Can Downside-Risk Measures Help to Explain the Reluctance of Households to Invest in XTFs? An Empirical Study Using the SHS-base

Andreas Oehlera and H. Philipp Wangerb

Abstract
XTFs are plain-vanilla Exchange Traded Funds (ETFs) which replicate a broad, internationally diversified market index. Academics usually recommend households to employ XTFs. However, this recommendation seems to be widely ignored by households. In order to find a possible explanation for this, we investigate whether evaluating risk according to downside-risk measures, which reflect households’ interpretation of risk more closely than the standard deviation of returns, yields less performance enhancements and thus fewer incentives to invest in XTFs. Since households widely show similar portfolio compositions, we apply stylized household portfolios with different asset class concentrations and combine them with security holding data which is representative for German households. In that regard, 47,388 securities are covered. To the best of our knowledge, we are the first to investigate performance enhancements induced by XTFs according to downside-risk measures while considering multiple relevant asset classes of household portfolios. Our results show that, first, none of the applied downside-risk measures leads to consistently lower enhancements compared to when the standard deviation is applied. Thus, we conclude that none of the downside-risk measures can help to explain the reluctance of households to invest in XTFs. Second, the results show that all stylized household portfolios can enhance portfolio performance from employing XTFs, regardless of the underlying risk measure. This substantiates the recommendation of academics to employ XTFs and shows that the advice holds true for households that evaluate risk according to downside-risk measures while taking into account multiple relevant asset classes.

JEL Classification: D14, D81, D91, G11, G41
Key Words: XTF, Household Finance, Downside-risk measures, Multiple Asset Classes, Portfolio Performance Enhancements

a Full Professor and Chair of Finance, Bamberg University, Kaerntenstraße 7, 96052 Bamberg, Germany, Phone: (+49) 951-863-2537, Fax: (+49) 951-863-2538, e-mail: andreas.oehler@uni-bamberg.de.
b Department of Finance, Bamberg University, Kaerntenstraße 7, 96052 Bamberg, Germany, Phone: (+49) 951-863-2537, Fax: (+49) 951-863-2538, e-mail: hans.wanger@uni-bamberg.de, corresponding author.

This paper uses data from the Deutsche Bundesbank’s Securities Holdings Statistics-Base. The results published and the related observations and analysis may not correspond to results or analysis of the data producers. We would like to thank Deutsche Bundesbank, especially Martin Eisele, for providing the dataset of the Securities Holdings Statistics-Base. In addition, we would like to thank seminar participants at Bamberg University in Bamberg, Germany for helpful comments and suggestions. All remaining errors are our own.
Can Downside-Risk Measures Help to Explain the Reluctance of Households to Invest in XTFs? An Empirical Study Using the SHS-base

Abstract
XTFs are plain-vanilla Exchange Traded Funds (ETFs) which replicate a broad, internationally diversified market index. Academics usually recommend households to employ XTFs. However, this recommendation seems to be widely ignored by households. In order to find a possible explanation for this, we investigate whether evaluating risk according to downside-risk measures, which reflect households’ interpretation of risk more closely than the standard deviation of returns, yields less performance enhancements and thus fewer incentives to invest in XTFs. Since households widely show similar portfolio compositions, we apply stylized household portfolios with different asset class concentrations and combine them with security holding data which is representative for German households. In that regard, 47,388 securities are covered. To the best of our knowledge, we are the first to investigate performance enhancements induced by XTFs according to downside-risk measures while considering multiple relevant asset classes of household portfolios. Our results show that, first, none of the applied downside-risk measures lead to consistently lower enhancements compared to when the standard deviation is applied. Thus, we conclude that none of the downside-risk measures can help to explain the reluctance of households to invest in XTFs. Second, the results show that all stylized household portfolios can enhance portfolio performance from employing XTFs, regardless of the underlying risk measure. This substantiates the recommendation of academics to employ XTFs and shows that the advice holds true for households that evaluate risk according to downside-risk measures while taking into account multiple relevant asset classes.

JEL Classification: D14, D81, D91, G11, G41

Key Words: XTF, Household Finance, Downside-risk measures, Multiple Asset Classes, Portfolio Performance Enhancements
1 Introduction

Households fail to outperform a respective market index on average (see e.g. Barber and Odean, 2000; 2001; French, 2008; von Gaudecker, 2015). A common advice of academics thus is to employ well-diversified, low-cost passive index funds or Exchange Traded Funds (ETFs) (see e.g. Malkiel, 2003; French, 2008; Huang and Lin, 2011; Jacobs et al., 2014; Bhattacharya et al., 2017; Elton et al., 2019). However, although assets under management of ETFs largely increased in the last decade (see e.g. Investment Company Institute, 2018; Deutsche Bundesbank, 2018), individuals who invest in ETFs on average do not improve portfolio performance due to poor ETF selection and timing (see Clifford et al., 2014; Bhattacharya et al., 2017). In this context, separating ETFs that are appropriate to follow the common advice of academics from the extensive ETF universe seems necessary. Therefore, Oehler and Wanger (2020) define XTFs as plain-vanilla ETFs which pursue a passive investment approach and replicate a broad, internationally diversified market index.

Studies which outline potential enhancements from XTFs mostly rely on a mean-variance (MV) framework in a neoclassical manner (see Markowitz, 1952). In a MV-framework, risk is measured by return variance, or standard deviation (SD), which assumes, among others, that returns are normally distributed and that households weight positive and negative deviations around the mean equally.

However, in reality returns are hardly normally distributed (see e.g. Mandelbrot, 1963; French et al., 1987). Moreover, Kahneman and Tversky’s (1979) Prospect Theory suggests that households actually distinct between gains and losses relative to an individual reference point or target, and show a higher sensitivity towards losses than towards gains. This phenomenon is called loss-aversion (see e.g. Thaler, 1980; Benartzi and Thaler, 1995). Correspondingly, Unser (2000) and Veld and Veld-Merkoulova (2008) find that households perceive risk rather in terms of downside-risk than in terms of variance and apply several versions of the general form of Lower-Partial-Moments (LPMs) to evaluate risk.
The SD is a special case of the LPM (see e.g. Harlow and Rao, 1989; Harlow, 1991). Risk evaluations according to both measures, however, differ from each other if returns are non-normally distributed, i.e. skewed, and if the reference point which is incorporated in the LPM varies from the mean return used in the SD (see e.g. Harlow and Rao, 1989; Harlow, 1991; Marmer and Ng, 1993; Grootveld and Hallerbach, 1999; Jarrow and Zhao, 2006). If SD does not adequately reflect a household’s interpretation of risk, a MV-framework might ascertain the risk/return-enhancements from XTFs insufficiently. The goal of this analysis is to assess risk/return-enhancements from XTFs while employing assumptions that approximate households’ actual risk evaluation more accurately than implied by the MV-framework. If risk/return-frameworks that are based on downside-risk measures yield fewer risk/return-enhancements, XTFs might appear less attractive, consequently motivating households to refuse making a respective investment. This raises the question: Can downside-risk measures help to explain the reluctance of households to invest in XTFs?

To answer this question, we compare possible risk/return-enhancements when risk is measured by SD, to risk/return-enhancements when households show loss-aversion and evaluate risk in terms of downside-risk measures. First, we estimate risk/return-profiles of households’ portfolios. Most households rely on the recommendation of financial advisors (see e.g. DAB Bank, 2004; Hackethal et al., 2011), however, financial advisors build very similar portfolios instead of individualizing portfolio compositions according to households’ characteristics (see Foerster et al., 2017). Thus, we rely on stylized portfolio compositions of households, so called Household Portfolio Types (HPTs) (see Oehler and Wanger, 2020). HPTs include multiple relevant asset classes and are clusters of the relative share of the total portfolio value invested into a certain asset class, i.e. the asset class weight. Considering multiple asset classes is crucial since the allocation of a portfolio’s assets among the asset classes mainly determines overall
portfolio performance (see Brinson et al., 1986; Ibbotson and Kaplan, 2000).\(^1\) We then use a random sampling process that selects securities from the Securities Holdings Statistics (SHS)-base and assign them to HPTs. In this way, we built 1,000 portfolios per HPT. Our security sample of the SHS-base of the German central bank (Deutsche Bundesbank) includes 47,388 individual securities.

Second, we select downside-risk measures that are more likely to reflect households’ actual perception of risk than SD. The downside-risk measures are Probability of Shortfall, i.e. LPM Zero (LPM0), Expected Value of Shortfall, i.e. LPM One (LPM1), Semi-standard deviation, i.e. LPM Two ($\sqrt{\text{LPM}^2}$), and Maximum Drawdown (MDD). Third, for the investigation of risk/return-enhancements from XTFs, we establish benchmark portfolios which consist of one stock XTF, one bond XTF, and a “safe”\(^2\) investment exhibiting a fixed interest rate. We build a separate benchmark portfolio for each HPT by adjusting the asset class weights of the benchmark portfolio according to the asset class concentration of the HPT. Since the XTFs employed in the benchmark portfolios replicate market returns and market return represents the most important reference for households from which they measure downside-risk (see Veld and Veld-Merkoulova, 2008), we employ the returns of the benchmark portfolio as target returns.

Fourth, risk/return-enhancements from XTFs are determined by risk-adjusting the benchmark portfolios to the risk of each household portfolio and measuring the Return Difference (RD) (see Oehler and Wanger, 2020; based on Calvet et al., 2007).

Our contribution to the literature is twofold. First, to the best of our knowledge, we are the first to investigate risk/return-enhancements from XTFs according to several downside-risk

\(^{1}\) The authors refer to this as “investment policy”. In this context, the term “asset allocation” is commonly employed as well. However, this term is not unambiguously defined and used in the literature and the investment practice. To avoid misconceptions, we rely on “asset class weights” in the following which denotes the relative share of the total portfolio value invested into certain asset classes.

