Idiosyncratic financial risk and a reevaluation of the market risk-return tradeoff *

Sung Je Byun[†] Johnathan A. Loudis[‡] Lawrence D.W. Schmidt[§]

November 2022

Abstract

A value-weighted portfolio of US stocks is not a well-diversified portfolio. While a substantial amount of the variation in index returns can be explained by a single dominant factor, index returns are also driven by nontrivial, time-varying exposures to weaker factors and "granular residuals" – idiosyncratic shocks to large firms that aren't diversified away. We argue, both theoretically and empirically, that these additional components can contaminate tests of the risk-return tradeoff. We then reevaluate the current consensus for a weak market risk-return tradeoff in the US stock market using an alternative index unaffected by them. In the time series, we find stronger evidence for a relation between the risk premium and variance of the market after these corrections. In the cross-section, we find evidence that making these corrections generates larger cross-sectional variation in market betas, and that this exposure to market risk explains a much larger share of variation in expected returns. Finally, in line with our theory, correcting for these errors eliminates the ability of size factors to improve pricing within a large set of standard factor models.

* *JEL classification*: C15; C58; G12; G17

* *Keywords*: Risk-return trade-off; Idiosyncratic risk; Empirical asset pricing

*This is a substantially revised version of an earlier manuscript which was circulated under the title, "Real risk or paper risk? Mis-measured factors, granular measurement errors, and empirical asset pricing tests." We are grateful to John Campbell, John Cochrane, Xavier Gabaix, Stefano Giglio, Daniel Greenwald, Leonid Kogan, Ralph Koijen, Arvind Krishnamurthy, Eben Lazarus, Lira Mota, Jonathan Parker, Seth Pruitt, Cesare Robotti, Andrea Tamoni and seminar participants at MIT, Stanford, the University of Rochester, Rutgers, the NBER Asset Pricing meeting, and the BI/SHOF asset pricing conference for helpful discussions related to the paper. We thank Brice Green, Luxi Han, Maziar Kazemi, Paymon Khorrami, Kyle Kost, and Huben Liu for outstanding research assistance. The views expressed are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

†Financial Industry Studies Department, Federal Reserve Bank of Dallas SungJe.Byun@dal.frb.org.

‡Mendoza College of Business, University of Notre Dame jloudis@nd.edu.

§MIT Sloan School of Management ldws@mit.edu.

1 Introduction

The asset pricing literature struggles to reconcile two seemingly contradictory empirical results about the riskiness of the stock market. On one hand, risk premia on broad stock market indexes are large – puzzlingly high relative to predictions from many macroeconomic models – suggesting investors view market returns as being very risky. On the other hand, measures of the quantity of market risk vary considerably yet are only tenuously linked with expected returns empirically. In the time series, the empirical relationship between the risk premium and conditional variance of the overall stock market is somewhat weak. In the cross-section, measures of covariance with the market line up poorly with differences in expected excess returns across stocks. Unlike the equity premium puzzle, these findings suggest that the market return – or at least a sizable component of the market return – is not very risky.

While it is possible to reconcile these empirical results with theory by introducing additional sources of risk and/or time-variation in risk prices, this paper considers a simple and complementary explanation related to the composition of the value-weighted equity market index itself. We argue that the market portfolio is not a well diversified portfolio, for two reasons. First, some sectors and systematic (e.g., industry-specific) factors likely receive substantially higher weights in the aggregate stock market return – a value-weighted average of publicly traded firm returns – relative to their weights in the overall economy. Second, some very large firms receive nontrivial large weights in the portfolio. Thus, the market return includes a "granular residual," a weighted average of idiosyncratic shocks to large firms that do not diversify away (Gabaix, 2011). A component of the value-weighted market return thus reflects shocks that are potentially unrelated to broader macroeconomic conditions, creating a disconnect between macroeconomic fundamentals and market returns.

In a static model in which investors only hold public equity, the marginal value of wealth (i.e., the stochastic discount factor) is an affine transformation of the stock market return. However, in the data, publicly traded US equities only comprise a small fraction of the total value of resources available to finance a typical investor's consumption, especially after incorporating the value of illiquid/nontradable assets such as private business assets, human capital, and real estate. When most investor wealth is concentrated in these alternative asset classes, their returns dominate the marginal value of wealth. In this case, stock market returns that do not comove with returns on these other components of wealth become less important (Mayers, 1973; Roll, 1977). In turn, investors may require little compensation for holding these additional stock market-specific risks.

In this paper, we reexamine evidence for a market risk-return tradeoff using two different market indices. The first is the standard value-weighted market factor (VMF) which is ubiquitous in the literature and which can be motivated by the Capital Asset Pricing Model (CAPM). The second is a dominant market factor (DMF), which we construct as the linear projection of the VMF onto the first principal component of a large cross-section of equal-weighted, characteristics sorted portfolios. Our alternative index (the DMF) is constructed with the objective of purging two sources of variation in the aggregate market return (the VMF) discussed above: large, time-varying factor exposures specific to highly valuable firms/industries and a granular residual. We argue that these two forces create an "idiosyncratic financial (risk) factor" (IFF, where $IFF \equiv VMF - DMF$), reflecting the macro-financial disconnect.

Our paper's main contribution is to show empirically that the DMF performs better than the VMF in tests of the market risk-return tradeoff and along many other metrics. Our empirical results are also consistent with predictions from our theoretical model, which provides assumptions (for which we find strong empirical support) under which the IFF is priced differently than the DMF. We therefore propose both an explanation for and a potential resolution to the puzzling evidence on the weak market risk-return tradeoff in both the time series and the cross-section. We also provide a simple explanation for the size anomaly (Fama and French, 1992) and find evidence that one need not necessarily include size factors in workhorse empirical asset pricing models once the DMF is used in place of the VMF.

