolio Choice

In this Section, we compute the consumption-portfolio decision of an agent satisfying the specification outlined in Section 2.2 of the main text. We focus on the decision of a single agent and simplify notation by dropping the agent index i. We also asume without loss of generality that the agent is born at date 0.

The consumption-portfolio decision problem is defined as follows. The agent lives and consumes in periods $t = 0, \ldots, T$. At the beginning of every period t, she receives stochastic labor income L_t. The cash on hand available to her, W_t, is the sum of labor income, L_t, and the value at date t of previous investments in financial assets.

The agent uses cash on hand to consume C_t and invests the remainder, $W_t - C_t$, in financial assets. She can trade a riskless asset with net rate of return R_f and the stocks $j \in \{1, \ldots, J\}$. The stocks have excess returns $R_{j,t+1}^e$ between dates t, which we stack into the column vector $\boldsymbol{R_{t+1}^e}$. Let $\boldsymbol{\alpha_t} \in \mathbb{R}^J$ denote the weights of stocks in the agent's portfolio of stocks and the riskless asset. We emphasize that the components of $\boldsymbol{\alpha_t}$ need not add up to unity. By contrast, we denote by $\boldsymbol{\omega_t} = \boldsymbol{\alpha_t}/(\mathbf{1}'\boldsymbol{\alpha_t})$ the vector containing the weights of individual stocks in the agent's stock portfolio.

With this notation, the cash on hand available to the agent is

$$W_{t+1} = L_{t+1} + (W_t - C_t)\left(1 + R_f + \sum_{j=1}^{J} \alpha_{j,t}\, R^e_{j,t+1}\right) \tag{IA-13}$$

at the beginning of period $t+1$.

The agent selects the consumption-portfolio plan $\{(C_t, \boldsymbol{\alpha_t})\}$ that maximizes the lifetime utility

$$\mathbb{E}_0\left[\sum_{t=0}^{T} \beta^t\, u(C_t)\right], \tag{IA-14}$$

subject to the budget constraint (IA-13).

We make the following assumptions on the stochastic processes driving the economy. Labor income is given by:

$$L_t = L_{t-1}\, g_t, \tag{IA-15}$$

where the growth rates g_t are independent and identically distributed. We also assume that there exists a state vector $\boldsymbol{y_t} \in \mathbb{R}^K$ that drives returns. The growth rate g_t and the state vector $\boldsymbol{y_t}$ are known to the agent at the beginning of period t. Conditional on $\boldsymbol{y_t}$, the distribution of $(R_{1,t+1}, \ldots, R_{J,t+1}, g_{t+1})$ is multivariate lognormal.

Let $J(t, W_t, L_t, \boldsymbol{y_t})$ denote the value function of the agent at t. The value function satisfies the Bellman equation

$$J(t, W_t, L_t, \boldsymbol{y_t}) = \max_{\{\boldsymbol{\alpha_t}, C_t\}} \left[u(C_t) + \delta\, \mathbb{E}_t J(t+1, W_{t+1}, L_{t+1}, \boldsymbol{y_{t+1}})\right]$$

At the terminal date T, the value function is $J(T, W_T, L_T, \boldsymbol{y_T}) = u(W_T)$.

The optimal portfolio $\boldsymbol{\alpha_t}$ is a linear combination of the tangency portfolio, a portfolio providing a hedge against labor income, and a portfolio providing a hedge against time-varying investment opportunities, as the following proposition shows.

Proposition IA.1. *The optimal portfolio* $\boldsymbol{\alpha_t}$ *is approximately given by*

$$\boldsymbol{\alpha_t} = -\frac{J_W}{(W_t - C_t)J_{WW}}\boldsymbol{\Sigma_t^{-1}}(\boldsymbol{\mu_t} - R_f\mathbf{1}) - \frac{L_t}{W_t - C_t}\left(1 + \frac{J_{WL}}{J_{WW}}\right)\boldsymbol{\Sigma_t^{-1}}\,\boldsymbol{b_t} - \frac{\boldsymbol{\Sigma_t^{-1}}\,\boldsymbol{D_t}\,\boldsymbol{J_{Wy}}}{(W_t - C_t)\,J_{WW}},$$

where the derivatives of the value function are evaluated at $(\mathbb{E}_t W_{t+1}, \mathbb{E}_t L_{t+1}, \boldsymbol{\mathbb{E}_t y_{t+1}})$,

$$\begin{aligned} \boldsymbol{b_t} &= \mathbb{E}_t[(g_{t+1} - \mathbb{E}_t g_{t+1})\boldsymbol{R^e_{t+1}}], \\ \boldsymbol{D_t} &= \mathbb{E}_t[\boldsymbol{R^e_{t+1}}\,(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}})'], \end{aligned}$$

for every $t < T$.

Proof. We note that wealth satisfies satisfies the following moment identities:

$$\mathbb{E}_t(W_{t+1}) = L_t\mu_g + (W_t - C_t)\left[1 + R_f + \boldsymbol{\alpha_t'}(\boldsymbol{\mu_t} - R_f\,\mathbf{1})\right], \tag{IA-16}$$

$$\mathrm{Var}_t(W_{t+1}) = L_t^2\,\sigma_g^2 + (W_t - C_t)^2\,\boldsymbol{\alpha_t'\Sigma\alpha_t} + 2(W_t - C_t)\,L_t\,\boldsymbol{\alpha_t'}\,\boldsymbol{b_t}, \tag{IA-17}$$

$$\mathrm{Cov}_t(W_{t+1}, L_{t+1}) = L_t^2\,\sigma_g^2 + L_t\,(W_t - C_t)\,\boldsymbol{\alpha_t' b_t}, \tag{IA-18}$$

$$\mathbb{E}_t[(W_{t+1} - \mathbb{E}_t W_{t+1})\,(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1})'}] = L_t\,\boldsymbol{f_t} + (W_t - C_t)\,\boldsymbol{\alpha_t' D_t}, \tag{IA-19}$$

where $\mu_g = \mathbb{E}_t(g_{t+1})$, $\sigma_g^2 = \mathrm{Var}_t(g_{t+1})$, and

$$\boldsymbol{f_t} = \mathbb{E}_t[g_{t+1}\,(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}})'].$$

We consider a quadratic expansion of $J(t+1, W_{t+1}, L_{t+1}, \boldsymbol{y_{t+1}})$ around the conditional mean vector $(\mathbb{E}_t W_{t+1}, \mathbb{E}_t L_{t+1}, \boldsymbol{\mathbb{E}_t y_{t+1}})$. We obtain

$$\begin{aligned} J(t+1, W_{t+1}, L_{t+1}, \boldsymbol{y_{t+1}}) \;\approx\; & J(t+1, \mathbb{E}_t W_{t+1}, \mathbb{E}_t L_{t+1}, \boldsymbol{\mathbb{E}_t y_{t+1}}) + \frac{1}{2}(W_{t+1} - \mathbb{E}_t W_{t+1})^2 J_{WW} \\ & + J_{WL}\,(W_{t+1} - \mathbb{E}_t W_{t+1})\,(L_{t+1} - \mathbb{E}_t L_{t+1}) \\ & + \boldsymbol{J'_{Wy}}\,(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}})\,(W_{t+1} - \mathbb{E}_t W_{t+1}) + Q_{t+1} + \varepsilon_{t+1}, \end{aligned}$$

where $\varepsilon_{t+1} = J_W\,(W_{t+1} - \mathbb{E}_t W_{t+1}) + J_L\,(L_{t+1} - \mathbb{E}_t L_{t+1}) + \boldsymbol{J'_y\,(y_{t+1} - \mathbb{E}_t y_{t+1})}$, and

$$\begin{aligned} Q_{t+1} \;=\; & \frac{1}{2}(L_{t+1} - \mathbb{E}_t L_{t+1})^2\,J_{LL} + \frac{1}{2}(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}})'\,\boldsymbol{J_{yy}}\,(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}}) \\ & + \boldsymbol{J'_{Ly}}\,(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}})\,(L_{t+1} - \mathbb{E}_t L_{t+1}). \end{aligned}$$

We note that $\mathbb{E}_t \varepsilon_{t+1} = 0$.