\(^{2}\) The quotes around “safe” indicate that the respective investments are only relatively safe compared to the remaining assets and not riskless in a neoclassical sense in which default is not considered. For better readability, we leave out the quotes in the following.
measures while considering multiple relevant asset classes of households’ portfolios. Second, we use representative security holding data of households. The results indicate, first, that the risk/return-enhancements determined in a MV-framework are significantly different from those determined in the downside-risk/return-frameworks. However, none of the downside-risk/return-frameworks yields consistently lower risk/return-enhancements compared to the MV-framework across all stylized portfolios. We conclude that none of the employed downside-risk measures can help to explain the reluctance of households to invest in XTFs. Second, all risk/return-frameworks indicate that households can enhance their portfolio performance by employing XTFs, (see e.g. Malkiel, 2003; French, 2008; Huang and Lin, 2011; Jacobs et al., 2014; Bhattacharya et al., 2017) regardless of whether risk is evaluated according to SD or the LPM-based downside-risk measures. This supports the recommendation of academics to employ XTFs and indicates that the advice holds true when households evaluate risk according to downside-risk measures and consider multiple relevant asset classes.

2 Related Literature

Portfolios of households frequently show their preference for small numbers of individual securities (see e.g. Polkovnichenko, 2005; Goetzmann and Kumar, 2008) and domestic assets (see e.g. French and Poterba, 1991; Oehler et al., 2008; Baltzer et al., 2015), which can lead to suboptimal levels of portfolio diversification and increase portfolio risk. To improve portfolio performance, academics typically recommend to employ XTFs (see e.g. Malkiel, 2003; French, 2008; Huang and Lin, 2011; Jacobs et al., 2014; Bhattacharya et al., 2017). Most studies thereby measure risk in terms of the return variance, or SD. SD assumes, among other things, normally distributed returns and does not involve the skewness of returns. However, normally distributed returns are hardly a tenable assumption (see e.g. Mandelbrot, 1963; Fama, 1965; Christie, 1982; French et al., 1987). Since securities like stocks and bonds can achieve unlimited gains, but
losses are limited to the invested amount of capital, the asymmetrical nature of returns is straightforward (see e.g. Grootveld and Hallerbach, 1999). Various studies document the preference of individuals for asymmetric return distributions with positive skewness (see Arditti, 1967; Barberis and Huang, 2008; Kumar, 2009). In order to raise the chance to obtain positively skewed returns, it can be reasonable from an investor’s perspective to abstain from expanding portfolio diversification (see Simkowitz and Beedles, 1978; Conine and Tamarkin, 1981; Hueng and Yau, 2006; Mitton and Vorkink, 2007; Kim, 2015). Consequently, the preference for skewness can motivate households to refuse to invest in diversified XTFs.

As opposed to SD which symmetrically weights return variations as risk, Prospect Theory by Kahneman and Tversky (1979) provides evidence that investors actually distinct between gains and losses relative to a specific reference point, or target. Investors assign approximately twice the weight to losses than they assign to gains which represents their loss-aversion (see Kahneman et al., 1990, Tversky and Kahneman, 1992; Benartzi and Thaler, 1995). Increased levels of loss-aversion lower the probability of an investor participating in risky assets (see Dimmock and Kouwenberg, 2010). In this regard, further studies argue that downside-risk, i.e. the returns which range below a specific target of the investor, approximates households’ interpretation of risk more closely (see e.g. De Bondt, 1998; Bertsimas et al., 2004).

Considering below-target-returns, or downside-risk, as an undesirable event also seems to be a more intuitive interpretation of risk compared to the symmetrical understanding of risk implied by SD.

Measures of downside-risk correspondingly take return fluctuations below a specific target into account. Some of these measures seem particularly adequate in this context. Unser (2000) reports that the LPM0 best represents individual investors’ risk perception. LPM0 describes the probability for the occurrence of a below-target-return. This measure is closely related to the Safety First principle by Roy (1952) as the disaster level described therein corresponds to LPM0’s target return.
Veld and Veld-Merkoulova (2008) find that investors implicitly employ more than one risk measure and that investors’ risk perception directly influences their selection of risky investments. In a downside-risk framework, the previously applied variance and SD are equivalent to Semi-variance and the square root of the Semi-variance, i.e. \( \sqrt{LPM^2} \). Semi-variance is preferred by stock-investors while bond-investors focus on LPM0. Investors who care more about the underperformance relative to the market return than about the initial investment rather employ LPM1. LPM1 is composed by the probability of a below-the-target return multiplied with the extent of the deviation from the target. Veld and Veld-Merkoulova (2008) outline, however, that semi-variance most often represents investors’ risk perception. \( \sqrt{LPM^2} \) was, like the mean-variance criterion in security selection (see Markowitz, 1952), proposed by Markowitz (1959). Both SD and \( \sqrt{LPM^2} \) assume the same degree of risk aversion which allows comparing them to each other. SD can be interpreted as a special case of \( \sqrt{LPM^2} \) (see e.g. Harlow and Rao, 1989; Harlow, 1991). Using \( \sqrt{LPM^2} \) would be equivalent to using SD if returns were symmetrically distributed and the target return for \( \sqrt{LPM^2} \) was equivalent to the mean return of a portfolio (see Harlow, 1991). In turn, both measures differ if returns are asymmetrically distributed (see also Jarrow and Zhao, 2006) or the choice of a household’s target return does not coincide with the portfolio mean (see e.g. Ang and Chua, 1979).

The downside-risk measures above can be subsumed under the general notation of the LPMs (see Bawa, 1975; Fishburn, 1977). This notation enables specifying the LPM0, LPM1 and \( \sqrt{LPM^2} \). Both Unser (2000) and Veld and Veld-Merkoulova (2008) conclude that downside-risk measures better reflect investors’ perceived risk than the variance which is far less commonly employed by investors. Consequently, we control for LPM0, LPM1 and \( \sqrt{LPM^2} \) as downside-risk measures in our analysis.

Another risk measure which is commonly used in investment practice and evaluates risk in terms of downside-risk is the MDD. MDD is defined as the performance loss, measured from
a previous peak or reference point (see Bradford and Siliski, 2016). MDD measures risk in terms of returns which makes it easy to understand for households. Additionally, MDD is frequently offered to investors by financial service providers. It is thus likely that households also employ MDD in their investment decisions. Furthermore, MDD is well-suited for risk evaluations against a benchmark (see e.g. Bradford and Siliski, 2016). As opposed to the LPM-based downside-risk measures, MDD is not based on the LPM notation and can consist of periods other than one month as implied by the remaining risk measures. Risk/return-enhancements according to MDD thus cannot be compared directly with those of LPM-based downside-risk measures. Nevertheless, MDD allows us to control for a different and relevant downside-risk measure and serves as robustness regarding possible enhancements from XTFs.

Documented target returns from which households measure downside-risk are the initial price of an investment, the risk-free rate of return or the market return (see Unser, 2000; Veld and Veld-Merkoulova, 2008). However, Veld and Veld-Merkoulova (2008) find that the market return represents the most important benchmark for households.3 XTFs replicate market returns and are employed in the benchmark portfolios. Moreover, the benchmark portfolios which employ XTFs represent the reference, or target, from which risk/return-enhancements are determined. Consequently, we apply the returns of the benchmark portfolios as target returns to measure downside-risk.

Existing studies which investigate households’ portfolio performance (see e.g. Barber and Odean, 2000; 2001; Polkovnichenko, 2005; Goetzmann and Kumar, 2008) or compare portfolio performance according to traditional versus downside-risk measures (see e.g. Grootveld and Hallerbach, 1999; Jarrow and Zhao, 2006; Hoffmann and Post, 2017) only include, to the best of our knowledge, a few asset classes, particularly stocks. However, the allocation of investments across asset classes is the most important determinant with regard to overall

---

3 Veld and Veld-Merkoulova (2008) differentiate in their questionnaire between stated and actual choice of the target return. In terms of actual choice, they document market return as the most relevant for households.
portfolio performance (see Brinson et al., 1986; 1995; Ibbotson and Kaplan, 2000). Employing HPTs in our analysis is thus particularly suitable since they include multiple relevant asset classes which involve, besides stocks and bonds, safe financial assets and several mutual fund types. These asset classes have been shown to be particularly relevant in households’ portfolios and, thus, have a major impact on households’ portfolio performance (see e.g. Calvet et al., 2007; von Gaudecker, 2015).

Simultaneously examining all asset classes included in the HPTs and establishing separate benchmark portfolios for each HPT is in line with the multi-layer portfolio framework. According to Oehler (2017, 2015), Oehler et al. (2018a), and Oehler and Horn (2019), households compile portfolios as multi-layer-portfolios. The framework is based on the Behavioral Portfolio Theory (BPT) by Shefrin and Statman (2000) and builds up on mental accounting as introduced in Prospect Theory (see Kahneman and Tversky, 1979; Thaler, 1999). According to the multi-layer portfolio framework, households compile portfolios as layered pyramids. The layers are arranged in decreasing order according to household’s financial needs. Each layer represents a certain financial need which is concretized by several more detailed and specific investment goals (see Oehler, 1995). Households can go on investing in the next upper layer as soon as the financial goals of the previous layer are satisfied. The bottom and middle layer contain basic and additional financial needs, respectively. Investments in the top layer, i.e. the speculation portfolio, pursue the goal of further income and speculation (including the possibility of total loss). All asset classes included in HPTs are asset classes of the speculation portfolio (see e.g. Oehler, 2015; 2017; Oehler et al., 2018a; Oehler and Horn, 2019).

Considering that the investments in the speculation portfolio pursue the same goal, we assume that it is reasonable from the perspective of households to jointly examine the performance of

---

4 Financial products which satisfy basic financial needs are, among others, liability and disability insurance, health care, or liquidity provisions, while additional financial needs can be covered by, among others, accident insurance or retirement provisions (see e.g. Oehler et al., 2018b).
the investments included in the speculation portfolio. Additionally, measuring downside-risk of the speculation portfolio seems important for households given the possibility of total loss for the investments included in this portfolio layer.

Applying this approach, we contribute to existing studies by investigating risk/return-enhancements from XTFs according to several downside-risk measures while considering multiple relevant asset classes of households’ portfolios.