More precisely, our paper makes five main contributions, four empirical and one theoretical. First, we provide five stylized facts related to the existence and effects of large firm concentration in the aggregate stock market, the VMF's exposure to common factors driving stock returns, and links between these common factors and real activity. These facts motivate a theoretical distinction between DMF and IFF risks. Second, we develop a parsimonious theoretical model to help clarify the distinction between risk prices on the DMF and the IFF, respectively, and to illustrate the impact of such a distinction on standard tests of the risk-return tradeoff. Third, we show that the DMF provides stronger evidence for a conditional risk-return tradeoff in the time series than the VMF. Fourth, we show that exposure to the DMF explains more cross-sectional variation in expected returns than exposure to the VMF in a multifactor setting. Fifth, we show that size factors in popular factor models are spanned when the DMF is used in place of the VMF in each model. This result is in line with the IFF theory we develop herein, which implies the size anomaly is related to the mis-measurement of market risk when using the VMF as opposed to the DMF. We now provide more details related to each of these contributions.

Related to our first main contribution, we begin by characterizing two facts related to distribution of weights across firms in the value-weighted market. First, we show that a small number of large firms constitute a large share of the value-weighted market.[1] Second, there are dramatic fluctuations in the industry composition across all firms and, particularly, across the largest firms in the equity market, which may not reflect the composition of the broader economy. These features can combine to generate time-varying VMF loadings on factors other than the DMF and by generating a granular residual, with neither strongly related to macroeconomic fundamentals. Thus, we call the combination of these two effects the IFF.

To motivate our definition of the DMF, our next two facts follow from performing a principal components analysis (PCA) to identify the VMF's exposure to various principal components (PCs) of a large set of equal-weighted characteristic-sorted portfolios.[2] These portfolios have a strong factor structure with the first PC being a level factor: all portfolios load positively on it and all portfolios receive positive weights in its construction. The first PC also explains a large fraction of VMF variation in the time series (84%) and can explain essentially all of its risk premium. In contrast, higher PCs only explain a modest amount of VMF variation. Further, the VMF's exposure to the first PC is stable over time. Consistent with changing composition of large firms inducing time-variation in VMF's systematic exposures, the VMF's

[1]For example, in November 2018 *The Economist* magazine recently reported that "some 37% of the rise in the value of all firms in the S&P 500 index since 2013 is explained by six of its members: Alphabet, Amazon, Apple, Facebook, Microsoft, and Netflix. About 28% of the rise in Chinese equities over the same period is owing to two firms: Alibaba and Tencent." Source: "Big Tech's sell-off", *The Economist*, November 1, 2018.

[2]Importantly, using equal-weighted portfolios effectively imposes that any granular residuals will be averaged out of the base portfolio returns by the law of large numbers. Therefore, the extracted PCs are effectively free of the granular residual component of the IFF. This approach follows a long history of literature that uses statistical factors estimated via PCA to explain the cross-section of stock returns. Early examples include Chamberlain and Rothschild (1983) and Connor and Korajczyk (1986, 1988). More recent examples include Kozak et al. (2018), Pukthuanthong et al. (2019), Kelly et al. (2019), and Giglio and Xiu (2021)

conditional loadings on higher PCs vary dramatically over time.

As our final stylized fact, we use local projections to study links between different components of the VMF and the real economy. First, we show that the projection of the VMF onto the first PC (i.e., our DMF) is highly predictive of changes in nine different macroeconomic time series in directions consistent with it earning a positive risk premium. However, a projection of VMF onto higher PCs is not reliably linked with any of these variables. This evidence points to a likely tenuous link between VMF's exposures to higher PCs and investors' marginal value of wealth.

Our second contribution is theoretical. We build a static model featuring a small number of very large stocks and a large number of small stocks which yields simple expressions for how exposures to the DMF and the IFF translate to risk premia. Our model features both of the above sources of the IFF: factor tilts specific to large firms and a granular residual. While the model nests the CAPM as a special case, in the empirically relevant case in which investors finance most of their consumption from sources of income other than US stock market wealth, the model makes a sharp distinction between the DMF and the IFF. When, as appears to be the case in the data, the IFF comoves little with other sources of non-stock market wealth, its risk price is small and even zero in some limiting cases. Furthermore, our model implies small-cap portfolio βs are biased downward when estimated with respect to the VMF, and the related a small-cap portfolio (IFF) αs are positive (negative), all of which we show holds in the data.

In our third main contribution, we show that the DMF yields a stronger risk-return tradeoff in the time series than does the VMF. We first replicate a familiar result from the literature: the coefficient relating the expected VMF return to its conditional variance is positive yet statistically insignificant.[3] Repeating the analysis using the DMF, the risk-return tradeoff coefficient is larger in magnitude and statistically significant across a wide variety of specifications. When the IFF is not priced, there are factors which contribute to the DMF's conditional variance but not its risk premium, thus making the VMF risk-return tradeoff weaker and unstable over time.

In our fourth main contribution, we find a stronger role for the DMF in explaining the cross-section of expected returns relative to the VMF. More precisely, working in a multifactor setting, we find that using the DMF in place of the VMF as a market

[3] This result is similar to findings of either negative or statistically weak coefficient estimates (Nelson 1991, Glosten et al. 1993, Koopman and Hol Uspensky 2002).

proxy yields similar pricing performance but attributes a substantially higher amount of variation in risk premia to market risk. The main mechanism behind this finding is straightforward: the presence of the IFF leads to quantitatively large changes in estimates of market risk resposures (i.e., βs) measured using the VMF relative to the DMF. We call the difference between the two beta estimates the “beta gap” and show analytically why the beta gap is expected to be more severe for small stocks relative to large stocks.

We find considerable differences between market risk exposures (i.e., βs) estimated using the VMF and DMF market factors. Per our estimates, it is not uncommon for portfolios of mostly small stocks to have market βs that are biased downwards by as much as 50%. Relatedly, we find that βs estimated using the VMF have a lower cross-sectional standard deviation than those estimated using the DMF. The cross-sectional standard deviation of the beta gap is large; at 0.13, it is roughly 2/3 as large as the cross-sectional standard deviation of VMF betas (0.19). Empirically, and consistent with our theory, average size of stocks in a portfolio is a very strong predictor of an attenuation bias in VMF betas relative to DMF betas.

Ex ante, increasing the dispersion in beta could worsen the ability of market risk exposures to explain risk premia. However, we find this not to be the case. Not only does the DMF yield a larger spread in βs, it also generates a higher market risk premium estimate (5.8%) compared to that using the VMF (4.9%). The combination of both facts (modestly higher price of risk and substantially higher dispersion in quantity of risk β) implies that a larger fraction of average returns in the cross section can be explained by exposure to market risk as captured by our DMF factor compared with using the VMF. In addition to holding across all test portfolios we investigate, similar results obtain uniformly across a broad set of characteristics-sorted portfolios. We also find similar results regardless of whether we control for other PC-based pricing factors, workhorse characteristics-based factors, or both.