The conditional expectation at t of the value function $J(t+1, W_{t+1}, L_{t+1}, \boldsymbol{y_{t+1}})$ satisfies

$$\begin{aligned} \mathbb{E}_t J(t+1, W_{t+1}, L_{t+1}, \boldsymbol{y_{t+1}}) \;\approx\; & J(t+1, \mathbb{E}_t W_{t+1}, \mathbb{E}_t L_{t+1}, \boldsymbol{\mathbb{E}_t y_{t+1}}) + \frac{1}{2} Var_t(W_{t+1})\,J_{WW} \\ & + J_{WL}\,\mathrm{Cov}_t(W_{t+1}, L_{t+1}) \\ & + \boldsymbol{J'_{Wy}}\,\mathbb{E}_t[(\boldsymbol{y_{t+1}} - \boldsymbol{\mathbb{E}_t y_{t+1}})\,(W_{t+1} - \mathbb{E}_t W_{t+1})] + \mathbb{E}_t Q_{t+1}. \end{aligned}$$

We infer that

$$\begin{aligned} \mathbb{E}_t J(t+1, W_{t+1}, L_{t+1}, \boldsymbol{y_{t+1}}) \;\approx\; & J(t+1, \mathbb{E}_t W_{t+1}, \mathbb{E}_t L_{t+1}, \boldsymbol{\mathbb{E}_t y_{t+1}}) \\ & + \frac{1}{2}[L_t^2\,\sigma_g^2 + (W_t - C_t)^2\,\boldsymbol{\alpha'_t \Sigma_t \alpha_t} + 2(W_t - C_t) L_t\,\boldsymbol{\alpha'_t b_t}] J_{WW} \\ & + J_{WL}\,\left[L_t^2\,\sigma_g^2 + L_t\,(W_t - C_t)\,\boldsymbol{\alpha'_t b_t}\right] \\ & + \boldsymbol{J'_{Wy}}\,\left[L_t\,\boldsymbol{f_t} + (W_t - C_t)\,\boldsymbol{D'_t \alpha_t}\right] + \mathbb{E}_t Q_{t+1}. \end{aligned}$$

The first-order condition with respect to $\boldsymbol{\alpha_t}$ is therefore[2]

$$(\boldsymbol{\mu_t} - R_f\mathbf{1})J_W + [(W_t - C_t)\boldsymbol{\Sigma_t}\,\boldsymbol{\alpha_t} + L_t\,\boldsymbol{b_t}]J_{WW} + J_{WL}\,L_t\,\boldsymbol{b_t} + \boldsymbol{D_t}\,\boldsymbol{J_{Wy}} = 0,$$

or equivalently

$$(\boldsymbol{\mu_t} - R_f\mathbf{1})J_W + (W_t - C_t)\,J_{WW}\,\boldsymbol{\Sigma_t}\,\boldsymbol{\alpha_t} + L_t\,(J_{WW} + J_{WL})\,\boldsymbol{b_t} + \boldsymbol{D_t}\,\boldsymbol{J_{Wy}} = 0.$$

We conclude that the proposition holds. ■

The expression for the optimal portfolio $\boldsymbol{\alpha_t}$ considerably simplifies in the last trading period $t = T - 1$, which is a cornerstone of the analysis. Indeed, since the value function is $J(T, W_T, L_T, \boldsymbol{y_T}) = u(W_T)$ at the terminal date, its partial derivatives satisfy $J_{WL} = 0$, $\boldsymbol{J_{Wy}} = 0$ at $(T, \mathbb{E}_{T-1}W_T, \mathbb{E}_{T-1}L_T, \boldsymbol{\mathbb{E}_{T-1}y_T})$. The optimal portfolio at date $T-1$ reduces to

$$\boldsymbol{\alpha_{T-1}} = -\frac{J_W}{(W_{T-1} - C_{T-1})J_{WW}}\boldsymbol{\Sigma_{T-1}^{-1}}(\boldsymbol{\mu_{T-1}} - R_f\mathbf{1}) - \frac{L_{T-1}}{W_{T-1} - C_{T-1}}\boldsymbol{\Sigma_{T-1}^{-1}}\,\boldsymbol{b_{T-1}}.$$

Since the agent stops trading at date $T-1$, the hedging demand against adverse variation in future investment opportunities drops out from the optimal portfolio. Furthermore, if $L_{T-1} = 0$, the agent does not need to hedge against labor income shocks and the optimal portfolio of stocks in the financial portfolio becomes proportional to the tangency portfolio: $\boldsymbol{\alpha_{T-1}} = -(W_{T-1} - C_{T-1})^{-1}\,J_{WW}^{-1}\,J_W\,\boldsymbol{\Sigma_{T-1}^{-1}}(\boldsymbol{\mu_{T-1}} - R_f\mathbf{1})$. The optimal portfolio of stocks in the agent's stock portfolio, $\boldsymbol{\omega_{T-1}} = \boldsymbol{\alpha_{T-1}}/\mathbf{1'}\boldsymbol{\alpha_{T-1}}$ is then equal to the tangency porfolio:

$$\boldsymbol{\omega_{T-1}} = \boldsymbol{\tau_{T-1}}.$$

A mature investor with a large financial wealth-to-labor income ratio therefore holds the mean-variance efficient portfolio $\boldsymbol{\tau_{T-1}}$ and has a maximal Sharpe ratio.

[2] We neglect the higher-order terms involving the derivatives of $\mathbb{E}_t(Q_{t+1})$ with respect to $\boldsymbol{\alpha_t}$.

If the agent has constant relative risk aversion (CRRA), $u(c) = C^{1-\gamma}/(1-\gamma)$, the optimization problem is homogenous in (W_t, L_t). Let A_t denote the age of the agent at date t. The optimal portfolio of stocks can then be written as

$$\boldsymbol{\omega_t} = \boldsymbol{\omega}\left(A_t, \frac{L_t}{W_t}, \boldsymbol{y_t}\right). \tag{IA-20}$$

We know in addition that $\boldsymbol{\omega}(T-1, 0, \boldsymbol{y_{T-1}}) = \boldsymbol{\tau_{T-1}}$ in the agent's last period of trading.

I.D. Linearization and Factor Structure

We now consider a cross-section of investors $i = 1, \ldots, N$ at date t. As is explained in the main text, each investor i is born at date b^i and dies at date $b^i + T$. Investors have the same CRRA utility, the same lifespan T and are exposed to same labor income growth g_t. These restrictions could be lifted in future work, but the current setup is parsimonious and sufficient to account for the main empirical regularities reported in the main text.