3 Data and Methodology

3.1 Estimation of risk/return-profiles of Household Portfolio Types (HPTs)

To estimate risk and return of households’ portfolios, we rely on stylized portfolio compositions of German households, i.e. HPTs (see Oehler and Wanger, 2020). HPTs are based on a representative household survey, the Panel on Household Finances (PHF), which was conducted in 2014 by the German central bank (Deutsche Bundesbank). HPTs are derived from clustering portfolios by asset class weights, i.e. the relative share of the total portfolio value invested into a certain asset class. HPTs include multiple relevant asset classes. These are the safe financial assets cash (CASH) and savings (SV), as well as the risky financial assets stock funds (SFs), bond funds (BFs), real estate funds (REFs), individual stocks (STs), and individual bonds (BDs). Each HPT exhibits a different asset class concentration.

To estimate risk and return of HPTs, we randomly select German households’ security holdings and assign them to the HPTs’ asset classes. German households’ security holding data is obtained from the SHS-base of the German central bank (Deutsche Bundesbank). The

---

5 This is in line with the tendency of households to jointly consider assets with similar outcome features (see Kahneman and Tversky, 1979; 1984; Tversky and Kahneman, 1981; Thaler, 1985; 1999). This behavioral feature stems from Prospect Theory (see Kahneman and Tversky, 1979) and is referred to as mental accounting (see e.g. Thaler, 1999). Thereafter, households’ propensity to distinct between gains and losses relative to a specific reference point involves that households, instead of a conjunct evaluation of the total portfolio, interpret certain investments separately from the remaining portfolio (see Kahneman and Tversky, 1979; Tversky and Kahneman, 1981; Thaler, 1999). According to the Capital Asset Pricing Model in a MV-framework, in contrast, investors build optimal portfolios by considering all investments jointly and only divide their wealth between a risk-free asset and the market portfolio (see Markowitz, 1952; Sharpe, 1964; Lintner, 1965; Mossin, 1966).
combination of HPTs with securities of the SHS-data fits well since both are representative for the German household sector and stem from the same provider (see Methodological Appendix for details on the preparation of the security data). Our final sample contains 47,388 securities and includes 7,552 SFs, 3,267 BF, 42 REFs, 22,225 BDs and 14,302 STs. For each security, the SHS-base further reveals the aggregated market value of shares owned by German households. Each security’s market value of shares relative to the aggregated market value of all securities included in the respective asset class is assumed as an indication for the distribution of a security among German households and, thus, is used as probability for the selection of a certain security in the random sampling process. This allows us to better approximate German households’ portfolios in terms of their real security holdings.

In each HPT portfolio, an equal number of securities is selected into the applied asset classes. For instance, for a portfolio size (i.e. the number of securities per portfolio) of nine securities, three mutual funds, three bonds and three stocks are obtained. Mutual funds are further divided into one SF, one BF, and one REF. All securities are equally weighted within the portfolio. For monthly interest rates on CASH and SV, German households’ deposits with different agreed maturities were deployed (see Methodological Appendix for details). In this way, we select 1,000 portfolios. To avoid selection bias, each random security selection of a portfolio is applied by all HPTs and weighted according to their respective asset class weights. In this way, we compute 3,000 HPT portfolios.

Within the observation period from January 2014 to December 2016, securities can expire (see Methodological Appendix). For reinvesting the amount of money that is available after the expiration, we involve transaction costs. As fixed transaction costs, we assume 10 Euros per transaction which reflects the approximate amount that large German online brokers charge their clients (see Stiftung Warentest, 2016). Additionally, for each transaction proportional transaction cost of 0.25 percent of the order value are included (see Lynch and Balduzzi, 2000).
An exception is the first month of the observation period, January 2014,\(^6\) in which the random portfolio selection is performed to establish HPTs.

3.2 Downside-risk Measures

The general notation of LPMs was introduced by Bawa (1975) and Fishburn (1977).\(^7\) This notation can be used to specify the applied downside-risk measures. It requires determining the degree of risk aversion \((n)\), i.e. the weight an investor assigns to negative return deviations, as well as the target return \((\tau)\) which is drawn from the benchmark portfolio of a household. The general LPM notation can be described by:

\[
LPM_{ni}(\tau) = \frac{1}{T} \sum_{t=1}^{T} \max[\tau - r_{it}, 0]^n.
\]

Thereby, \(r_{it}\) denotes the (discrete) return of portfolio \(i\) in month \(t\) \((t = 1, \ldots, T)\) and \(T\) the number of months. Depending on the definition of \(n\), the general LPM notation facilitates expressing the LPM0 (with \(n = 0\)), LPM1 (with \(n = 1\)), and \(\sqrt{LPM2}\) (with \(n = 2\)), of Portfolio \(i\) for the target return \(\tau).\(^8\) These three downside-risk measures can be specified by:

\[
\begin{align*}
(2) & \quad LPM_{n=0,i}(\tau) = \frac{1}{T} \sum_{t=1}^{T} \max[\tau - r_{it}, 0]^0 \\
(3) & \quad LPM_{n=1,i}(\tau) = \frac{1}{T} \sum_{t=1}^{T} \max[\tau - r_{it}, 0]^1 \\
(4) & \quad \sqrt{LPM_{n=2,i}(\tau)} = \sqrt{\text{semi\text{-}variance}_i(\tau)} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \max[\tau - r_{it}, 0]^2}.
\end{align*}
\]

\(\sqrt{LPM2}\) is, analogously to the relation between SD and the variance, defined as the square root of the semi-variance. We also employ the MDD in our analysis. MDD can be assembled by

\(^6\) The PHF-survey was conducted from the beginning of the second quarter of 2014. In line with Oehler and Wanger (2020) we assume that households’ responses refer to their financial situation of the previous quarter as of January 2014 and start our analysis from this date.

\(^7\) This general definition is also referred to as \(\alpha - \tau\) model in the literature (see e.g. Fishburn, 1977; Grootveld and Hallerbach, 1999).

\(^8\) Fishburn (1977) states that the model “may have a risk-seeking or ‘gambling’ aspect” (p. 119) when \(n < 0\). \(n = 1\) matches a risk-neutral investor and separates risk-preference from risk-averse behavior, and \(n = 2\) implies risk-aversion (see e.g. Fishburn, 1977; Nawrocki, 1999; Grootveld and Hallerbach, 1999).
first calculating point-in-time drawdowns \( D(T) \), where \( r_{it} \) is the return of portfolio \( i \) in month \( t \) and \( r_{iz} \) is the return of portfolio \( i \) in month \( z \) (see Bradford and Siliski, 2016):

\[
D(T) = \min \left( \frac{\prod_{t=\max_0,\ldots,T-1}^{T} (1+r_{it})}{\max_{t \in 0,\ldots,T-1} \prod_{z=0}^{T} (1+r_{iz})} - 1, 0 \right).
\]

In formula (5), the numerator shows the cumulative portfolio return at time \( T \). The denominator indicates the high-water mark, i.e. the maximum realized cumulative return at any time \( t < T \).

As MDD merely refers to cumulative returns, it is straightforward to calculate. After determining the time series \( D = [D_0, D_1, \ldots, D_T] \) of point-in-time drawdowns, \( MDD(T) \) can be ascertained as the largest point-in-time drawdown (formally expressed by the minimum since drawdowns represent negative returns) determined in \( D \) (see Bradford and Siliski, 2016):

\[
MDD(T) = \min_{t \in 0,\ldots,T} \left[ \min_{D(T) \in D} D(T) = \right.
\]

\[
= \min_{t \in 0,\ldots,T} \left( \min_{\max_{t \in 0,\ldots,T-1} \prod_{z=0}^{T} (1+r_{iz})} \prod_{t=\max_0,\ldots,T-1}^{T} (1+r_{it}) - 1, 0 \right).
\]

As opposed to SD, MDD does not require any assumptions about the distribution of returns. In contrast to the LPM-based downside-risk measures, MDD does not directly include a target return. MDD of the benchmark portfolios, however, can separately be compared to the MDD of HPTs’ portfolios.

Determining downside-risk according to the measures above requires target returns. In our analysis, we use separate target returns for each HPT. As target returns, we employ the mean return of the benchmark portfolios. Thereby, we assume that households apply past returns, relying on the time prior to the observation period in order to establish target returns. The longest pre-observation period with available data for the assets included in the benchmark portfolios is five years, i.e. the period from January 2009 to December 2013. Incorporating the mean benchmark portfolio returns of this period allows computing the downside-risk for the HPTs’ portfolios (1,000 portfolios for each HPT) and the benchmarks portfolios.
3.3 Benchmark Portfolios

We employ benchmark portfolios which include XTFs to determine household portfolio risk/return-enhancements from XTFs. Furthermore, XTFs are used to estimate the target return that is used by the downside-risk measures. Asset class weights of the benchmark portfolios are derived from HPTs’ asset class concentration among stocks (ST and SF), bonds (BD und BF) and safe financial assets. Market returns for the stock XTF and the bond XTF are drawn from ETFs covering the Morgan Stanley Capital International (MSCI) World Index and the Markit iBoxx Euro Sovereign Index, respectively; the fixed interest rate for the safe investment is gathered from the rate on SVs (see Methodological Appendix for details). By choosing SV instead of CASH as benchmark for safe financial assets, we rely on the assumption that households rather prefer investments with longer-term maturities and (marginally) higher fixed interest rates. To avoid semblance of precision and to ensure an easy and memorable benchmark construction for households, more general asset class weights were applied for the benchmark portfolios.

As households cannot obtain information about future target returns, we rely on the period prior to the beginning of the observation period in January 2014. Therefore, we employ monthly return data of the past five years, i.e. from January 2009 to December 2013. We assume this as a reasonable period since highlighting the performance of the past five years to customers is very common among financial data service providers, websites or banks. To compute the target returns for this period, we use the mean returns of the stock XTF, bond XTF, and savings, weighting them according to asset class weights of the derived benchmark portfolios. In the present context of loss-averse households and their substantial amounts invested among safe financial assets, a particular threat to returns might generally arise from inflation. We thus involve respective inflation rates for Germany when assessing portfolio returns.⁹

---

⁹ Monthly inflation rates for Germany were computed by monthly changes of the Harmonised Index of Consumer Prices as of Eurostat (2019). For the period between January 2009 and December 2013, we ascertain an average
3.4 Measurement of Risk/return-Enhancements

The measure of risk/return-enhancements in this analysis needs to be applicable for all included risk/return-frameworks. In addition, the performance measure should yield reasonable and interpretable results from the perspective of households which employ multiple asset classes in their portfolio. We thus rely on a measure that ascertains performance in absolute terms and uses the returns as a common basis to determine performance differences between the HPTs’ portfolios and the benchmark portfolio since this variable is equally established and included in all risk/return-frameworks (in reference to the Return Loss by Calvet et al., 2007; see Methodological Appendix for details).