When the share of wealth in the stock market is relatively small compared to other assets, our theoretical model implies the IFF has a relatively small risk price. We test this implication in the data by running bivariate cross-sectional regressions that include separately estimated loadings on both the DMF and the IFF in the second stage. We show the IFF either carries a negative risk premium or one that is indistinguishable from zero depending on the specification. In other words, its risk premium is drastically different than that on the market factor, providing direct evidence to our claim that the market prices the IFF differently than the DMF.

In our fifth main contribution, we further investigate the importance of the IFF and the beta gap for assessing the size anomaly. In particular, when using the VMF to estimate market risk for size-sorted portfolios, our theoretical results imply small-cap portfolio βs are biased downwards making these assets appear less risky than they actually are. Although there are many ways to document the importance of this effect, we choose to explore one particularly poignant one related to prominent factor models in the empirical asset pricing literature.

Following insights from Banz (1981) and Fama and French (1992, 1993), the size anomaly has received much attention and most factor models contain factors related to market capitalization (i.e., size factors). First, using workhorse empirical factor models, pricing errors on a large set of standard portfolios are similar if we drop size factors when we use the DMF as the market proxy. These pricing errors are substantially larger when we instead use the VMF, suggesting that size factors help to address the econometric biases we identify. Second, when we drop size factors from these factor models and use the DMF as the market factor, the pared-down models can explain size factor average returns (i.e., the size factors are spanned by these modified models). This is not the case when the VMF is used as the market factor.

In other spanning tests, the IFF has large negative alphas when we use the VMF as the market factor and a strong negative loading on size factors, consistent with the IFF being unpriced and the negative alpha being generated by the negative size factor loadings. Furthermore, these IFF alphas increase in magnitude when we omit size factors from the models. In contrast, the IFF is approximately spanned when we use the DMF, which is the case regardless of whether we include or omit size factors from the models and is consistent with the IFF being unpriced. These results provide strong evidence that the existence of the size premium can largely be explained by mis-measured market risk of small stocks, which is in line with our theory and other empirical findings.

Related Literature: Our results relate to two literatures – both are quite mature so we will not survey them here – that have found mixed results when testing for an intertemporal risk-return tradeoff as well as a relationship between market risk and expected returns in the cross-section. Many papers in the former literature have focused on various approaches to better estimate conditional variance, and results have been found to be sensitive to these modeling choices.[4] In contrast to these implementation issues, our critique relates to a fundamental source of misspecification and applies even when a researcher has access to a "perfect" measure of conditional volatility of the VMF. Accordingly, we find that adjusting the VMF to purge it of the IFF leads to much stronger results.

In the empirical cross-sectional asset pricing literature (see, e.g., Nagel, 2013, for a survey), the potential challenges associated with a mis-measured market proxy date back to Roll's (1977) critique, namely that the correct market proxy is unobservable and tests of the univariate CAPM may fail without it.[5] The apparent failure of the CAPM has spawned a plethora of research. One tack has been to create more accurate constructions of the investor wealth portfolio, an approach that has had mixed results (see, e.g., Stambaugh, 1982; Jagannathan and Wang, 1996; Cederburg and O'Doherty, 2019). In a related paper, Kothari et al. (1995) argue that using an annual horizon for measuring market betas yields more accurate estimates, and show that these betas imply an economically and statistically significant risk-return tradeoff in the cross-section whether one uses the value- or equal-weighted market portfolio as a proxy for the investor wealth portfolio.

Another tack has been to accept that a single-factor model cannot explain the cross-section of returns and to try to identify other pricing factors related to anomalies (see, e.g., Fama and French, 1993, and the many papers that followed). Rather than try to identify specific anomaly-related factors to resolve this issue, Giglio and Xiu (2021) propose a method that uses principal components analysis to consistently

[4]Using GARCH-type volatility forecasting models, for example, Bollerslev et al. (1988), Chou (1988), Baillie and DeGennaro (1990), Campbell and Hentschel (1992), and Chou et al. (1992) find a positive risk-return relation, whereas Nelson (1991) and Glosten et al. (1993) find a negative relation when using extended GARCH models with so-called leverage effects in forming volatility forecasts. Focusing on the long-run relation between the two, Lundblad (2007) provides a positive coefficient estimate from a sample spanning from 1836 to 2003. However, using a stochastic volatility (SV) model, Koopman and Hol Uspensky (2002) and Brandt and Kang (2004) provide empirical evidence for a negative relation from stock indices in the United Kingdom, the United States and Japan (former) and from CRSP value-weighted index (latter). Using alternative approaches to GARCH or SV volatility forecasting models, French et al. (1987), Harvey (1989), Ludvigson and Ng (2007) and Ghysels et al. (2005) find a positive relation, while Pagan and Hong (1991) and Whitelaw (1994) find a negative relation between excess returns of stock index and its variance. Hedegaard and Hodrick (2016) find that using multiple sets of nonoverlapping returns improves precision of estimates and find evidence consistent with a positive risk return tradeoff. Finally,Goyal and Santa-Clara (2003) find that, although the conditional market variance is not a reliable forecaster of market returns, an average of the cross-sectional variance of individual stocks is.

[5]See also Malevergne et al. (2009) for a related discussion.

estimate risk premia in the cross-section, including the market risk premium. Lewis and Santosh (2021) take an alternative tack to resolve the weak risk return tradeoff. They argue that using portfolios that better reflect the marginal investor's portfolio can help resolve the weak empirical risk-return tradeoff when using the VMF.