The stock portfolio of each investor i in period t is given by:

$$\boldsymbol{\omega_t^i} = \boldsymbol{\omega}\left(A_t^i, \frac{L_t^i}{W_t^i}, \boldsymbol{y_t}\right),$$

as equation (IA-20) implies. We linearize $\boldsymbol{\omega}(\,\cdot\,,\,\cdot\,, \boldsymbol{y_t})$ around $A_t^i = T-1$ and $L_t^i/W_t^i = 0$:

$$\boldsymbol{\omega_t^i} = \boldsymbol{\tau_t} + (T-1-A_t^i)\,\boldsymbol{d_{1,t}} + \frac{L_t^i}{W_t^i}\,\boldsymbol{d_{2,t}}.$$

We now explain how this linearization can be achieved.

The linearization is based on two principles. First, the agent trades frequently between the start and the end of her trading life, that is between dates 0 and $T-1$ for an agent born at date 0. The behavior of the agent between 0 and $T-1$ can therefore be approximated by a continuous-time model when the trading frequency is sufficiently high. Second, the agent stops trading in the last period of her life, which has a fixed length equal to one unit of time. This second condition is technically useful because it implies that the agent remains sensitive to income and wealth at date $T-1$.[3] The corresponding utility function can therefore be written as

$$\sum_{s=0}^{h\,(T-1)} \beta^{s/h} u(C_s) + \beta^T u(C_T),$$

where h denotes the number of trading periods per unit of time.

The auxiliary continuous-time economy is defined as follows. We consider an agent with lifetime utility:

$$\mathbb{E}_0 \left[\int_0^{T-1} e^{-\rho\,(s-t)}\, u(c_s)\, ds + e^{-\rho\,(T-t)}\, u(c_T) \right]. \tag{IA-21}$$

The Bernoulli utility $u(\cdot)$ is same as in the discrete time model. The time discount rate is $\rho = -\ln(\beta)$, where β is the psychological discount factor of the discrete-time model.

The agent trades the riskless asset and the N stocks in continuous time between dates 0 and $T-1$. We assume for simplicity that the instantaneous riskless rate is $r = \ln(1+R_f)$. The joint dynamics of the stock returns, labor income flow, and state vector $\boldsymbol{y_t}$ are given by:

$$\boldsymbol{dR_t} = \boldsymbol{\mu(y_t)}\, dt + \boldsymbol{\sigma(y_t)}\, \boldsymbol{dz_t}, \tag{IA-22}$$

$$dL_t = L_t\,(\mu_L\, dt + \boldsymbol{\sigma_L}\, \boldsymbol{dz_t}), \tag{IA-23}$$

$$\boldsymbol{dy_t} = \boldsymbol{\mu_y(y_t)}\, dt + \boldsymbol{\sigma_y(y_t)}\, \boldsymbol{dz_t} \tag{IA-24}$$

We assume that the dynamics of these variables in discrete time, as explained in Section I.C, converge to their continuous-time equivalent, as given by IA-22 to IA-24, when the number of trading periods h goes to infinity. Since the convergence of discrete-time processes to

[3] If the length of the last period of life were allowed to converge to zero, the stock portfolio would converges to the tangency portfolio and the derivative of the optimal stock portfolio with respect to the labor income-to-wealth ratio would be zero in the limiting economy. A second order expansion of $\boldsymbol{\omega}$ would therefore be required to capture the link between the stock portfolio of wealth. Our assumption that the last investment period remains finite therefore allows us instead to rely on a more parsimonious, first-order Taylor expansion of $\boldsymbol{\omega_{T-1}}$.

Itô diffusions is the subject of a vast literature, we do not need to make specific parametric assumptions and only require that the convergence property holds.

We note that

$$\begin{aligned} \mathbb{E}_{\boldsymbol{t}}(\boldsymbol{dR_t}\, dL_t) &= \boldsymbol{b_t}\, dt, \\ \mathbb{E}_{\boldsymbol{t}}(\boldsymbol{dR_t}\, \boldsymbol{dy_t'}) &= \boldsymbol{D_t}\, dt, \\ \mathbb{E}_{\boldsymbol{t}}(\boldsymbol{dR_t}\, \boldsymbol{dR_t'}) &= \boldsymbol{\Sigma_t}\, dt, \end{aligned}$$

where $\boldsymbol{b_t} = \boldsymbol{\sigma(y_t)}\, \boldsymbol{\sigma_L'}$, $\boldsymbol{D_t} = \boldsymbol{\sigma(y_t)}\, \boldsymbol{\sigma_y(y_t)'}$, and $\boldsymbol{\Sigma_t} = \boldsymbol{\sigma(y_t)\sigma(y_t)'}$.

We denote by W_t the wealth, by c_t the consumption, and by α_t the vector of risky asset weights of the agent at t. The budget constraint is:

$$dW_t = \left[L_t + W_t\, r + W_t\, \boldsymbol{\alpha_t'}(\boldsymbol{\mu} - r\mathbf{1}) - c_t\right]\, dt + W_t\, \boldsymbol{\alpha_t'\sigma dz_t} \tag{IA-25}$$

at every instant t. The agent chooses $\{(\alpha_s, c_s)\}$ that maximize the lifetime utility (IA-21) under the budget constraint (IA-25). We denote by $V(t, L_t, W_t, y_t; t)$ the corresponding value function.

Since the agent has CRRA utility, $u(C) = C^{1-\gamma}/(1-\gamma)$, the homogeneity of the problem implies that the optimal portfolio of the continuous-time problem can be written as:

$$\boldsymbol{\omega_t} = \boldsymbol{\omega}\left(A_t, \frac{L_t}{W_t}, \boldsymbol{y_t}\right). \tag{IA-26}$$

Furthermore, we know that $\boldsymbol{\omega_t}$ coincides with the portfolio of the discrete-time economy if $A_t = T-1$. In particular, $\boldsymbol{\omega}(T-1, 0, \boldsymbol{y_t}) = \boldsymbol{\tau}_t$. We linearize $\boldsymbol{\omega}$ around $(T-1, 0, \boldsymbol{y_t})$ and obtain:

$$\boldsymbol{\omega}\left(t, \frac{L_t}{W_t}, \boldsymbol{y_t}\right) = \boldsymbol{\tau_t} + (T-1-t)\, \boldsymbol{d_{1,t}} + \frac{L_t}{W_t}\, \boldsymbol{d_{2,t}}$$

where

$$\boldsymbol{d_{1,t}} = -\frac{\partial \boldsymbol{\omega}}{\partial A}(T-1, 0, \boldsymbol{y_t}),$$

$$\boldsymbol{d_{2,t}} = \frac{\partial \boldsymbol{\omega}}{\partial L/W}(T-1, 0, \boldsymbol{y_t}).$$

The cross-section of portfolios can therefore be approximated by a three-factor model, in which the portfolio of each investor is the linear combination of the tangency portfolio, a portfolio linked to age, and a portfolio linked to wealth.

Furthermore, Proposition IA.1 has a direct counterpart in continuous time.

Proposition IA.2. *The optimal portfolio is given by*

$$\boldsymbol{\alpha_t} = -\frac{V_W}{W_t\, V_{WW}}\boldsymbol{\Sigma_t^{-1}}(\boldsymbol{\mu_t} - r\mathbf{1}) - \frac{L_t}{W_t}\frac{V_{LW}}{V_{WW}}\boldsymbol{\Sigma_t^{-1}}\boldsymbol{b_t} - \frac{\boldsymbol{\Sigma_t^{-1}}\, \boldsymbol{D_t}\, \boldsymbol{V_{Wy}}}{V_{WW}},$$

at every $t \leq T-1$.