In a MV-framework, risk-adjusting presumes an equal interest rate regarding borrowing as well as the risk-free investment. However, real costs for borrowing and a (quasi) risk-free investment typically vary from each other. To closer approximate households’ risk/return-enhancements from XTFs, we use distinct rates for borrowing and the risk-free investment. We apply SVs as (quasi) risk-free investment \( (r_{SV}) \) which means a monthly rate of 0.12 percent. A proxy for the borrowing rate is drawn from the security lending rate \( (r_{SL}) \) by Stiftung Warentest (2013), a leading German consumer foundation. According to a comparison of security loans offered by large German banks, households had to pay a rate of 5.5 percent per annum (0.46 percent per month) on average for a loan on security lending.

Risk-adjusting is accomplished by shifting equal parts of the wealth invested in the portfolio under consideration into SVs (to decrease portfolio risk), or by taking a security loan of the rate of \( r_{SL} \) and investing (in equal parts) in the portfolio under consideration (to increase portfolio risk), respectively.

In this way, we determine so-called Return Differences (RD). For the applied risk measures, RD based on the risk of HPT portfolio \( i \) \( (RD_{HPTi}) \) can formally be described by:

---

monthly inflation rate of 0.14 percent; for the period between January 2014 and December 2016, we determine a rate of 0.05 percent.
Thereby, $RD_{HPTi}$ is ascertained by the mean return of the benchmark portfolio ($\mu_{BM}$) which is risk-adjusted to the risk of the HPT portfolio $i$ ($\mu_{BM,HPTi}$), minus the mean return of the HPT portfolio $i$ ($\mu_{HPTi}$). Results for all risk measures are computed in monthly terms. $\mu_{BM,HPT}$ is determined for SD and the LPM-based risk measures by (see Scholz and Wilkens, 2005; see Matulich, 1975 regarding differing lending and borrowing rates)

$$
\mu_{BM,HPTi} = \begin{cases} 
    r_{SL} + \frac{\mu_{BM} - r_{SL}}{risk_{BM}} \text{risk}_{HPTi}, & \text{if } risk_{BM} < risk_{HPTi} \\
    r_{SV} + \frac{\mu_{BM} - r_{SV}}{risk_{BM}} \text{risk}_{HPTi}, & \text{if } risk_{BM} > risk_{HPTi} \\
    \mu_{BM}, & \text{if } risk_{BM} = risk_{HPTi}
\end{cases}
$$

where SD and the LPM-based risk measures are generally denoted as risk and differing rates for borrowing and the risk-free investment are considered.

### 4 Results and discussion

#### 4.1 Return distribution of HPTs

The cluster analysis of German households’ portfolio composition by Oehler and Wanger (2020) yields three HPTs. Table 1 reveals the total portfolio value (VALUEpf), i.e. the aggregated amount of money invested across all securities, and the asset class weights of the HPTs, i.e. the relative share of VALUEpf invested in each asset class (average value of all households assigned to a cluster). HPT 1 includes most of the households of the PHF-survey (total sample size: 1,052 households). In addition, HPT 1 reveals the smallest VALUEpf (163,900 Euros), a clear focus on safe financial assets making up 70.5 percent of VALUEpf as well as considerable amounts in SF and ST. Households assigned to HPT 2 exhibit on average a higher VALUEpf (214,200 Euros), invest about one third of VALUEpf among safe financial assets and allocate a larger fraction to mutual funds (i.e. SF, BF and REF). HPT 3 shows the highest VALUEpf (312,900 Euros), the least amount invested in safe financial assets (25.0 percent) and a heavy portfolio concentration of 63.7 percent invested into ST.
The distribution of HPTs’ portfolio returns (1,000 portfolios per HPT) are described in Tables 2a and 2b.\textsuperscript{10} Table 2a pools monthly portfolio returns for each HPT and shows summary statistics. Table 2b considers each portfolio separately and reveals summary statistics on a portfolio basis for each HPT. According to Table 2a and 2b, portfolios of HPT 1 exhibit on average the lowest mean return per month of approximately 0.15 percent, while portfolios of HPT 2 (HPT 3) achieve 0.22 percent (0.47 percent), respectively.

Differences in risk evaluation according to SD and downside-risk measures particularly arise if returns are skewed and non-normally distributed (see e.g. Harlow, 1991; Jarrow and Zhao, 2006). Table 2b shows that portfolios of HPT 1 reveal on average the highest skewness value (0.82) compared to the skewness values of HPT 2 (0.40) and HPT 3 (0.10). Additionally, each portfolio’s return distribution is tested for normality. According to the Jarque-Bera test, however, the assumption of normally distributed portfolio returns must be rejected for 34.1 percent (49.7 percent) of the portfolios of HPT 1 at the one (five) percent significance level. For portfolios of HPT 2, the assumption of normally distributed portfolio returns must be rejected for 14.0 percent (22.1 percent) and only 2.9 percent (5.2 percent) for portfolios of HPT 3 at the one (five) percent significance level, respectively. Hence, the test indicates that most portfolio returns of HPT 2 and HPT 3 follow a normal distribution. Nevertheless, a substantial percentage of portfolio returns seems not normally distributed. The results on skewness and return distribution suggest that possible differences in risk evaluation between SD and downside-risk measures are more likely to occur for portfolios of HPT 1 than for portfolios of HPT 2 and HPT 3.

\textsuperscript{10} For each portfolio, 36 months of returns are available which yields 108,000 returns on aggregate across the entire HPTs’ portfolios. The highest and lowest 0.1 percent of the returns were excluded as random checks suggested that some extreme outliers are driven by data errors and might lead to misinterpretations.
4.2 Risk and return of benchmark portfolios and HPTs’ portfolios

Table 3 outlines the asset class weights, the reference period, and the assets applied in the benchmark portfolios. While the assumed asset class weights of the benchmark portfolio of HPT 2 are equally distributed among safe financial assets, bonds, and stocks, the benchmark portfolios of HPT 1 and HPT 3 represent asset class concentrations of 70 percent in safe financial assets and stocks, respectively. The target returns are drawn from the five-year pre-observation period from which households are assumed to draw their target return expectations (i.e. January 2009 to December 2013). In this way, we ascertain a mean target return for HPT 1 (HPT 2, HPT 3) of 0.33 percent per month (0.52 percent, 0.87 percent). The target returns all differ from the mean values of HPTs’ portfolio returns. This suggests that differences in risk evaluation occur.

[ Please insert Table 3 here ]

Risk and return of the benchmark portfolios and the HPTs’ portfolios are outlined in Table 4a and Table 4b. The benchmark portfolios are not (yet) risk-adjusted. The mean return of all benchmark portfolios is higher than the mean return of the respective HPTs’ portfolios. Skewness values of the benchmark portfolios are below the average skewness values of the respective HPT portfolios. This effect might be due to higher levels of diversification inherent to the applied benchmark portfolios which are based on XTFs (see e.g. Hueng and Yau, 2006; Mitton and Vorkink, 2007; Kim, 2015).

[ Please insert Table 4a here ]

Benchmark portfolios of HPT 1 and HPT 2 reveal (slightly) higher SDs than the respective HPT portfolios on average. In contrast, when considering the LPM0, LPM1, $\sqrt{LPM2}$ and MDD as measure of risk, the benchmark portfolios of all HPTs exhibit lower risk than the respective HPT portfolios on average. This implies that loss-averse households that implicitly evaluate risk according to downside-risk measures and hold – instead of their current portfolio – their benchmark portfolio, obtain on average higher returns and less risk.
4.3 Return Differences (RDs)

RD is the performance measure employed to determine risk/return-enhancements from XTFs. RDs according to the MV-, mean-LPM0 (M-LPM0)-, mean-LPM1 (M-LPM1)- and mean-$\sqrt{LPM2}$ (M-LPM2)-framework are presented in the upper part of the Tables 5a to 5c. HPT 1 shows the least RDs on average (see Table 5a). If the benchmark portfolio is risk-adjusted to the risk of HPT 1’s portfolios (upper part of the Table 5a), RDs range between 0.14 percent and 0.15 percent monthly return. RDs for HPT 2 are generally higher than for HPT 1 (see Table 5b). The least RDs for HPT 2, amounting to 0.27 percent monthly return, occur in a MV-framework, while the highest average RDs, amounting to 0.36 percent, can be obtained in a M-LPM1-framework. RDs of HPT 3 are higher than those of the portfolios of HPT 2 (see Table 5c). The minimum and maximum enhancements can be obtained in a M-LPM0- and M-LPM1-framework amounting to 0.36 percent and 0.58 percent monthly return, respectively.

As a first robustness test, we compute RDs which are based on the risk of the respective benchmark portfolio. This is the reverse case compared to the initial RDs and involves risk-adjusting the HPTs’ portfolios to the risk of the respective benchmark portfolio (based on the Risk-adjusted Performance measure by Modigliani and Modigliani, 1997; see bottom part of the Tables 5a to 5c). For most HPTs’ portfolios, this means to deleverage as they reveal higher risk than the benchmark portfolio. Overall, the robustness RDs remain positive. Deviations from the initial RDs mainly occur since the basis of risk is now on average lower and the costs for deleveraging HPTs’ portfolios differ from the costs to leverage the benchmark portfolio to compute the initial RDs.

As a second robustness test, we control for risk/return-enhancements according to the MDD (see upper part of the Tables 5a to 5c). As opposed to the other risk measures, the periods which
define the MDD of a portfolio may spread over more than one month and the starting and ending month of each MDD can differ. Therefore, MDDs cannot be compared directly with the enhancements of the remaining risk measures. Nevertheless, MDD allows controlling for an additional, non-LPM-based downside-risk measure that indicates possible enhancements from XTFs in terms of returns. However, controlling for MDDs does not change our results. The MDDs of all HPTs are on average positive while HPT 1 (HPT 3) shows, according to RDs, the least (largest) average MDD. So far, we conclude that XTFs can on average enhance the portfolios of all HPTs, regardless of the applied risk measure.