Relatedly, multiple recent papers emphasize machine-learning or other data-driven approaches to identifying common factors in expected and realized returns (see, e.g., Kelly et al., 2019; Freyberger et al., 2020; Lettau and Pelger, 2020a,b; Harvey and Liu, 2021; Gu et al., 2020; Kozak et al., 2020; Bryzgalova et al., 2022), where a common thread is that many of the pricing factors uncovered by these procedures are not value-weighted portfolios. Our paper also relates to a broader literature on misspecification in linear factor models in asset pricing, such as papers studying effects of spurious/weak factors (see, e.g., Kan and Zhang, 1999; Kleibergen, 2009; Bryzgalova, 2015; Giglio et al., 2021), due to the disconnect between characteristics-sorted factors and the asset covariance structure (Daniel et al., 2020), or due to the way in which factors are constructed (Cremers et al., 2013). We contribute to this line of research by identifying two specific sources of variation in the VMF that likely cause a disconnect between it and variation in the overall investor wealth portfolio, provide evidence for this, and provide evidence that our adjusted market factor (the DMF) performs better in standard tests of the risk-return tradeoff than the VMF.

Finally, the size anomaly has played a prominent role in the cross-sectional asset pricing literature. There is still an active debate on whether the anomaly exists at all, and, if so, what sources of risk it captures.[6] However, as Berk (1995) points out, because size is measured using market capitalization and (all else equal) market capitalization is related to a firm's discount rate, the size anomaly may simply capture any component of discount rates that are not captured by a particular misspecified model. Our results are consistent with this line of reasoning as we propose one specific form of misspecification in the market factor. In line with implications from our analysis, once we correct for this misspecification, the size factors across all models

[6]Not only have there been periods when the size anomaly has yielded large average returns for investors, but the existence of this anomaly can have significant implications for allocational efficiency in the real economy (van Binsbergen and Opp, 2019). There is both an empirical debate on the strength of the size anomaly (for a summary of and contribution to this debate, see Asness et al., 2018), and also a theoretical debate on what sorts of risks the size premium may capture. For instance, small-cap stocks are relatively illiquid (Amihud and Mendelson, 1986) and Acharya and Pedersen (2005) argue that a stock's discount rate is linked to both its own liquidity as well as the covariance between its own return and market liquidity.

we study are spanned by the other factors in the models.

The remainder of the paper is organized as follows. Section 2 presents some preliminary details related to how we identify the IFF and the DMF. Section 3 summarizes our stylized facts related to the value-weighted market and the IFF. Section 4 provides a motivating theoretical model in which the IFF is priced differently than market risk. In sections 5 and 6, we develop analytical results to characterize the econometric biases introduced by the IFF then propose and implement tests to correct them in the time series and cross section, respectively. Section 7 presents some results related to controlling for factors from prominent factor models, and section 8 concludes.

2 Preliminaries

In this section, we begin by describing a simple linear asset pricing framework that serves as an organizing structure for our theoretical and empirical results. Next, we formally define the DMF and the IFF. Finally, we describe our empirical strategy for estimating these quantities and the data we use for this purpose.

2.1 Maintained assumptions

We begin by defining the VMF. Let $r_{i,t+1}$ denote the monthly *excess* return on individual stock i from month t to $t+1$. The VMF is a value-weighted average of individual excess returns,

$$VMF_{t+1} = \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}, \tag{1}$$

where $w_{i,t}$ is the predetermined value-weight of a stock i's return based on its market capitalization and N_t is a number of individual stocks in the stock index return in period t. We will show how the VMF may deviate from the DMF due to the existence of large firms whose idiosyncratic shocks may not be fully diversified in the index, and whose time-varying composition may lead to time varying loading on additional factors that are not of primary interest.

Next, we make the standard assumption that shocks to expected excess returns on each asset, i, is described by a linear factor model

$$\widetilde{r}_{i,t+1} = \beta_i \cdot \widetilde{f}_{t+1} + \zeta_i' \cdot \widetilde{g}_{t+1} + \eta_{i,t+1}, \tag{2}$$

where we use the tilde notation to denote deviations in excess returns from expected values (e.g., $\tilde{r}_{i,t+1} \equiv r_{i,t+1} - \mathbb{E}_t[r_{i,t+1}]$) to emphasize the fact that (at this point) we do not take a stand on the model of expected returns,[7] f_{t+1} is a scalar random variable that represents the DMF, β_i is asset i's exposure to f_{t+1}, g_{t+1} is a vector of additional common factors where $E(g_{t+1}|f_{t+1}) = 0$, ζ_i is asset i's exposure to g_{t+1}, and $\eta_{i,t+1}$ is an idiosyncratic shock such that $\mathbb{E}[\eta_{i,t+1}|f_{t+1}, g_{t+1}] = 0$ and has finite variance.[8]

2.2 Defining DMF and IFF

Under the strong assumptions that yield the Capital Asset Pricing Model, a value-weighted index of all risky assets is the tangency portfolio. Owing to this intellectual history, it is customary to define f_{t+1} in equation (2) as the VMF and use this as a level factor in essentially all factor models. In the next section, we highlight five empirical facts that motivate the perspective that the VMF may not itself be an ideal factor for this purpose.

With this perspective in mind, we define f_{t+1} via the following rotation/normalization restrictions:

$$\mathbb{E}\left[\sum_{i=1}^{N_t} w_{i,t}\beta_i\right] = 1, \text{ and} \tag{3}$$

$$\mathbb{E}\left[\frac{1}{N_t}\sum_{i=1}^{N_t} \zeta_i\right] = 0. \tag{4}$$

In other words, the scale of f_{t+1} is defined so that the VMF has an unconditional average loading of 1 on it, and the remaining factors (g_{t+1}) are rotated so that the average stock has zero loading on these orthogonal factors. Therefore, f_{t+1} captures the common factor exposure of the average stock and g_{t+1} capture tilts away from f_{t+1}.

If we have a balanced panel of asset returns ($N_t = N$ for all t) with constant factor

[7]To do so, we would also have to make assumptions about the stochastic discount factor.