Proof. The optimal policy function satisfies the Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} 0 = \quad \max\nolimits_{c_t, \boldsymbol{\alpha_t}} \quad & [u(c_t)dt - \rho V dt + V_t\, dt + V_L\ \mathbb{E}_t(dL_t) + V_W \mathbb{E}_t(dW_t) \\ & + \boldsymbol{V_y' \mathbb{E}_t(dy_t)} + \tfrac{1}{2}V_{WW}\,(dW_t)^2 + \tfrac{1}{2}V_{LL}(dL_t)^2 + \tfrac{1}{2}\boldsymbol{dy_t' V_{yy} dy_t} \\ & + \boldsymbol{V_{Wy}' dy_t}\, dW_t + V_{LW}\, dL_t\, dW_t + \boldsymbol{V_{Ly}' dy_t}\, dL_t] \end{aligned}$$

or equivalently

$$\begin{aligned} 0 = \quad \max\nolimits_{c_t, \boldsymbol{\alpha_t}} \quad & [u(c_t) - \rho V + V_t + V_L\ L_t\ \mu_L + V_W\,[L_t + W_t\, r + W_t\, \boldsymbol{\alpha_t}'(\boldsymbol{\mu} - r) - c_t] \\ & + \boldsymbol{V_y' \mu_y(y_t)} + \tfrac{1}{2}V_{WW}\, W_t^2\, \boldsymbol{\alpha_t' \Sigma_t \alpha_t} + \tfrac{1}{2}V_{LL}\ L_t^2\, \|\boldsymbol{\sigma_L}\|^2 \\ & + \tfrac{1}{2}\mathrm{tr}(\boldsymbol{\sigma_y'}\, \boldsymbol{V_{yy'}}\, \boldsymbol{\sigma_y}) + W_t \boldsymbol{\alpha_t' \sigma \sigma_y'}\, \boldsymbol{V_{Wy}} + L_t\, W_t \boldsymbol{\alpha_t' \sigma \sigma_L'} V_{LW} \\ & + L_t \boldsymbol{V_{Ly}' \sigma_y \sigma_L'}] \end{aligned}$$

where $\Sigma = \sigma\sigma'$.

We write the first-order condition with respect to $\boldsymbol{\alpha_t}$:

$$V_W \, W_t \, (\boldsymbol{\mu} - r \, \mathbf{1}) + V_{WW} \, W_t^2 \boldsymbol{\Sigma_t \alpha_t} + W_t \boldsymbol{D_t} \, \boldsymbol{V_{Wy}} + L_t W_t \, \boldsymbol{b_t} \, V_{LW} = 0$$

and conclude that the proposition holds. ■

II. Appendix to Section 3: Data and Construction of Equity Factors

II.A. Firm Factors

In Section 3.3 of the main text, we summarize how we construct Norwegian factors from firm characteristics. We now provide more details about the construction methodology and present statistics about the pricing performance of these factors.

The market portfolio is the value-weighted portfolio of all the stocks in the universe of Norwegian stocks we consider. For the risk-free rate, we follow Oedegaard (2020) and use the monthly interbank rate, NIBOR, which we download from his website.[4]

The construction of firm factors is based on the following general rules. The momentum factor is rebalanced monthly. Other firm factors are rebalanced yearly. In the calculation of all factors, stocks are weighted by their free-float adjusted market weights (NOSHFF).[5] We now turn to the details of construction of each factor.

The size factor, SMB_t, is based on each firm's market value of equity in year t, which we obtain by multiplying the end-of-June closing price by the number of shares outstanding. Our share price variable corrects for stock splits, dividends, and other corporate events. If

[4] http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html

[5] In cases where the free-float number of share is not available, we use the total number of shares.

the closing price is unavailable, we proxy it by the bid or ask price. The long leg of the size factor contains stocks that are larger than the median, while the short leg contains stocks smaller than the median.

The value, profitability, and investment factors are defined from sorted portfolios as follows. We rank stocks by the selected characteristic (book-to-market ratio, gross profit, investment growth) at the end of year $t-1$. The Low (L) portfolio contains stocks below the 30^{th} percentile, the Middle (M) portfolio contains stocks between the 30^{th} and the 70^{th} percentiles, and the High (H) portfolio contains stocks above the 70^{th} percentile. The factor is the portfolio that is long the High (H) portfolio and short the Low (L) portfolio. Our methodology is similar to Stambaugh and Yuan (2017), with the exception that they use the 20^{th} and 80^{th} percentiles as thresholds. Our motivation for using the 30^{th} and 70^{th} percentiles stems from the relatively small cross-section of stocks listed on the Oslo Stock Exchange. We choose percentiles that allow us to include more stocks in the High and Low portfolios, thereby ensuring that our factors are well diversified.

The value factor, HML_t, is based on firms' book-to-market ratio. We calculate the book value of equity using accounting data from NHH for the 1997 to 2011 period and from Datastream for 2012 onward. For the 1997-2011 sub-period, the book value of equity is the sum of the book value of stockholders' equity (*Sum egenkapital*) and the balance sheet deferred taxes and investment credit (*Utsatt skatt*), net of capital raised through preference shares (*Preferansekapital*). The book-to-market ratio for year t is the ratio of the book value of equity at the end of December of year t-1 to the market value of equity at the end of year t-1. For the period starting in 2012, we rely on Datastream and use common equity (Datastream code WC03501) on December 31 divided by the closing market value of the last trading day in December. If this information is not available, we use the inverse of the price-to-book value ratio (Datastream code WC09304).

Cohen and Polk (1998) and Cohen et al. (2003) show that, if book-to-market ratios are decomposed into an across-industry component and a within-industry component, then only the within-industry component is priced. Building on their finding, we adjust for industry effects by subtracting the industry's mean book-to-market ratio from each firm's book-to-market ratio. Industries are defined by 11 distinct 2-digit GICS codes. To reduce noise, we only subtract the industry average from firm book-to-market ratios if the industry average is based on five or more companies in a given year. We use the demeaned book-to-market ratio to form the Low, Medium, and High portfolios of the value factor. The industry adjustment leads to a small increase in the Sharpe Ratio of the resulting value factor.

The profitability factor, RMW_t, is defined as in Novy-Marx (2013). For the 1997-2011 sub-period, gross profit, GP_t, is the difference between total revenue (*Sum inntekter*) and cost of goods sold (*Vareforbruk*). Let $TA_{t-1,Dec}$ denote the value of total assets (*Sum eiendeler*). The profitability characteristic available in June of year t is $PROF_{t,June} = GP_{t-1,Dec}/TA_{t-1,Dec}$. For the 2012-2018 subsample, we define gross profit $GP_{t-1,Dec}$ as revenues (Datastream code WC01001) minus cost of goods sold (Datastream code WC01051).

The investment factor, CMA_t, is defined as in Fama and French (2015). We define a firm's investment growth from year $t-2$ to year $t-1$ as the growth rate of total assets $INV_{t,June} = (TA_{t-1,Dec} - TA_{t-2,Dec})/TA_{t-2,Dec}$.

The momentum factor, MOM_t, is constructed as in Carhart (1997). A stock's momentum characteristic is its geometric return over the previous 12 months, excluding the most recent month from consideration: $MOM_t = \prod_{\ell=2}^{12}(1 + r_{t-\ell}) - 1$. We sort stocks according to their momentum. The Low (L) portfolio contains stocks up to the 30^{th} percentile and the High (H) portfolio contains stocks above the 70^{th}. The momentum factor is long the High portfolio and short the Low portfolio.