When comparing mean RDs across the employed risk/return-frameworks and HPTs, most RDs of downside-risk/return-frameworks range above those of the MV-framework. However, RDs do not seem to substantially vary in absolute terms. Consider, for example, the mean RDs according to the MV- and M-LPM2-framework. Monthly RDs for HPT 1 of 0.1400 and 0.1546 percent return, respectively, translate into approximately 1.69 and 1.87 percent annual RD. Based on a VALUEpf of 163,900 Euros for HPT 1, the difference in RD of 0.18 percent results in 292 Euros per annum which does not seem to be an essential amount relative to VALUEpf.11 In addition, these values have been computed ex-post. Considering that households have to derive expected portfolio returns under ambiguity ex-ante, these values could also be in the range of forecast errors.

In this regard, we examine if RDs of the downside-risk/return-frameworks statistically differ from the RDs of the MV-framework. Therefore, we first assess for the RDs of each risk/return-framework and HPT, if the RDs follow a normal distribution. According to the Shapiro-Wilk test, the assumption that RDs follow a normal distribution must be rejected at the one percent significance level for all risk/return-frameworks and HPTs. Based on this result, we apply a paired, two-sided Wilcoxon signed rank test. The test shows that the assumption that RDs of

11 The respective annual amounts for HPT 2 and HPT 3 reach 1,053 and 891 Euros and allow, given a VALUEpf of 214,200 and 312,900 Euro, the same conclusion.
the downside-risk/return-frameworks follow the same distribution as RDs of the MV-framework must be rejected at the one percent significance level for all initial RDs (see upper part of Tables 5a to 5c). Except for the M-LPM0-framework in HPT 1 and HPT 2, the test indicates the same result for the robustness RDs (see bottom part of Tables 5a to 5c). Overall, most RDs of downside-risk/return-frameworks statistically differ from those of the MV-framework although the differences in terms of percentage points of return appear rather small. We also check if one of the downside-risk/return-frameworks yields consistently lower RDs across all HPTs compared to the MV-framework. For households which evaluate risk and return according to that specific downside-risk/return-framework, XTFs might then appear less attractive. Thus, the regarding downside-risk measure could help to explain why certain households refuse to invest in XTFs. To investigate this, we employ a paired, one-sided Wilcoxon signed rank test and test the hypothesis that the RDs of downside-risk measures are statistically lower than those of the MV-framework. Except for RDs of two frameworks in HPT 3,\(^\text{12}\) however, the test reveals that RDs of downside-risk/return-frameworks are not significantly lower. In contrast, most downside-risk/return-frameworks show statistically higher RDs than the MV-framework (see also mean RDs in Tables 5a to 5c). We therefore conclude that none of the employed downside-risk measures can help to explain the reluctance of households to invest in XTFs. Moreover, instead of indicating fewer incentives to invest in XTFs, it rather appears promising to invest in XTFs from the perspective of households that evaluate risk according to downside-risk measures.

In context of our result of statistically different RDs, Jarrow and Zhao (2006) find little differences between optimal portfolios in a MV- and LPM-framework if the returns are nearly normally distributed. The authors thereby use portfolio optimization techniques to construct stock/bond-portfolios instead of a random portfolio selection of households’ portfolios. Das et

\(^{12}\) RDs in the M-LPM0- (initial and robustness RDs) and the M-LPM2-framework (only robustness RDs) of HPT 3 are statistically lower than those of the MV-framework at the one percent significance level.
al. (2010) and Pfiffelmann et al. (2016) demonstrate that portfolios which are optimal in a BPT-related framework can also be MV-efficient. We contribute to these studies by also showing, although statistically different, small absolute differences in risk/return-enhancements between downside-risk/return-frameworks and the MV-framework. As opposed to previous studies, we use randomly constructed HPTs and an easily investable benchmark XTF portfolio instead of portfolio optimization techniques which might lead to different extents of enhancements.

Other studies outline differences in households’ interpretation of risk. A reason for this might be a different emphasis and research design. For instance, while the experiments of Unser (2000) and Veld and Veld-Merkoulova (2008) ask participants from an ex-ante perspective to evaluate the risk of (hypothetical) security returns, we assume existing portfolios and determine the (downside-) risk from an ex-post perspective.

Regarding the puzzling reluctance of households to invest in XTFs, two alternative explanations are conceivable. First, financial advice is prevalent among German households (see DAB Bank, 2004; Hackethal et al., 2011). A reason why households hardly employ XTFs may be that financial advisors barely recommend them to do so. This is based on the finding that the incentive structure of financial advisors typically motivates them to recommend high-fee products (see e.g. Christoffersen et al., 2013; von Gaudecker, 2015; Chalmers and Reuter, 2015; Egan, 2019) as opposed to XTFs which typically charge low fees instead.

Second, knowledge illusion might add an explanation. Baars and Goedde-Menke (2019) find that individuals tend to distinguish between different sources of risk based on their perceived expertise. Because perceived expertise does not necessarily coincide with the actual expertise of an individual, making decisions under risk can involve what the authors refer to as “knowledge illusion”. As one source of risk, they discuss the home bias puzzle. The authors argue that the higher expertise which individuals perceive towards geographically close investments (see e.g. Kilka and Weber, 2000; Ackert et al., 2005) can induce knowledge illusion since the preference for geographical proximity is usually not based on value-relevant
information (see Seasholes and Zhu, 2010). In this way, knowledge illusion increases the attraction of investments which are geographically close to the individual and can be associated with less diversified portfolios (see also Dimmock et al., 2018). Hence, knowledge illusion can add an explanation why households refrain from XTFs since XTFs imply, as opposed to geographical proximity, investing internationally for which households perceive, in reference to Baars and Goedde-Menke (2019), less expertise and fewer incentives.

4.4 Limitations

A first limitation of our analysis is the relatively short observation period. If a specific stock market that has, for example, a major influence on the employed stock XTF, but not on the HPTs, performs exceptionally well during the observation period, RDs might overestimate the risk/return-enhancements from XTFs. Second, we assumed that households’ target returns are fixed throughout the entire observation period. However, households might adjust their desired target return over time. Third, by relying on stylized portfolio compositions of households we do not cover the entire diversity of individual household portfolios. Nevertheless, given the diversity in households’ portfolios and the availability of household portfolio data, we believe our procedure represents a best estimate.

5 Conclusions

The goal of this paper is to investigate if households can obtain less risk/return-enhancements from employing XTFs under assumptions that approximate their actual risk evaluation more closely than the MV-framework which is applied in most existing studies. Less enhancements offer fewer incentives for households to invest in XTFs. We raise the question whether downside-risk measures can help to explain the reluctance of households to invest in XTFs. To the best of our knowledge, we are the first to analyze risk/return-enhancements from XTFs according to several downside-risk measures while considering multiple relevant asset classes.
of household portfolios. A further advantage is that our approach is based on representative household portfolio data.

Our results show, first, that the risk/return-enhancements according to each downside-risk measure are statistically different from the risk/return-enhancements when the SD is the used measure of risk. However, none of the applied downside-risk measures leads to consistently lower risk/return-enhancements on average compared to SD. We thus conclude that none of the employed downside-risk measures can help to explain the reluctance of households to invest in XTFs. Second, all risk/return-frameworks, regardless of whether SD or downside-risk is the underlying measure of risk, indicate that households can enhance their portfolio performance by employing XTFs. This substantiates the common recommendation of academics to employ XTFs and suggests that the advice holds true if households evaluate risk according to downside-risk measures and take multiple relevant asset classes into account.
References


Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (BMELV)].
Wissenschaftliche Studie im Auftrag des früheren Bundesministeriums für Ernährung,
Landwirtschaft und Verbraucherschutz (BMELV).


### Table 1: Household Portfolio Types (HPTs) of German households’ portfolios

<table>
<thead>
<tr>
<th>HPT (N=1,052)</th>
<th>VALUEpf</th>
<th>CASH&amp;SV</th>
<th>CASH</th>
<th>SV</th>
<th>SF</th>
<th>BF</th>
<th>REF</th>
<th>BD</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPT 1 (n=463)</td>
<td>avg.</td>
<td>163,900</td>
<td>70.5</td>
<td>8.0</td>
<td>62.5</td>
<td>7.4</td>
<td>3.5</td>
<td>4.1</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>stDev.</td>
<td>509,900</td>
<td>12.9</td>
<td>9.0</td>
<td>13.6</td>
<td>11.2</td>
<td>8.5</td>
<td>10.3</td>
<td>10.3</td>
</tr>
<tr>
<td>HPT 2 (n=366)</td>
<td>avg.</td>
<td>214,200</td>
<td>32.4</td>
<td>16.7</td>
<td>15.7</td>
<td>25.4</td>
<td>10.0</td>
<td>8.1</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td>stDev.</td>
<td>436,300</td>
<td>22.9</td>
<td>20.2</td>
<td>12.7</td>
<td>26.9</td>
<td>19.1</td>
<td>17.8</td>
<td>25.1</td>
</tr>
<tr>
<td>HPT 3 (n=223)</td>
<td>avg.</td>
<td>312,900</td>
<td>25.0</td>
<td>8.7</td>
<td>16.3</td>
<td>5.1</td>
<td>0.9</td>
<td>1.3</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>stDev.</td>
<td>759,500</td>
<td>17.7</td>
<td>10.8</td>
<td>15.6</td>
<td>10.0</td>
<td>3.8</td>
<td>4.6</td>
<td>10.4</td>
</tr>
</tbody>
</table>

Notes: This table shows the results of Oehler and Wanger (2020) who perform a K-Means cluster analysis according to German households’ asset class weights using a subsample of the PHF-survey of Deutsche Bundesbank. The clusters represent stylized portfolio compositions of German households, so-called Household Portfolio Types (HPTs). For all households which were assigned to a certain cluster, the table reports the average standard deviation and the average percentage shares of households’ wealth invested into cash (CASH), savings (SV), stock funds (SF), bond funds (BF), real estate funds (REF), bonds (BD) and stocks (ST) relative to households’ total portfolio value of all included asset classes (VALUEpf) in Euros. CASH&SV aggregates safe financial assets.
Table 2a: Return distribution of HPTs’ monthly portfolio returns

<table>
<thead>
<tr>
<th></th>
<th>HPT 1</th>
<th>HPT 2</th>
<th>HPT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean value [%]</td>
<td>0.1505</td>
<td>0.2200</td>
<td>0.4732</td>
</tr>
<tr>
<td>Median [%]</td>
<td>0.0653</td>
<td>0.1719</td>
<td>0.4368</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.9145</td>
<td>0.5978</td>
<td>0.1399</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.1341</td>
<td>3.2044</td>
<td>0.6039</td>
</tr>
</tbody>
</table>

Notes: The table reports the mean value, median, skewness, and kurtosis of HPTs’ monthly portfolio returns for the period from January 2014 to December 2016. The measures consider the returns of all portfolios of a respective HPT jointly.