[8]This is also a static factor model (i.e., constant coefficients), which is standard in the literature when working with test assets that are portfolios as will be the case in our empirical work. This framework has been a workhorse for empirical asset pricing studies and encompasses more specialized models whether they be based on theory such as the CAPM (Sharpe, 1964; Lintner, 1965; and Black, 1972) or statistical arguments such as the APT (Ross, 1976). Furthermore, it encompasses the characteristics-based factor models that have come to dominate the empirical cross-sectional asset pricing literature (e.g., Fama and French, 1993, among the many others that followed).

loadings, then returns on an equal-weighted portfolio of assets satisfy[9]

$$\widetilde{r}_{ew,t+1} = \frac{1}{N}\sum_{i=1}^{N}\left(\beta_i \cdot \widetilde{f}_{t+1} + \zeta_i' \cdot \widetilde{g}_{t+1} + \eta_{i,t+1}\right) = \beta_{ew} \cdot \widetilde{f}_{t+1} + 0 \cdot \widetilde{g}_{t+1} + \frac{1}{N}\sum_{i=1}^{N}\eta_{i,t+1}, \quad (5)$$

where $\beta_{ew} \equiv \frac{1}{N}\sum_{i=1}^{N}\beta_i$. Because $E_t[\eta_{i,t+1}] = 0$, we know that $\widetilde{r}_{ew,t+1}$ will be an *unbiased* estimator of $\widetilde{f}_{t+1}$. Since the number of stocks N is large, under some regularity conditions on the idiosyncratic shocks $\{\eta_{i,t+1}\}_{i=1}^{N}$, the law of large numbers may hold almost exactly and $\widetilde{r}_{ew,t+1}$ is also *consistent* for $\widetilde{f}_{t+1}$ up to a constant of proportionality, i.e.,

$$\operatorname*{plim}_{N\to\infty} \widetilde{r}_{ew,t+1} = \beta_{ew}\widetilde{f}_{t+1} \iff \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\eta_{i,t+1} = 0, \quad (6)$$

in which case $\widetilde{r}_{ew,t+1}$ essentially reveals (a scaled version of) $\widetilde{f}_{t+1}$. With the additional assumption of no arbitrage (which we impose later to pin down expressions for risk premia given the assumed factor structure), $r_{ew,t+1}$ is also a scaled version of f_{t+1}.

If we perform a similar aggregation exercise for the VMF itself, which is a linear combination of individual stocks with loadings $\beta_{m,t} \equiv \sum_{i=1}^{N_t} w_{it}\beta_i$ and $\zeta_{m,t} \equiv \sum_{i=1}^{N_t} w_{it}\zeta_i$ on f_{t+1} and g_{t+1}, respectively, we find that

$$\begin{aligned}
\widetilde{VMF}_{t+1} &= \beta_{m,t} \cdot \widetilde{f}_{t+1} + \zeta_{m,t}' \, \widetilde{g}_{t+1} + \sum_{i=1}^{N_t} w_{it}\eta_{i,t+1} \\
&\equiv \widetilde{f}_{t+1} + \underbrace{(\beta_{m,t} - 1) \cdot \widetilde{f}_{t+1}}_{\substack{\text{time-varying exposure} \\ \text{to true market factor } \widetilde{f}_{t+1}}} + \underbrace{\zeta_{m,t}' \, \widetilde{g}_{t+1}}_{\substack{\text{(potentially) time-varying} \\ \text{exposure to other factors } \widetilde{g}_{t+1}}} + \underbrace{\eta_{t+1}}_{\substack{\text{granular} \\ \text{residual}}} \\
&\equiv \widetilde{f}_{t+1} + \underbrace{\widetilde{\epsilon}_{t+1}}_{\substack{\text{(innovation in) idiosyncratic} \\ \text{financial risk factor}}} .
\end{aligned} \quad (7)$$

We define $\epsilon_{t+1} \equiv VMF_{t+1} - f_{t+1}$ as the idiosyncratic financial risk factor ("IFF"), and it is the sum of three distinct components. The first captures the fact that the current set of large stocks may have conditional loadings on f_{t+1} that are above or below the time series average loading of 1. The second captures a value weighted average of conditional loadings on other orthogonal factors g_{t+1} (a deviation relative

[9]This follows from our assumptions that $E[\sum_{i=1}^{N} w_{i,t}\beta_i] = 1$ and $\frac{1}{N}\sum_{i=1}^{N}\zeta_i = 0$ and $N_t = N$.

to the equal-weighted average ζ_i that is normalized to zero). As the type of economic activities conducted by the market's largest firms changes considerably over time, these factor exposures likely vary over time.

Following Gabaix (2011), we will refer to the third component, η_{t+1}, as the "granular residual", as it captures the contribution to the VMF from idiosyncratic shocks that are not fully diversified away. Whether or not the conditional variance of η_{t+1} collapses to zero depends on the extent to which the law of large numbers applies, and some necessary conditions are easy to test.[10] As we demonstrate and discuss further below, this condition likely does not hold for the CRSP universe of stocks in the US, which is the dominant data source for most empirical asset pricing tests.

2.3 Identifying DMF and IFF empirically

We have now have defined the DMF and the IFF given our our linear asset pricing framework in equation (2). To identify the DMF and the IFF empirically, we follow standard practice in the literature and perform a PCA on excess returns on a broad set of standard equal-weighted characteristics-sorted portfolios so that each PC is a linear combination of the base portfolio excess returns.[11] We then use the resulting PCs as (rotated) proxies for f_{t+1} and g_{t+1} in equation (2).[12] Using equal-weighted portfolios ensures the PCs are orthogonal to the granular residual, η_{t+1}.[13] Furthermore, the first PC explains 85% of the variation in the equal-weighted portfolio returns and 80% of the variation in value-weighted portfolio returns.[14] We obtain all portfolio data used in our PCA from Kenneth French's Data Library and provide a summary of the portfolios in Online Appendix Table OA.1. The table reports a standard set of 372 portfolios having data from 1963-2021 that we use throughout the paper.[15] We

[10]For example, if $\underline{\kappa} < Var[\eta_{i,t+1}] < \overline{\kappa} < \infty$ for all stocks and $\eta_{i,t+1} \perp \eta_{j,t+1}$ for all $i \neq j$, a necessary condition for $\eta_{t+1} \xrightarrow{p} 0$ is $\lim_{N_t \to \infty} \sum_{i=1}^{N_t} w_{i,t}^2 = 0$.

[11]Although the PCA is performed using demeaned excess returns, we apply the PCA weighting matrix directly to portfolio excess returns to construct the PCs used in our tests so that these represent zero net cost (ex-post) tradable portfolios that have non-zero average excess returns.

[12]See Online Appendix section OA.3 for more motivation behind using PCs for this purpose.