Table IA.1 reports pricing statistics for the firm factors in Norway between 1997 and 2018. The Sharpe ratio of the market factor, MKT_t, is 0.32. The profitability, investment, and momentum factors produce statistically significant CAPM-alphas during the sample period.

III. Appendix to Section 4: Pricing Performance of Investor Factors

In this Section, we propose alternative specifications of investor factors and evaluate their pricing performance.

III.A. Age and Wealth Factors Constructed from Investor Stockholdings

In Section 3.2 of the main text, we construct investor factors in two steps. First, we estimate for each stock the average age and wealth of its individual investors. This step defines *stock-level* characteristics of the retail investor base. Second, for each of these stock-level characteristics, we sort stock by the characteristic and construct the portfolio that is long stocks in the top 30^{th} percentile and short stocks in the bottom 30^{th} percentile.

We now present an alternative method that builds investor factors directly from investor stock holdings, without requiring the calculation of *stock-level* investor characteristics. The alternative construction proceeds as follows. We compute the age variable, age_t^i, and net worth, $wealth_t^i$, of each investor i as defined in Section 3.2 of the main text. For each investor characteristic $c_t^i \in \{age_t^i, wealth_t^i\}$, we construct an alternative pricing factor by following the methodology outlined in Section 2.2 of the main text. Specifically, let $\mathcal{I}^H$ denote the set of investors who have a value of c_t^i above a high threshold c_t^H, and let $\mathcal{I}^L$ denote the set of investors with a value of c_t^i below a low threshold, c_t^L. The choice of thresholds is discussed further in the section.

We assign to each investor $i \in \mathcal{I}^H$ the positive weight

$$z_{i,c,t} = +\frac{E_{i,t}}{\sum_{i' \in \mathcal{I}^H} E_{i',t}},$$

where $E_{i,t}$ denotes the investor's wealth directly invested in equities at time t. Likewise, we assign to each investor $i \in \mathcal{I}^L$ the negative weight $z_{i,c,t} = -E_{i,t}/(\sum_{i' \in \mathcal{I}^L} E_{i',t})$. A zero weight is assigned to other investors $i \notin \mathcal{I}^H \cup \mathcal{I}^L$.

The alternative investor factor is the weighted sum of the investors' portfolio holdings:

$$\boldsymbol{g_{c,t}} = \sum_{i=1}^{I} z_{i,c,t}\, \boldsymbol{\omega_t^i}.$$

The specification of the investor weights guarantees that $\sum_{i=1}^{I} z_{i,c,t} = 0$, so that $\mathbf{1}'\boldsymbol{g_{c,t}} = 0$.

For the alternative age factor, we set the high threshold to the top 20$^{\text{th}}$ percentile of the age distribution every year and the low threshold to the bottom 20$^{\text{th}}$ percentile. The alternative age factor is therefore long the portfolios of mature investors and short the portfolios of young investors.

For the alternative wealth factor, we set the high threshold to the 90$^{\text{th}}$ percentile and the bottom threshold to the 50$^{\text{th}}$ percentile of the distribution of net worth. The wealth factor is therefore long the portfolios of wealthy investors and short the portfolios of less wealthy investors.

Table IA.2 presents pricing statistics on the alternative investor factors. Consistent with the baseline results in Table 2 of the main text, the alternative age and wealth factors have positive expected returns, positive alphas, and negative betas. We note that the expected return and alpha of the alternative factors are smaller than the baseline results in Table 2. The explanation is that the long and short legs of the alternative factors do not differ from each other as much as their baseline equivalents. Under the alternative construction presented in this section, the long and short legs are aggregate portfolios of investor holdings and thus include many similar stocks (albeit in different proportions). By contrast, the baseline method generates no overlap between stocks in the long leg and stocks in the short leg, which produces a greater return spread between the two legs.

III.B. Age and Wealth Factors Constructed from a Subsample of Investors

In Section 5.1 of the main text, we split the sample of investors into two sub-samples. The first sample includes two-thirds of investors and is used to reconstruct the age and wealth factors. The second sample is used to run regressions of investor factor tilts on their socioeconomic characteristics.

We obtain the sub-samples as follows. We split the investor population into 12 groups according to their wealth. The groups are specified in Section 3.2 of the main text. We then randomly draw two-thirds of investors from each group to form the first sub-sample. The wealth stratification guarantees that both sub-samples contain investors from the right tail of the skewed wealth distribution.

Table IA.3 presents the summary statistics of the investor factors constructed from the sub-samples. The results are equivalent to Table 3 of the main text, which uses the full sample. Specifically, we find that the age and wealth factors have positive expected returns, positive alphas, and negative betas. We note that the factors constructed from the sub-sample of investors are statistically less powerful than the full-sample factors, as evidenced by the smaller t-values.

III.C. Income-to-Wealth Ratio and Retirement Factors

Under the ICAPM developed in Section 2.2 of the main text, investor factors can be recovered by sorting investors by age and by the income-to-wealth ratio. We now show that the combination of the market factor, the age factor, and a factor constructed from investors' income-to-wealth ratio delivers similar pricing results as the baseline model. We also provide an alternative specification of the age factor as an additional robustness test.

We proceed with the two-step approach developed in Section 3.2 of the main text. Let $iwr_t^i = income_t^i / wealth_t^i$ denote the income-to-wealth ratio of investor i at time t. We calculate the income-to-wealth ratio characteristic of stock j as:

$$\mathrm{IW}_{j,t} = \frac{\sum_{i=1}^{I} N_{j,t}^i \, iwr_t^i}{\sum_{i=1}^{I} N_{j,t}^i} \tag{IA-27}$$

for every j and t.

For the alternative age factor, we create a dummy variable ret_t^i that is equal to unity if the investor is above 62 at time t and zero otherwise, corresponding to the usual retirement age. We then calculate the retiree characteristic of a stock as

$$\mathrm{Retiree}_{j,t} = \frac{\sum_{i=1}^{I} N_{j,t}^i \, ret_t^i}{\sum_{i=1}^{I} N_{j,t}^i} \tag{IA-28}$$

for every j and t.

We then sort stocks according to their characteristic $C_{j,t} \in \{\mathrm{Retiree}_{j,t}, \mathrm{IW}_{j,t}\}$ to form value-weighted portfolios. For the alternative age factor, we construct portfolios of stocks in the bottom 30% (L), middle 40% (M), and top 30% (H) of the $\mathrm{Retiree}_{j,t}$ distribution. For the income-to-wealth factor, we repeat the procedure but in the reverse order in order to maintain a parallel with the baseline wealth factor. That is, we refer to H as the portfolio of stocks in the *bottom* 30% of $\mathrm{IW}_{j,t}$, M as the portfolio of stocks in the middle 40% (M), and

L as the portfolio in the top 30%. For each characteristic, the investor factor is defined as the portfolio that is long H and short L.

In Table IA.4, we report pricing results obtained with these factors. The retirement and income-to-wealth factors both have significant expected returns, alphas, and betas. The retirement factor produces almost the same pricing results as the baseline age factor used in the main text. The income-to-wealth factor also delivers similar results as the baseline wealth factor.

III.D. Spanning Regressions

We next evaluate the pricing performance of the three alternative three-factor models presented above by running the spanning regressions described in Section 4.2 of the main text. Specifically, we ask whether the alternative three-factor models span the investment, profitability, and momentum factors that generate significant CAPM alpha in 1997-2018.