Table 2b: Return distribution of HPT portfolios’ monthly portfolio returns

<table>
<thead>
<tr>
<th></th>
<th>HPT 1</th>
<th>HPT 2</th>
<th>HPT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean value [%]</td>
<td>0.1517</td>
<td>0.2225</td>
<td>0.4747</td>
</tr>
<tr>
<td>Median [%]</td>
<td>0.0712</td>
<td>0.1989</td>
<td>0.3951</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.8212</td>
<td>0.3513</td>
<td>0.1026</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.3751</td>
<td>1.0377</td>
<td>0.1440</td>
</tr>
</tbody>
</table>

Based on the Jarque-Bera test, the assumption that HPT 1’s portfolio returns are normally distributed must be rejected for 34.1 percent (49.7 percent) of the portfolios at the 1 percent (5 percent) significance level.

<table>
<thead>
<tr>
<th></th>
<th>HPT 2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean value [%]</td>
<td>0.2225</td>
<td>0.1505</td>
<td>0.4747</td>
</tr>
<tr>
<td>Median [%]</td>
<td>0.1989</td>
<td>0.0712</td>
<td>0.3951</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.3951</td>
<td>0.8212</td>
<td>0.1026</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.0377</td>
<td>1.3751</td>
<td>0.1440</td>
</tr>
</tbody>
</table>

Based on the Jarque-Bera test, the assumption that HPT 2’s portfolio returns are normally distributed must be rejected for 14.0 percent (22.1 percent) of the portfolios at the 1 percent (5 percent) significance level.

<table>
<thead>
<tr>
<th></th>
<th>HPT 3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean value [%]</td>
<td>0.4747</td>
<td>0.1505</td>
<td>0.4747</td>
</tr>
<tr>
<td>Median [%]</td>
<td>0.4530</td>
<td>0.0712</td>
<td>0.3951</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.1026</td>
<td>0.8212</td>
<td>0.1026</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.1440</td>
<td>1.3751</td>
<td>0.1440</td>
</tr>
</tbody>
</table>

Based on the Jarque-Bera test, the assumption that HPT 3’s portfolio returns are normally distributed must be rejected for 2.9 percent (5.2 percent) of the portfolios at the 1 percent (5 percent) significance level.

Notes: The table presents the mean value, median, skewness and kurtosis of HPTs’ monthly portfolio returns as of January 2014 to December 2016. Each of the measures was first calculated for all portfolios separately. The table provides the mean, median, standard deviation, minimum and maximum of each measure across all portfolios. Example: Across all portfolios of HPT 3, the standard deviation of portfolio means is 0.4053 percent.
Table 3: Construction of benchmark portfolios

<table>
<thead>
<tr>
<th>Reference asset classes</th>
<th>Asset class weights</th>
<th>Reference period</th>
<th>Illustration</th>
</tr>
</thead>
<tbody>
<tr>
<td>(according to HPTs' asset classes)</td>
<td>HPT 1</td>
<td>HPT 2</td>
<td>HPT 3</td>
</tr>
<tr>
<td>safe financial assets</td>
<td>70.0</td>
<td>33.3</td>
<td>30.0</td>
</tr>
<tr>
<td>(cash + savings)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bonds</td>
<td>10.0</td>
<td>33.3</td>
<td>0.0</td>
</tr>
<tr>
<td>(bond funds + individual bonds)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stocks</td>
<td>20.0</td>
<td>33.3</td>
<td>70.0</td>
</tr>
<tr>
<td>(stock funds + individual stocks)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This Table provides details on the construction of the benchmark portfolios which are used in the analysis to determine risk/return-enhancements from XTFs during the observation period between January 2014 and December 2016. The benchmark portfolios are also employed to ascertain target returns which are applied in the LPM-based downside-risk measures. The mean target returns are drawn from the pre-observation period. This period is defined by the maximum period of time with available return data for all constituents. The weighted average interest rates on German households’ deposits (savings) are applied as benchmark returns for safe financial assets (see Deutsche Bundesbank, 2019). Respective security data on the stock XTF (MSCI World ETF) and bond XTF (Markit iBoxx Euro Sovereign Index ETF) are gathered from Thomson Reuters Datastream.
Table 4a: Risk/return-profile of benchmark portfolios

<table>
<thead>
<tr>
<th></th>
<th>benchmark HPT 1</th>
<th>benchmark HPT 2</th>
<th>benchmark HPT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean value [%]</td>
<td>0.3210</td>
<td>0.5390</td>
<td>0.8030</td>
</tr>
<tr>
<td>Median [%]</td>
<td>0.1720</td>
<td>0.5640</td>
<td>0.7890</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.4795</td>
<td>-0.0736</td>
<td>-0.1405</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.9138</td>
<td>0.7599</td>
<td>0.7206</td>
</tr>
<tr>
<td>Standard deviation [%]</td>
<td>0.9645</td>
<td>1.5118</td>
<td>2.6103</td>
</tr>
<tr>
<td>LPM0 [%]</td>
<td>55.6</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>LPM1 [%]</td>
<td>0.3470</td>
<td>0.5319</td>
<td>0.9743</td>
</tr>
<tr>
<td>$\sqrt{LPM2}$ [%]</td>
<td>0.6218</td>
<td>1.0530</td>
<td>1.8859</td>
</tr>
<tr>
<td>MDD [%]</td>
<td>3.6813</td>
<td>6.2464</td>
<td>9.7569</td>
</tr>
<tr>
<td>Mean target return used in LPM-based risk measures [%]</td>
<td>0.3250</td>
<td>0.5160</td>
<td>0.8650</td>
</tr>
</tbody>
</table>

Notes: The table shows the mean value, median, skewness, kurtosis, standard deviation, Lower-Partial-Moment Zero (LPM0), LPM One (LPM1), LPM Two ($\sqrt{LPM2}$) and the Maximum Drawdown (MDD) of the monthly return distribution of each benchmark portfolio for the period from January 2014 to December 2016. The target returns which are employed in the LPM-based risk measures is the monthly mean benchmark portfolio return from January 2009 to December 2013.

Table 4b: Risk/return-profile of HPT portfolios

<table>
<thead>
<tr>
<th></th>
<th>HPT 1</th>
<th>HPT 2</th>
<th>HPT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean value [%]</td>
<td>0.1517</td>
<td>0.2225</td>
<td>0.4747</td>
</tr>
<tr>
<td>Standard deviation [%]</td>
<td>0.8927</td>
<td>1.4330</td>
<td>3.5432</td>
</tr>
<tr>
<td>LPM0 [%]</td>
<td>65.1</td>
<td>63.7</td>
<td>55.0</td>
</tr>
<tr>
<td>LPM1 [%]</td>
<td>0.4312</td>
<td>0.6966</td>
<td>1.5888</td>
</tr>
<tr>
<td>$\sqrt{LPM2}$ [%]</td>
<td>0.6611</td>
<td>1.1283</td>
<td>2.6669</td>
</tr>
<tr>
<td>MDD [%]</td>
<td>3.8220</td>
<td>6.5721</td>
<td>15.5763</td>
</tr>
<tr>
<td>Mean target return used in LPM-based risk measures [%]</td>
<td>0.3250</td>
<td>0.5160</td>
<td>0.8650</td>
</tr>
</tbody>
</table>

Notes: The table provides respective values for the mean, standard deviation, Lower-Partial-Moment Zero (LPM0), LPM One (LPM1), LPM Two ($\sqrt{LPM2}$) and the Maximum Drawdown (MDD). The values were first calculated for each HPT portfolios’ monthly return distribution separately for the period of January 2014 to December 2016. The values in the table are means of each measure (across all portfolios of a respective HPT). The target returns which are employed in the LPM-based risk measures is the monthly mean benchmark portfolio return from January 2009 to December 2013.
Table 5a: Risk/return-enhancements according to applied risk measures (HPT 1)

<table>
<thead>
<tr>
<th></th>
<th>HPT 1</th>
<th>MV</th>
<th>M-LPM0</th>
<th>M-LPM1</th>
<th>M-LPM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean return of benchmark portfolios (risk-adj.) [%]</td>
<td></td>
<td>0.2917</td>
<td>0.3052</td>
<td>0.2997</td>
<td>0.3063</td>
</tr>
<tr>
<td>Mean return of HPT 1 portfolios [%]</td>
<td></td>
<td>0.1517</td>
<td>0.1517</td>
<td>0.1517</td>
<td>0.1517</td>
</tr>
<tr>
<td>Mean return of benchmark portfolio [%]</td>
<td></td>
<td>0.3210</td>
<td>0.3210</td>
<td>0.3210</td>
<td>0.3210</td>
</tr>
<tr>
<td>Mean return of HPT 1 portfolios (risk-adj.) [%]</td>
<td></td>
<td>0.1895</td>
<td>0.1822</td>
<td>0.1848</td>
<td>0.1843</td>
</tr>
<tr>
<td>Mean RD [%]</td>
<td></td>
<td>0.1400</td>
<td>0.1535</td>
<td>0.1480</td>
<td>0.1546</td>
</tr>
<tr>
<td>Two-sided Wilcoxon test between MV- and M-LPM-frameworks: p-value</td>
<td>- / -</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>MDD of benchmark portfolio [%]</td>
<td></td>
<td>3.6813</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                      |       |       |        |        |        |
| Mean RD [%]          |       | 0.1407 |
| (0.0977)             |       | (1.0355) |

|                      |       |        |        |        |        |
|                      |       | Mean RD [%] |        |        |        |
| Two-sided Wilcoxon test between MV- and M-LPM-frameworks: p-value | - / - | 0.8006 | < 0.001 | < 0.001 |
Table 5b: Risk/return-enhancements according to applied risk measures (HPT 2)