[13]This is true as long as the equal-weighted anomaly portfolios are well-diversified such that the law of large numbers holds. In this case, asset granular residuals, $\eta_{i,t+1}$, will be diversified away as in the case of $r_{ew,t+1}$ in equation (5). Because PCs are a linear combination of these equal-weighted portfolios, the PCs themselves will also be orthogonal to $\eta_{i,t+1}$.

[14]See Online Appendix subsection OA.3.2.

[15]We choose these portfolios for our PCA to balance a few tradeoffs that we discuss further in Online Appendix section OA.3.

define the DMF to be the projection of the VMF onto the first PC, which ensures that the DMF satisfies our restriction in equation (3).[16] The IFF is then defined to be their difference, $IFF \equiv VMF - DMF$, consistent with its definition in equation (7).

3 Motivating evidence

In this section, we present five motivating facts that support the existence of the IFF as we have defined it in the previous section. We then describe some of the time series properties of our estimated IFF.

3.1 Five motivating facts

The extent to which the IFF is quantitatively large is an empirical question. To support our discussion and empirical tests that follow, we provide some empirical evidence from the U.S. stock market that suggests this is likely the case.

3.1.1 Fact 1: The distribution of VMF weights is highly concentrated

There are at least two reasons the standard value-weighted market may not reflect the average stock, a desirable property for a level factor. First, the equity market is comprised of stocks with a fat-tailed size distribution that may cause the idiosyncratic component, $\eta_{i,t+1}$, in equation (2) to have a non-trivial influence on the value-weighted market return. Second, as we show in the next subsection, the industry composition changes drastically over time. This is particularly true among the largest stocks.

Figure 1 highlights market concentration among the largest stocks in the CRSP value-weighted index. The figure plots the fraction of market value associated with the largest 5, 10, 25, and 50 stocks, respectively.[17] Throughout the entire sample

[16] See Online Appendix section OA.3.3 for more details related to this projection.

[17] We also compute an alternative measure of concentration which allows the number of stocks in the bin to grow as the number of listed stocks increases. Specifically, we compute the share of stocks in the top 2.5th percentile of the market value, where this measure sums up weights for the largest stocks whose rank in the CRSP market value distribution is less than 2.5% times the number of stocks in the NYSE for each month. Here we use the count of stocks in the NYSE (shaded in red on the right axis) to avoid mechanical increases in our concentration measure due to a large number of (mostly small) AMEX and NASDAQ stocks added during the second half of our sample period. The concentration measure lacks a downward trend, hovering between 25 and 40 percent of the total value of the stock market over the 1926-2021 sample period.

period from 1926 to 2021, the largest five stocks are associated with between 10 and 25 percent of overall market value. The next 20 stocks are also associated with a substantial fraction of market value, so that the combined weight of the top 25 stocks ranges from a high of almost 50% early in the sample to a low of about 20% during the 1980s. As of 2021, this weight is about 35%. Stocks 26-50 constitute another nontrivial percentage of total market value with similar overall time series variation as the top 25 stocks.

These shares decline somewhat over the sample in part because the CRSP universe expands, as is clear from the number of constituent stocks plotted on the right axis. Nonetheless, the market is highly concentrated throughout the sample, and it is striking that one can barely detect any changes in these weights even when the number of stocks included in the index jumps dramatically upward.

Given this concentration, it is likely the value-weighted market is not well diversified with respect to the idiosyncratic components of individual stock returns represented by $\eta_{i,t+1}$ in equation (2). Gabaix (2011) shows how such "granular residuals" likely do not diversify when firms have a fat-tailed size distribution, in which case idiosyncratic shocks affect aggregate quantities. To the extent the aggregated granular residual (i.e., the value-weighted sum of $\eta_{i,t+1}$'s, which we refer to as η_t) carries a different price of risk than the DMF, it can cause issues for standard asset pricing inference when using the VMF, which we discuss in subsequent sections below. We also provide some motivation for why one might expect the granular residual to be priced differently than the DMF in a theoretical model that we develop below.

3.1.2 Fact 2: VMF industry composition varies drastically over time

Next, we show that the industry composition of the VMF fluctuates drastically over time. To the extent different industries load differently on the factors in equation (2) this could lead to time-varying loadings of the VMF on these factors. Furthermore, the industry composition of the largest stocks changes even more drastically over time, implying that the idiosyncratic shocks that constitute the aggregated granular residual, η_{t+1}, change over time as well.

For each of the Fama-French 10 industries, Figure 2 plots the fraction of their of market capitalizations among the top 2.5% stocks by size (Panel 1), among all other stocks (Panel 2), and finally among all stocks in the CRSP universe (Panel 3).[18]

[18]These threshold rules typically yield between 20-35 firms included in the top 2.5% of stocks,

Comparing Panels 1 and 2, we note that the industry composition of the "megacap" group exhibits much more pronounced time series fluctuations relative to the set of smaller firms and its industry distribution is quite unrepresentative of the rest of the market. For instance, firms in the energy and durables categories are associated with almost half of the market value in the top group over the first two thirds of the sample (see Panel 1), whereas the two categories combined make up less than 20% of the sample of smaller firms throughout the entire sample period (see Panel 2). Likewise, tech firms are somewhat overrepresented, especially later in the sample where they rise to comprising approximately 60% of the total megacap composition in Panel 1 in 2021, whereas this industry makes up less than 20% of the sample of smaller firms in the same sample period (see Panel 2). Panel 3 shows that the overall VMF industry composition is somewhere between Panels 1 and 2, inheriting some of the spikes in concentration from Panel 1. Furthermore, the large overall fluctuations in the VMF's industry composition may lead to the value-weighted market loading differently on factors over time, for which we show empirical evidence in subsection 2.2.

3.1.3 Fact 3: The VMF is priced by the first PC

Table 1 provides results from regressions of the CRSP value-weighted (Panel A) and equal-weighted (Panel B) market index excess returns on various numbers of PCs from the equal-weighted portfolio PCA. Vriation in the value-weighted (equal-weighted) market return is well-described by its projection on the first PC with an R-squared value of 0.84 (0.95), which increases to 0.95 (0.99) when three PCs are included in the regression. The first PC also does a good job explaining average returns on both the value- and equal-weighted indexes, with each having statistically and economically insignificant annualized alphas of -0.53% and -0.46%, respectively. These alphas can be compared to average annualized excess returns of 6.79% and 9.36%, respectively. In other words, the first PC approximately prices both the VMF and the equal-weighted market.