Table IA.5, Panel A, reports the results of the spanning regressions for the market factor and the baseline age and wealth factors. The results are the same as those reported in Table 3, Panel B, in the main text. They serve as a useful benchmark for the current analysis.

In Panel B of Table IA.5, we use the age and wealth factors constructed directly from investor direct stockholdings, as explained in Section III.A of this Appendix. The results are similar to those in Panel A but slightly less powerful. Other than momentum, the firm profitability and investment factors are spanned by our alternative three-factor model.

In Panel C of Table IA.5, we use the age and wealth factors constructed from the subsamples that contains two-thirds of investors, as explained in Section III.B of this Appendix. The results are again similar to those in Panels A and B but not as statistically powerful due to the smaller investor subset.

In Panel D of Table IA.5, we use the retirement and income-to-wealth factors described in Section III.C of this Appendix. The results are in line with those in the earlier panels and robust. Indeed, firm factor alphas are all statistically insignificant.

In Panel E of Table IA.5, we follow the procedure of Gibbons, Ross, and Shanken (1989) and test the null hypothesis that the intercepts of all three firm factors are jointly equal to zero. We run this test for every benchmark investor-factor model presented in the earlier panels. In most cases, we fail to reject the null hypothesis that pricing errors of firm factors are jointly zero at the 5% significance level. The p-value of these tests is highest for the baseline specification of the age and wealth factors.

In summary, this robustness analysis confirms that our investor factors are robust to alternative specifications. It also confirms that the baseline specification performs best in pricing the cross-section of stock returns.

IV. Appendix to Section 5: The Cross-Section of Household Portfolio Tilts

IV.A. Calculation of Labor Income Risk

We calculate income risk by adopting the methodology of Guvenen et al. (2017). An investor's income beta is defined as the sensitivity of her labor income growth to real GDP growth. We retrieve the real GDP series from Statistics Norway.[6]

We construct the sample of workers as follows. For a given year t, we include all individuals who meet two conditions: (i) a minimum income of NOK 5,000, and (ii) a reported income

[6]This measure refers to Mainland Norway and excludes general government.

above NOK 5,000 in at least three out of the five previous years. We distinguish between retirees and individuals who are between 26 and 65 years old and are actively working.

We define an individual's real earnings, $Y_{i,t}$, as the ratio of her nominal earnings to the consumer price index, and we denote log real earnings by $y_{i,t} = \ln(Y_{i,t})$. The first difference, $\Delta\, y_{i,t}$, is our measure of real earnings growth. We estimate an individual's permanent income by the average of real earnings $Y_{i,t}$ over the 5-year period between years $t-6$ to $t-2$. This choice of dates ensures that there is no overlap between the period over which earnings growth is computed (years $t-1$ and t) and the period over which average earnings are computed (years $t-6$ to $t-2$). As a result of this gap, we are able to rule out any mechanical correlation between earnings growth in year t and historical average earnings.

We calculate an individual's income risk in several steps. First, we classify individuals into $g = 1, 2, 3, ..., G-1$ groups based on permanent income (12 groups) and employment sector (20 groups). We also include retirement as a separate group, which produces 241 groups in total. Consistent with the main text, the 12 permanent-income groups include the first 9 deciles of the permanent income distribution (groups 1-9), the 90^{th}-99^{th} percentiles (group 10), the 99^{th}-99.9^{th} percentiles (group 11), and the top 0.1% (group 12). The sector categorization is explained in detail below. We filter out groups with at least 1000 individuals and end up with 220 groups.

For each group, 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:

$$\Delta y_{i,t} = a_g + \beta_g^{GDP} \Delta GDP_t + \varepsilon_{i,t}. \tag{IA-29}$$

The regression yields a slope coefficient β_g^{GDP} for each group. We assign this coefficient to all individuals in the group and use it as a proxy for their exposure to macroeconomic risk.

Figure IA.1 displays the distribution of income betas for four permanent income groups: the 40th-50th percentiles, the 60th-70th percentiles, the 80th-90th percentiles, and the top 0.1%s. For each group, the distribution in income betas is economically significant and ranges from an income beta of -0.5 (retirees) to more than unity (household services, finance, petroleum, technical services). We note that income betas also differ across the permanent income distribution. For example, the income beta of individuals working in finance exceeds 2 for high-income individuals but is below 1 for the median-income individual.

We assign an industry to each individual based on SIC codes. In the following description, we explain the industry classification scheme that is used by Statistics Norway and then describe the level of aggregation we use.

Our industry classification is based on the Norwegian Standard Industrial Classification (SIC2007), which has been used in Norway's official statistics since 2008. The first four digits of SIC2007 codes are identical to the EU's industrial classification, NACE Rev. 2.[7] As in other countries, SIC2007 codes are modified to incorporate a fifth digit that reflects local industrial conditions. Before 2008, the industry classification was based on NACE Rev. 1.1 (the 2002 update of NACE). We convert old classifications to new ones by using a linking table available on Statistics Norway's website, which ensures that our codes are consistent over time.

Our aggregation scheme involves two main steps. First, we follow the recommended standards in the National Accounts' second revision (A64 Rev. 2). However, in some cases Statistics Norway split some industries into two or more sub-industries. In these cases, we follow Statistics Norway. Overall, our procedure generates 70 industries, compared with 66 in A64 Rev. 2. We then aggregate the resulting industries into 20 main industries on

[7] The description ias available at https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

the basis of our knowledge about the Norwegian economy and existing aggregations. An overview over the disaggregated and aggregated industries applied in our study is provided in Table IA.6.

REFERENCES

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

Cohen, Randolph B., and Christopher Polk, 1998, An investigation of the impact of industry factors in asset-pricing tests, Working paper, Harvard Business School.

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *Journal of Finance* 58, 609–641.

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

Gibbons, Michael R, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–52.

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.

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

Oedegaard, Bernt Arne, 2020, Empirics of the Oslo Stock Exchange: Basic descriptive results, 1980-2019, Working paper, University of Stavanger.

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

Table IA.1
Return Performance of Firm Factors

This table reports statistics of the return performance of five firm factors estimated from the universe of Norwegian stocks in 1997-2018. We report the monthly value-weighted average excess returns (percentage rate), standard deviations, Sharpe Ratios, and CAPM-alphas (Alpha) together with t-statistics of the CAPM-alphas. The market factor is the value-weighted portfolio of all the stocks in the universe. The size factor SMB is based on market value at the end of June each year. It takes a long position in the companies that are above the median in the size distribution, and a short position in the companies that are below the median. All other factors are defined as the return differential between 70th and 30th percentile portfolios of stocks sorted on different characteristics. The momentum factor MOM is based on the geometric return over the previous 12 months in which the most recent month is left out. The size factor SMB is based on market value at the end of June each year. The value factor as HML is based on industry adjusted book-to-market ratio at the end of December in the previous year. The profitability factor RMW is based on the ratio gross profits to total assets. The investment factor CMA is based on growth in total assets.