<table>
<thead>
<tr>
<th>HPT 2</th>
<th>MV</th>
<th>M-LPM0</th>
<th>M-LPM1</th>
<th>M-LPM2</th>
<th>M-MDD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean return of benchmark portfolios (risk-adj.) [%]</td>
<td>0.4893</td>
<td>0.5751</td>
<td>0.5798</td>
<td>0.5290</td>
</tr>
<tr>
<td></td>
<td>Mean return of HPT 2 portfolios [%]</td>
<td>0.2225</td>
<td>0.2225</td>
<td>0.2225</td>
<td>0.2225</td>
</tr>
<tr>
<td></td>
<td>Mean RD [%]</td>
<td>0.2668</td>
<td>0.3526</td>
<td>0.3573</td>
<td>0.3065</td>
</tr>
<tr>
<td></td>
<td>(0.1563)</td>
<td>(0.1978)</td>
<td>(0.1912)</td>
<td>(0.1836)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Two-sided Wilcoxon test between MV- and M-LPM-frameworks: p-value</td>
<td>- / -</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

|       | Mean return of benchmark portfolio [%] | 0.5390  | 0.5390  | 0.5390  | 0.5390  |               |       |
|-------| Mean return of HPT 2 portfolios (risk-adj.) [%] | 0.2459  | 0.2008  | 0.1914  | 0.2306  |               |       |
|       | Mean RD [%] | 0.2931  | 0.3382  | 0.3476  | 0.3084  |               |       |
|       | (0.1255)    | (0.2040) | (0.2191) | (0.2669) |               |       |
|       | Two-sided Wilcoxon test between MV- and M-LPM-frameworks: p-value | - / -  | 0.0161  | < 0.001 | < 0.001 |               |       |
### Table 5c: Risk/return-enhancements according to applied risk measures (HPT 3)

<table>
<thead>
<tr>
<th>HPT 3</th>
<th>MV</th>
<th>M-LPM0</th>
<th>M-LPM1</th>
<th>M-LPM2</th>
<th>M-MDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean return of benchmark portfolios (risk-adj.) [%]</td>
<td>0.9447</td>
<td>0.8388</td>
<td>1.0542</td>
<td>0.9672</td>
<td>15.5763</td>
</tr>
<tr>
<td>Mean return of HPT 3 portfolios [%]</td>
<td>0.4747</td>
<td>0.4747</td>
<td>0.4747</td>
<td>0.4747</td>
<td>9.7569</td>
</tr>
<tr>
<td>Mean RD [%]</td>
<td>0.4700</td>
<td>0.3641</td>
<td>0.5795</td>
<td>0.4925</td>
<td>5.8194</td>
</tr>
<tr>
<td>Two-sided Wilcoxon test between MV- and M-LPM-frameworks: p-value</td>
<td>- / -</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Mean return of benchmark portfolio [%]</td>
<td>0.8030</td>
<td>0.8030</td>
<td>0.8030</td>
<td>0.8030</td>
<td>0.0020</td>
</tr>
<tr>
<td>Mean return of HPT 3 portfolios (risk-adj.) [%]</td>
<td>0.3713</td>
<td>0.4652</td>
<td>0.3548</td>
<td>0.3888</td>
<td></td>
</tr>
<tr>
<td>Mean RD [%]</td>
<td>0.4317</td>
<td>0.3378</td>
<td>0.4482</td>
<td>0.4142</td>
<td></td>
</tr>
<tr>
<td>Two-sided Wilcoxon test between MV- and M-LPM-frameworks: p-value</td>
<td>- / -</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.0020</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The tables 5a to 5c provide for each HPT the mean (risk-adjusted) returns and mean Return Differences (RDs) which are determined in a mean-variance (MV), mean-Lower-Partial-Moment Zero (M-LPM0), mean-LPM One (M-LPM1), mean-LPM Two (M-LPM2) and mean-Maximum Drawdown (M-MDD) framework. Below RDs, the standard deviation of RDs is placed into parentheses. Each table is divided into two sections. The upper section reveals the RDs between the benchmark and the respective HPT portfolios while the benchmark portfolio is risk-adjusted to the risk of the HPT portfolios (initial RD). The bottom section outlines the reverse case, i.e. RDs between the benchmark and the respective HPT portfolios while each HPT portfolio is risk-adjusted to the risk of its benchmark portfolio (robustness RD). Cases in which the risk-adjusting of HPT portfolios would both increase the risk and reduce the return the portfolio are excluded. HPTs’ portfolio means are based on monthly portfolio returns from January 2014 to December 2016. As robustness, RDs according to the MDD are provided. Since the period of time for each MDD is individual, MDDs were not risk-adjusted. Example: When risk-adjusting the benchmark portfolio of HPT 3 to the risk of all 1,000 HPT 3 portfolios, the average monthly mean return of the risk-adjusted benchmark portfolios in a M-LPM2 framework is 0.9672 percent. Comparing this with the average mean return of the HPT 3 portfolios (0.4747 percent) yields a mean RD of 0.4925 percent monthly return which implies an enhancement in risk-adjusted returns. Additionally, we provide p-values for a paired, two-sided Wilcoxon signed rank test which was used to test statistical difference between the RDs according to the MV- versus the employed downside-risk/return-frameworks.
Methodological Appendix

Methodological Appendix to chapter 3.1: Estimation of risk/return-profiles of HPTs

A typical challenge when examining households’ investment portfolios is to gather information about households’ financial situation that is both detailed and representative.\(^{13}\) The SHS-base of Deutsche Bundesbank provides detailed security holding data which are representative for the German household sector. The SHS-base collects, on a monthly basis, the entire listed stocks, mutual funds and debt securities which all financial institutions domiciled in Germany store of German clients. The securities are captured by International Securities Identification Number (ISIN). Deutsche Bundesbank also provides for each security the aggregated market value of shares owned by German households. This is the aggregated number of shares owned by German households, times the current market price (end-of-month) in Euro (for a detailed description of the acquisition of the SHS-data see Bade et al., 2017). The special feature of the aggregated market values is that they only take shares held by German households into account. Hence, aggregated market values are a reasonable indication for the distribution of a security among German households on an aggregate basis.\(^{14}\)

For each security of the SHS-base, respective asset class information and daily total return price data (including dividends, payouts and interest payments) were requested from Thomson Reuters Datastream. This allows us to categorize each security into SF, BF, REF, BD\(^{15}\) or ST, and to calculate discrete monthly security returns. If there were no data available or if a security could not be clearly categorized into one of the employed asset classes, it was dropped from the data set. We exclude securities when they reveal negative market values of shares, since this

---

\(^{13}\) See e.g. von Gaudecker (2015) for a discussion regarding this aspect.

\(^{14}\) Securities which are held at less than three institutions are excluded from the data set. They can barely be assumed to be representatively spread among German households. Moreover, please note that we do not analyze households’ portfolios on the individual level. With regard to current market prices, we assume households to be price takers and to follow the pricing by professional investors.

\(^{15}\) BDs include bonds with a fixed coupon (straight bonds and zero bonds) as well as bonds with a variable coupon (floating bonds and index linked bonds). Returns for the latter group of bonds were calculated using total return price data from Thomson Reuters Datastream. Of the total 22,225 individual bonds in our sample, 14,291 are designated as straight bonds, 1,247 as zero bonds, 6,660 as floating bonds and 27 as index linked bonds.
indicates bankruptcy and implies that the shares of the respective security cannot be purchased anymore by households.

The security sample also includes mutual funds which follow a mixed asset strategy, i.e. mixed funds. Mixed funds are, according to their market value of their fund shares held by German households, among the most widespread mutual funds of German households. Each mixed fund is categorized as SF or BF. The categorization fits with questionnaire of the PHF-survey. Therein, households who hold mutual funds were asked about the asset class which their mutual fund predominantly invests in. For the categorization, we calculate correlations of the returns of all mixed funds included in the sample with twelve stock and nine bond market indices (see Table 6 for the applied indices). The highest correlation value with a respective stock or bond index determines the categorization. Initially, we identify 1,970 mixed funds in total. Thereof, 49 mixed funds were excluded since they reveal less than three months of return which prevents a proper calculation of correlations. We end up with 1,921 mixed funds in our sample. Thereof, 1,540 are assigned to SFs and 381 are assigned to BFs.
### Table 6: Stock and bond indices for the categorization of mixed funds