Furthermore, adding more PCs to the regressions does not change the alphas appreciably and only marginally increases the R-squared values. If anything, the decomposition of the VMF risk premium suggests that exposures to the second and

though the count is somewhat higher in the 1990s (peaking at 45) and is closer to 15 prior to 1940. Results are similar for other cutoffs.

third PCs – the main factors which add sizable explanatory power for the VMF beyond the first PC – contribute essentially zero (a statistically insignificant 17 bp) and -1.6% per annum to VMF expected returns in the time series, respectively.

3.1.4 Fact 4: The VMF has a stable loading on the first PC but time-varying loadings on higher PCs

To the extent the market composition changes over time through time variation in, for instance, its industry composition, it may have time-varying loadings on factors that explain the cross-section of returns similar to the assumed time-varying loadings, $\zeta_{m,t}$, in equation (2). To explore this possibility, we plot coefficient estimates from rolling regressions of the value-weighted (equal-weighted) market excess return on the first five PCs constructed using equal-weighted portfolios in Figure 3, Panel 1 (Panel 2). Coefficients on the first PC are relatively stable in both cases. However, both the value- and equal-weighted market generate time-varying loadings on higher PCs, which is in line with the notion that compositional changes in the market may result in it having time-varying loadings on some factors.

Taken together, results in Table 1 and Figure 3 show that the first PC is able to explain variation in the equal-weighted market excess return well, the market loading on this first PC is fairly stable over time, and the market loadings on higher PCs are time varying. This helps justify our definition of the DMF in subsection 2.3 as a projection of the VMF onto the first PC.

3.1.5 Fact 5: The DMF is strongly linked to macroeconomic conditions but a projection of the VMF onto higher PCs is not

In this subsection, we examine whether different components of the VMF have differential predictive power for macroeconomic aggregates by estimating impulse response functions (IRFs) via local projections. In particular, we project the VMF onto the DMF and up to the first 10 PCs and then estimate IRFs with respect to the DMF and projections of the VMF onto the higher PCs for nine macroeconomic aggregates.[19,20] The IFF includes several components, but here we emphasize only

[19]Our choice of macroeconomic aggregates are log industrial production (IP) growth, log employment growth, initial claims divided by lagged employment, log per capita consumption growth, log compensation growth, the ADS business conditions index (Aruoba et al., 2009), the unemployment rate, log GDP growth, and the Chicago Fed National Activity Index (CFNAI).

[20]Additional details on the IRF estimation can be found in Online Appendix OA.4. One point to note is that we must use a quarterly frequency to conform with the availability of the macroeconomic

its projection onto the first two higher PCs. We do so because this is the key term in the IFF which generates sizable differences in estimated market risk exposures.[21] We find the DMF reliably predicts economic activity, which is not the case for the projection of the IFF onto higher PCs.

Figure 4 shows the IRF estimates for projections onto the first three PCs over the 1963-2019 sample.[22] The blue line and blue confidence bands correspond to using the DMF. The red ones correspond to using the projection of the VMF onto both the second and third PCs, which is equivalent to projecting the IFF onto these PCs. Comparing the sets of blue and red series, it is clear that the DMF is more informative about future macroeconomic aggregates than the projected VMF, as the former typically yields statistically significant impulse response coefficients in the anticipated direction. For instance, a one-standard-deviation increase in the DMF generates a statistically significant 0.5% (0.7%) increase in consumption (GDP) after 10 quarters. On the other hand, impulse responses to the VMF projection are typically statistically insignificant and often go in a direction consistent with the factor hedging increases in a representative investor's marginal utility. For instance, a one-standard-deviation increase in the this component of the VMF generates 0.5% *decreases* in consumption and GDP after 10 quarters, albeit in a (typically) statistically insignificant manner. While this evidence is only suggestive, it provides motivation that, while the DMF is likely to be tightly linked with investors' marginal value of wealth, such a link is tenuous for the projection of the VMF onto higher PCs.

aggregates, so we aggregate the monthly DMF and VMF projections within each quarter for this analysis.

[21] We show that a projection of the VMF onto PCs yields market risk measurements that are similar to those based on the VMF itself in Online Appendix Figure OA.5, which also implies the granular residual component of the VMF is small (consistent with results in Online Appendix Table OA.2) and does not impact the estimate of market risk much.

[22] We omit data from the COVID pandemic because of the outliers in many of the macroeconomic time series of interest. However, we provide results using the full 1963-2021 sample in Online Appendix Figure OA.1. We also provide results when projecting the VMF onto up to 10 PCs in Online Appendix Figure OA.2 and when the projections are done in a rolling fashion in Online Appendix Figure OA.3. Finally, we provide results including those for an estimate of the GR (the residual from the regression) obtained when projecting the VMF onto five PCs when projections are done in a rolling fashion in Online Appendix Figure OA.4. These results are quantitatively and qualitatively similar to those presented in the main text when projecting the VMF onto the DMF and the first two higher PCs.

3.2 Time series properties of the IFF

In this subsection, we describe some of the basic properties of our estimated DMF and the associated IFF. Our main specification constructs the DMF as a projection of the VMF onto the first principal component as described in Subsection 2.2 using data from 1963-2021. We also construct an alternative DMF by projecting the VMF on the CRSP equal-weighted market and use this in some analyses.[23] Given a proxy for the DMF, the associated IFF is constructed as VMF - DMF.