		Factor					
	Period	MKT	SMB	HML	RMW	CMA	MOM
Mean	1997-2018	0.56	-0.13	0.09	0.85	0.54	0.94
Stddv	1997-2018	5.95	4.07	5.10	5.86	5.42	7.44
SR	1997-2018	0.32	-0.11	0.06	0.50	0.34	0.44
Alpha	1997-2018		-0.02	0.16	0.96	0.66	1.05
t(Alpha)	1997-2018		-0.09	0.51	2.70	2.01	2.32

Table IA.2
Age and Wealth Factors from Investor Portfolio Holdings

This table reports statistics of the return performance of alternative age and wealth investor factors estimated from the universe of Norwegian stocks in 1997-2018. These alternative factors are constructed from the direct stock holdings of household investors. For the age factor, we split investors into three groups according to their age each year: investors below 20^{th} percentile (L), investors between the 20^{th} and 80^{th} percentiles (M), and investors above 80^{th} percentiles (H). For the wealth factor, we split investors into three groups according to their wealth: investors below the 50^{th} percentile (L), investors between the 50^{th} and 90^{th} percentiles (M), and investors above the 90^{th} percentile (H). Each factor is defined as H minus L. Panel A reports monthly value-weighted average excess returns for the low-, medium-, and high- portfolios for each investor characteristic. 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	0.40	0.61	0.60	0.20	1.01	1.54	1.78	1.76
Wealth	0.47	0.56	0.64	0.17	1.13	1.53	1.84	1.51
Panel B: Monthly CAPM Alphas								
	Alpha				t(Alpha)			
	L	M	H	H-L	L	M	H	H-L
Age	-0.16	0.06	0.11	0.26	-1.03	0.36	1.22	2.38
Wealth	-0.11	0.03	0.15	0.25	-0.66	0.27	1.20	2.59
Panel C: Monthly CAPM Betas								
	Beta				t(Beta)			
	L	M	H	H-L	L	M	H	H-L
Age	1.00	1.00	0.89	-0.10	39.48	38.24	61.17	-5.69
Wealth	1.04	0.96	0.89	-0.15	37.24	58.03	43.72	-9.24

Table IA.3
Age and Wealth Factors From a Smaller Investor Subset

This table reports statistics of the return performance of the age and wealth investor factors estimated from a smaller subset consisting of two thirds of Norwegian investors. The statistics are based on 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	0.23	1.01	0.98	0.76	0.46	2.29	2.69	2.15
Wealth	0.30	1.04	0.83	0.53	0.62	2.78	1.92	1.64
Panel B: Monthly CAPM Alphas								
	Alpha				t(Alpha)			
	L	M	H	H-L	L	M	H	H-L
Age	-0.66	0.12	0.18	0.84	-2.22	0.60	1.33	2.38
Wealth	-0.60	0.20	-0.01	0.60	-2.20	2.16	-0.02	1.85
Panel C: Monthly CAPM Betas								
	Beta				t(Beta)			
	L	M	H	H-L	L	M	H	H-L
Age	1.07	1.09	0.93	-0.14	21.57	33.98	40.85	-2.35
Wealth	1.11	1.00	0.99	-0.12	24.17	63.65	24.65	-2.27

Table IA.4
Retirement and Income-to-Wealth Factors

This table reports statistics of the return performance of retirement and income-to-wealth investor factors estimated 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. The retirement portfolios correspond to the bottom 30% (L), mid 40% (M), and top 30% (H) of stocks sorted by their retirement characteristic. The income-to-wealth portfolios correspond to the top 30% (L), mid 40% (M), and bottom 30% (H) of stocks sorted by their income-to-wealth characteristic. For each 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
Retirement	0.17	0.87	1.05	0.88	0.31	2.03	2.86	2.26
Income-to-wealth	0.58	1.01	1.27	0.68	1.27	2.58	3.10	2.05
Panel B: Monthly CAPM Alphas								
	Alpha				t(Alpha)			
	L	M	H	H-L	L	M	H	H-L
Retirement	-0.74	0.00	0.23	0.97	-2.20	-0.02	1.99	2.50
Income-to-wealth	-0.32	0.16	0.46	0.77	-1.42	1.21	2.04	2.34
Panel C: Monthly CAPM Betas								
	Beta				t(Beta)			
	L	M	H	H-L	L	M	H	H-L
Retirement	1.12	1.05	0.96	-0.16	19.79	32.80	49.19	-2.40
Income-to-wealth	1.10	1.01	0.94	-0.16	29.53	45.21	24.95	-2.88

Table IA.5
Spanning Regressions for Alternative Sets of Investor Factors

This table reports OLS spanning regressions of firm factors on the market factor and alternative specifications of the age and wealth factors in 1997-2018. The set of firm factors is limited to the factors that have statistically significant alphas in Panel A of Table III in the main text. Panel A reports tests for the baseline age and wealth factors and is equivalent to Table III, Panel B of the main text. Panel B reports tests for age and wealth factors constructed directly from the stock holdings of individual investors. Panel C reports tests for the age and wealth factors constructed from a random subset of 2/3 of investors. Panel D reports tests for factors based on investors' retirement status and income-to-wealth ratio. We report the intercept (alpha) and slope coefficients of the regressions along with their respective t-values and the regression R^2. For each of these tests, Panel E reports p-values for the Gibbons et al. (1989) test that all intercepts are equal to zero.

Panel A: Baseline Specification of Age and Wealth Factors

LHS	alpha	t(alpha)	MKT	t(MKT)	Age	t(Age)	Wealth	t(Wealth)	R^2
Factor:									
MOM	0.45	1.05	-0.11	-1.51	0.16	2.55	0.44	5.84	0.17
RMW	0.39	1.28	-0.11	-2.12	0.45	9.79	0.10	1.77	0.31
CMA	0.54	1.63	-0.20	-3.55	0.10	2.07	0.01	0.14	0.06

Panel B: Age and Wealth Constructed from Investor Stockholdings

LHS	alpha	t(alpha)	MKT	t(MKT)	Age	t(Age)	Wealth	t(Wealth)	R^2
Factor:									
MOM	0.90	1.98	-0.16	-1.85	0.82	2.27	-0.24	-0.60	0.04
RMW	0.55	1.73	-0.03	-0.56	1.42	5.63	0.13	0.47	0.24
CMA	0.47	1.46	-0.14	-2.26	0.65	2.54	0.06	0.23	0.10

Panel C: Age and Wealth Factors Based on 2/3 Investor Subset

LHS	alpha	t(alpha)	MKT	t(MKT)	Age	t(Age)	Wealth	t(Wealth)	R^2
Factor:									
MOM	0.89	1.99	-0.18	-2.32	-0.06	-0.70	0.35	3.97	0.07
RMW	0.60	1.83	-0.14	-2.60	0.42	7.14	0.01	0.18	0.20
CMA	0.63	1.90	-0.21	-3.85	0.09	1.49	-0.07	-1.12	0.06

Panel D: Investor Factors Are Retirement Status and Income-to-Wealth Ratio

LHS	alpha	t(alpha)	MKT	t(MKT)	Retiree	t(Retiree)	I/W	$t(I/W)$	R^2
Factor:									
MOM	0.80	1.75	-0.17	-2.16	0.19	2.72	0.08	0.95	0.05
RMW	0.54	1.68	-0.14	-2.52	0.44	8.94	-0.01	-0.16	0.26
CMA	0.36	1.11	-0.16	-2.96	0.09	1.90	0.27	4.66	0.13

(Continued)

Table IA.5 - *Continued*

Panel E: GRS Tests of Firm Factor Regressions on the Age, Wealth, and Market Factors			
LHS Factors	RHS Factors	GRS Statistic	p-value
RMW, CMA, MOM	MKT, Age, Wealth	1.49	0.21
RMW, CMA, MOM	MKT, Age, & Wealth from Stockholdings	2.41	0.07
RMW, CMA, MOM	MKT, Age & Wealth from 2/3 Subset	2.67	0.05
RMW, CMA, MOM	MKT, Retirement, I/W	1.84	0.14

Table IA.6
Industry Classification

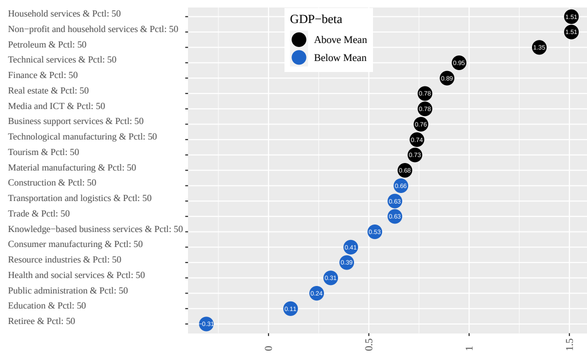
This table reports our classification of industries based on the A64 Rev. 2 and Nace codes.