<table>
<thead>
<tr>
<th>num.</th>
<th>Index category</th>
<th>Name</th>
<th>Ticker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stock index</td>
<td>MSCI NORTH AMERICA E</td>
<td>MSNAMRE</td>
</tr>
<tr>
<td>2</td>
<td>Stock index</td>
<td>MSCI EUROPE E</td>
<td>MSEXERPE</td>
</tr>
<tr>
<td>3</td>
<td>Stock index</td>
<td>MSCI AC ASIA PACIFIC E</td>
<td>MSAAPFEP</td>
</tr>
<tr>
<td>4</td>
<td>Stock index</td>
<td>MSCI EM E</td>
<td>MSEMKE</td>
</tr>
<tr>
<td>5</td>
<td>Stock index</td>
<td>MSCI AC WORLD :SM E</td>
<td>MSZAWFE</td>
</tr>
<tr>
<td>6</td>
<td>Stock index</td>
<td>MSCI AC WORLD :L E</td>
<td>MSLAWFE</td>
</tr>
<tr>
<td>7</td>
<td>Stock index</td>
<td>S&amp;P 500/CITIGROUP PURE VALUE</td>
<td>SP05PVA</td>
</tr>
<tr>
<td>8</td>
<td>Stock index</td>
<td>S&amp;P 500/CITIGROUP PURE GROWTH</td>
<td>SP05PGR</td>
</tr>
<tr>
<td>9</td>
<td>Stock index</td>
<td>S&amp;P 500 DIVIDENDS ARISTOCRATS</td>
<td>SP5DIAR</td>
</tr>
<tr>
<td>10</td>
<td>Stock index</td>
<td>DAX 30 PERFORMANCE</td>
<td>DAXINDEX</td>
</tr>
<tr>
<td>11</td>
<td>Stock index</td>
<td>EURO STOXX</td>
<td>DJEURST</td>
</tr>
<tr>
<td>12</td>
<td>Stock index</td>
<td>MSCI AC WORLD E</td>
<td>MSACWFE</td>
</tr>
<tr>
<td>13</td>
<td>Bond index</td>
<td>IBOXX EURO CORPORATES</td>
<td>IBCRPAL</td>
</tr>
<tr>
<td>14</td>
<td>Bond index</td>
<td>IBOXX EURO OVERALL</td>
<td>IBEURAL</td>
</tr>
<tr>
<td>15</td>
<td>Bond index</td>
<td>IBOXX EURO LIQUID SOVEREIGNS GLOBAL</td>
<td>IBELSVE</td>
</tr>
<tr>
<td>16</td>
<td>Bond index</td>
<td>IBOXX EURO LIQUID SOVEREIGNS CAPPED 2.5 - 5.5</td>
<td>IBELSCD</td>
</tr>
<tr>
<td>17</td>
<td>Bond index</td>
<td>IBOXX EURO LIQUID SOVEREIGNS CAPPED 5.5 - 10.5</td>
<td>IBELSCE</td>
</tr>
<tr>
<td>18</td>
<td>Bond index</td>
<td>IBOXX EURO LIQUID SOVEREIGNS CAPPED 10.5+</td>
<td>IBELSCC</td>
</tr>
<tr>
<td>19</td>
<td>Bond index</td>
<td>S&amp;P MUNICIPAL BOND HIGH YIELD INDEX</td>
<td>SMPUBHY</td>
</tr>
<tr>
<td>20</td>
<td>Bond index</td>
<td>S&amp;P 500 HIGH YIELD CORP BOND INDEX</td>
<td>SP5HYBI</td>
</tr>
<tr>
<td>21</td>
<td>Bond index</td>
<td>IBOXX EURO HY FIXED RATE</td>
<td>IBEHYFR</td>
</tr>
</tbody>
</table>

Notes: This table outlines details and tickers to the applied stock and bond indices which were used to calculate correlations with mixed funds and categorize the latter into stock and bond funds. Of each index, the total return version was employed from Thomson Reuters Datastream.

The resulting sample of security holdings is outlined in Table 7. The time span of available security returns can vary for each security and might be less than 36 months included in the observation period from January 2014 to December 2016. BDs show the largest differences between the entire securities of an asset class and the average numbers of available securities per month, which implicates that BDs reveal varying time spans most often compared to the other asset classes (e.g. due to the maturity of a bond).

Moreover, the Top 100 securities (by market value) exhibit a major part in every asset class. Among STs, they account on average for more than 75 percent of the market value of all STs. For BF, they exceed 60 percent and for SFs 50 percent of the respective asset class’ market value. REFs are few in number of securities, but the aggregated market value among all REFs indicates their importance compared to the remaining asset classes as they approximately reach the aggregated market value of BF (see Table 7).
Most households do not hold more than ten securities (see e.g. von Gaudecker, 2015; Bhattacharya et al., 2017). Considering the lack of information about future price movements (see e.g. Elton and Gruber, 1977), it seems to be a reasonable strategy from a households’ perspective to spread the wealth among their investments in equal parts (see e.g. De Wit, 1998; Benartzi and Thaler, 2001). This implies an equal division of three securities among stocks, bonds and mutual funds and yields nine securities for each portfolio in total, of which the three mutual funds are equally assigned to the mutual fund types, i.e. one security each for SF, BF and REF. The assumption on portfolio size, i.e. the number of securities per portfolio, corresponds to the base case scenario in Oehler and Wanger (2020). In contrast to the latter, we do not constrain the security holding data set to securities with the highest aggregated market value of shares owned by German households. This largely extends our security sample obtained from the SHS-base. The extension of the security data base is used to draw random samples of securities to compile HPTs and calculate risk/return-profiles. Following this method can substantiate the results in Oehler and Wanger (2020) and increase the generalizability of
the results of the present analysis. To avoid selection bias, each random portfolio selection (i.e. nine securities) is applied to every HPT. This means that the same security returns are applied on all three HPTs, however, adjusted by the HPTs’ respective asset class weights.

Regarding the monthly interest rates on CASH and SV, German households’ deposits with different agreed maturities are applied. Interest rates and outstanding amounts on these deposits allow calculating weighted (according to outstanding amounts) monthly interest rates (data obtained of Deutsche Bundesbank, 2019). The interest rate for SV was compiled as weighted average of households’ deposits with agreed maturities of up to one year, over one and up to two years, and over two years. For CASH, we choose the interest rate on overnight deposits. Compared to the common advice to buy-and-hold XTFs, households might refuse to accept the effort and transaction costs to replace their current portfolio with XTFs. An “easier-to-follow alternative” compared to the advice to buy-and-hold XTFs might, from a household’s perspective, be to avoid trading and to hold their current securities and only reinvest in the next security if necessary (e.g. if a bond expires). This means that the securities are held until they expire. Then, the available amount of money is reinvested in the subsequent month into the next security of the same asset class. The next security is, again, randomly selected according to the market value of shares owned by German households.

Methodological Appendix to chapter 3.3: Benchmark Portfolios

As an appropriate market index for stocks, the MSCI World Index is employed. This index is commonly used as an international benchmark for stocks (see e.g. Campbell, 2006; Calvet et al., 2007) and well-covered by respective XTFs (ISIN of the applied stock XTF: LU0392494562). As benchmark index for bonds, we choose the Markit iBoxx Euro Sovereign Index (see e.g. Jacobs et al., 2014) which is also replicated by an appropriate XTF and is readily attainable by German households (ISIN of the applied bond XTF: LU0290355717).
Due to difficulties in finding an adequate diversified real estate market index as well as a respective XTF, we do not employ REFs in our benchmark portfolios (see e.g. Jacobs et al., 2014). The last remaining assets of households’ (speculation) portfolio are safe financial assets (i.e. CASH and SV).

Methodological Appendix to chapter 3.4: Measurement of Risk/return-Enhancements

Besides computing the RD, we also ascertain the robustness RD which represents the reverse case of the initial RDs. In this case, the HPTs’ portfolios are risk-adjusted to the risk of the respective benchmark portfolio. The conceptual foundations of this procedure go back to the Risk-adjusted Performance measure of Modigliani and Modigliani (1997). Robustness RDs ($RD_{BMi}$) are based on the risk of the benchmark portfolio of HPT portfolio $i$ and are computed by:

$$RD_{BMi} = \mu_{BM} - \mu_{HPTi,BM}.$$  

Thereby, the monthly mean return of the HPT portfolio $i$ ($\mu_{HPTi}$), which is risk-adjusted to the risk of its respective benchmark ($\mu_{HPTi,BM}$), is defined by:

$$\mu_{HPTi,BM} = \begin{cases} r_{SL} + \frac{\mu_{HPTi}-r_{SL}}{\text{risk}_{HPTi}} \text{risk}_{BM}, & \text{if } \text{risk}_{HPTi} < \text{risk}_{BM} \\ r_{SV} + \frac{\mu_{HPTi}-r_{SV}}{\text{risk}_{HPTi}} \text{risk}_{BM}, & \text{if } \text{risk}_{HPTi} > \text{risk}_{BM} \\ \mu_{HPTi}, & \text{if } \text{risk}_{HPTi} = \text{risk}_{BM} \end{cases}$$

Considering the risk-adjusted performance measurement in the way described above, it might be the case that the risk of a certain HPT portfolio is lower than the risk of its corresponding benchmark, while the monthly mean return of the respective HPT portfolio is simultaneously lower than the monthly security lending rate. Risk-adjusting the HPT portfolio to the risk of the benchmark would then imply both a reduction in returns and an increase in risk. From the perspective of households, this does not seem to be an attractive strategy, particularly for those households whose portfolios reveal low portfolio risk and return. We thus exclude those cases to avoid misconceptions and ambiguous conclusions.
Both the RD and the robustness RD analogously assess performance differences between each HPT portfolio and the benchmark portfolio in terms of returns. In context of MV portfolio optimization, the continuous increase of the wealth invested in the risk-free asset leads to a proportional reduction or increase in portfolio risk.\textsuperscript{16} The proportional relation of portfolio separation has also been shown for LPM-based risk measures.\textsuperscript{17} In case of an asymmetric interpretation of risk, a risk-free investment is the maximum attainable rate of return which (quasi certainly) does not fluctuate below a certain level of return (see e.g. Hogan and Warren, 1974). In this respect, we assume that households prefer investments with longer-term maturities but higher fixed interest rates over the reverse case and choose the interest rate for SV. This allows analogously transferring the Risk-adjusted Performance by Modigliani and Modigliani (1997) on the LPM-based downside-risk measures in our analysis.

\textsuperscript{16} In the context of portfolio optimization, this separation is referred to as Tobin separation (see Tobin, 1958).

\textsuperscript{17} Hogan and Warren (1974) point out the validity of the relation for \( n = 2 \) and the risk-free asset as portfolio target return. Bawa and Lindenberg (1977) and Harlow and Rao (1989) demonstrate the relation for \( n = 1 \) and arbitrary choices for portfolio target returns. We consequently assume the proportional relation of portfolio separation for our analysis in which we employ compiled benchmarks of respective market returns as portfolio target returns and include the PL (\( n = 0 \)) as measure of risk. Please note in addition that the risk-free investment as target return of a portfolio should be separated from the risk-free investment as (quasi) downside-risk-free investment which allows risk-adjusting portfolios.