Our industries	A64 Rev. 2	Nace
Resource industries	Crop and animal production, hunting and related service activities	1
Resource industries	Forestry and logging	2
Resource industries	Fishing	3.3
Resource industries	Aquaculture	3.2
Resource industries	Mining and quarrying	5, 7-8, 9.9
Resource industries	Electricity, gas, steam and air conditioning supply	35
Petroleum	Oil and gas extraction	6
Petroleum	Transport via pipelines	49.5
Petroleum	Service activities incidental to oil and gas	9.1
Consumer manufacturing	Manufacture of food products, beverages and tobacco products	10-12
Consumer manufacturing	Manufacture of textiles, wearing apparel and leather products	14-15
Consumer manufacturing	Printing and reproduction of recorded media	18
Consumer manufacturing	Manufacture of furniture	31-32
Material manufacturing	Manufacture of wood and wood products, except furniture	16
Material manufacturing	Manufacture of paper and paper products	17
Material manufacturing	Refined petroleum, chemical and pharmaceutical products	19-21⋆⋆
Material manufacturing	Chemical commodities	20.11-20.15
Material manufacturing	Manufacture of rubber and plastic products	22
Material manufacturing	Manufacture of other non-metallic mineral products	23
Material manufacturing	Manufacture of basic metals	24
Material manufacturing	Fabricated metal products, except machinery and equipment	25
Technological manufacturing	Manufacture of computer, electronic and optical products	26
Technological manufacturing	Manufacture of electrical equipment	27
Technological manufacturing	Manufacture of machinery and equipment n.e.c.	28
Technological manufacturing	Manufacture of motor vehicles, trailers and semi-trailers	29
Technological manufacturing	Building of ships, oil platforms and modules	30
Technological manufacturing	Repair and installation of machinery and equipment	33
Public administration	Water collection, treatment and supply	36
Public administration	Sewerage	37-39
Public administration	Public administration central/local government	84
Public administration	Defence	84.2
Construction	Construction	41-43
Trade	Wholesale and retail trade and repair of motor vehicles	45
Trade	Wholesale trade, except of motor vehicles	46
Trade	Retail trade, except of motor vehicles	47
Transportation and logistics	Land transport, except transport via pipelines	49.1-49.4
Transportation and logistics	Inland water transport and supply	50.1-50.4⋆
Transportation and logistics	Air transport	51
Transportation and logistics	Warehousing and support activities for transportation	52
Transportation and logistics	Postal and courier activities	53
Tourism	Ocean transport	50.101, 50.201
Tourism	Accommodation and food service activities	55-56
Tourism	Travel agency and tour operator reservation service	79

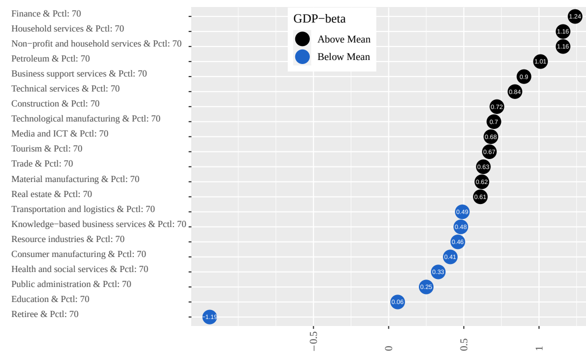
(Continued)

Table IA.6 - *Continued*

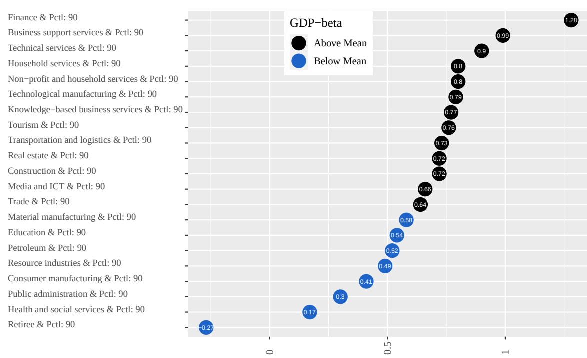
Our industries	A64 Rev. 2	Nace
Media and ICT	Publishing activities	58
Media and ICT	Motion picture and video program production, broadcasting	59-60
Media and ICT	Telecommunications	61
Media and ICT	Computer programming and related activities	62-63
Finance	Insurance, except compulsory social security	65
Finance	Activities auxiliary to financial services and insurance activities	66
Finance & consulting	Financial service activities (not insur.)	64
Finance& consulting	Other prof., scientific activities	74-75
Finance & consulting	Real estate activities	68
Knowledge-based business services	Legal and accounting activities	69-70
Knowledge-based business services	Scientific research and development	72
Knowledge-based business services	Advertising and market research	73
Technical services	Architectural and engineering consultancy activities	71
Business support services	Rental and leasing activities	77
Business support services	Employment activities	78
Business support services	Security and investigation activities	80-82
Education	Education	85
Health and social services	Human health activities	86
Health and social services	Social work activities	87-88
Non-profit and household services	Creative, arts and entertainment activities	90-92
Non-profit and household services	Sports activities and amusement and recreation activities	93
Non-profit and household services	Activities of membership organisations	94,99
Household services	Repair of computers and personal and household goods	95
Household services	Other personal service activities	96
Household services	Activities of households as employers	97



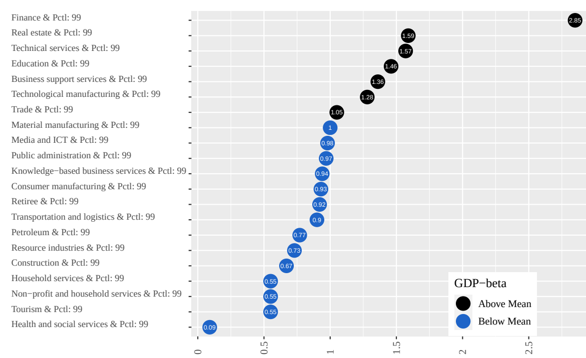
(a) Income Beta ($40^{th} - 50^{th}$) percentiles



(b) Income Beta ($60^{th} - 70^{th}$ percentiles)



(c) Income Beta ($80^{th} - 90^{th}$ percentiles)



(d) Income Beta ($99^{th} - 99.9^{th}$ percentiles)

Figure IA.1 Dispersion of income beta This figure plots income betas across sectors for four different permanent income groups of investors: the $40^{th} - 50^{th}$ percentiles (Panel A), the $60^{th} - 70^{th}$ percentiles (Panel B), the $80^{th} - 90^{th}$ percentiles (Panel C), and the $99^{th} - 99.9^{th}$ percentiles (Panel D). The income betas are 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 estimation is run on a panel of Norwegian individual investors in 1997-2018.