Table 2 reports summary statistics on the monthy excess return time series for the VMF, the DMF, and the IFF under various specifications.[24] We also report R-squared values from regressions of the VMF on the first PC and the equal-weighted market. Panels A, B, and C present results using data from the 1963-2021, 1963-1992, and 1993-2021 periods, respectively.[25] Our estimates indicate the IFF is negatively skewed and has a large kurtosis relative to the normal distribution. Furthermore, the DMF has a lower volatility than the VMF, is more positively skewed, and has relatively higher kurtosis. Results in the table indicate the ratio of IFF standard deviation to VMF standard deviation is 39.6%, implying that the "size" of the IFF is moderate compared to the VMF.[26]

4 Theoretical framework

In this section, we develop a simple theoretical model to characterize risk premia on DMF and VMF, respectively. In the first subsection, we develop a stylized model which nests the CAPM to motivate the assumption that components of the IFF have different risk prices than the DMF (and, in particular, why η_{t+1} likely carries a relatively small risk price) using a simple theoretical model. In the next subsection, we impose structure on the stochastic discount factor that allows us to derive expressions

[23] We use the CRSP equal-weighted market in place of the first PC because the first PC is well-known to be a level factor that places similar weights on all assets (like an equal-weighted index), and because the equal-weighted index is not likely to contain a sizable granular residual.

[24] We provide kernel density estimates of the VMF and the IFF in Online Appendix Figure OA.6 (Panel 1) for visualization.

[25] We choose 1992 as a cutoff for the first sub-period because it corresponds to the year before the publication of Fama and French (1993), which corresponds to the beginning of a period of proliferation for factor models and anomaly identification.

[26] We estimate the "size" of the GR in Online Appendix OA.2 and find the ratio of its standard deviation to VMF standard deviation is approximately 1-3% depending on the specification.

for the risk premium on a generic asset, i, as well as those on the DMF and the VMF based on our analytical framework described by equation (2).

4.1 A model in which IFF is priced differently than DMF

In this section, we outline some key results from a static theoretical model which explicitly characterizes risk prices on the DMF and the IFF, relegating additional details to Online Appendix OA.5. Our model features heterogeneous agents who hold both tradable financial wealth and an additional source of nontradable income. While this latter component captures many sources of additional resources available for consumption in practice, we will refer to it as labor income to fix ideas.

We assume that realized returns in the financial sector are governed by two orthogonal factors f and g similar to equation (2), except we omit time subscripts and make g is a scalar for simplicity.[27] Investors can trade two types of stocks in the financial sector: K large stocks and $N - K$ small stocks, where $N \gg K$. We assume that expected dividends are such that each large stock comprises $\frac{\bar{v}}{K}$ percent of the total stock market, where the analogous weight on each small stock is $\frac{1-\bar{v}}{N-K}$. Recall from section 3.1.1 that a very small number of large stocks consistently comprise a substantial fraction of market value, a share which barely changes even when thousands of smaller stocks are added to the CRSP universe. We focus on the limit in which K is fixed and $N \to \infty$. Thus, idiosyncratic risks to small stocks are perfectly diversified in VMF but a granular residual emerges from large stocks.

Loadings on both factors (β_i and ζ_i) are heterogeneous across stocks, and we separately define the factors as follows. f is a level factor normalized such that the value-weighted market has a loading of 1 on it. g is an additional systematic factor which, consistent with our definitions above, is rotated such that the average small stock loading on g equals zero. Large and small stock returns feature independent idiosyncratic error terms η_i that cannot be explained by the two factors, and aggregate to η in the value-weighted market.

The model implies a Mayers (1973)-like formula for expected excess returns,

$$\mathbb{E}[r_i] = RRA \cdot [(1-\phi)Cov(r_i, r_m)) + \phi Cov(r_i, r_h)]\,, \tag{8}$$

[27] Here, we keep g as a scalar for expositional simplicity, as our objective is to price the VMF's exposure to systematic factors orthogonal to DMF. It is straightforward to extend the model to allow for multiple factors. Also note that our static model abstracts away from hedging demands which would arise in a dynamic setting.

where RRA is relative risk aversion and ϕ captures the expected share of future consumption coming from nontraded income, which we will refer to as the human wealth share.[28] The VMF return, r_m, is given by

$$r_m = \mathbb{E}[r_m] + f - \mathbb{E}[f] + \zeta_m \cdot (g - \mathbb{E}[g]) + \eta \equiv f + \epsilon, \tag{9}$$

where $\zeta_m \equiv \frac{\bar{v}}{K}\sum_{i=1}^{K}\zeta_i$ captures the market's exposure to g and $\eta \equiv \frac{\bar{v}}{K}\sum_{i=1}^{K}\eta_i$ is the granular residual, coming solely from large stocks (since the law of large numbers holds for the small stock idiosyncratic component). In this model (and as in the data), r_m reflects the DMF (f) but also includes the IFF (ϵ), which varies due to large stocks' nonzero exposure to g and the granular residual η. The second priced factor, r_h, captures the innovation in labor income. Consistent with evidence in section 3.1.5 that the projection of VMF onto higher PCs did not predict labor income or other macroeconomic variables, we will suppose here that labor does not load on ϵ, in which case r_h is given by

$$r_h = \psi_f(f - \mathbb{E}[f]) + \psi_\xi \xi \tag{10}$$

where the ψ's are model constants and ξ is an aggregate labor return shock that is assumed to be uncorrelated with the financial sector.

Turning to equation (8), when there are no nontraded assets ($\phi = 0$), our model nests the standard CAPM, the stochastic discount factor is linear in r_m, and fluctuations coming from g or η shocks are priced identically to f shocks. However, when non-traded assets like human capital and private equity are the dominant source of current wealth and future consumption ($\phi \approx 1$), the link between the stochastic discount factor and the value-weighted index becomes very weak when the nontraded assets do not load on η. In other words, the volatility of fluctuations in wealth coming from the IFF becomes negligible. To see this formally, we characterize risk premia.

[28] Further, in Online Appendix OA.5, we also extend the model to allow for state-dependent non-diversifiable income risk in a manner similar to Schmidt (2016), another force which can increase the price of risk on labor income shocks. Conceptually, this amounts to reinterpreting r_h as capturing the innovation in average labor income plus a certainty equivalent capturing an individual's willingness to pay to avoid idiosyncratic income shocks.

When $N \to \infty$ and $Cov(r_h, \epsilon) = 0$,

$$\mathbb{E}[r_i] = \begin{cases} RRA \cdot \left[\underbrace{(1-\phi+\phi\psi_f)}_{f \text{ risk price}} \beta_i \sigma_f^2 + \underbrace{\zeta_m (1-\phi)}_{g \text{ risk price}} \zeta_i \sigma_g^2 + \underbrace{(1-\phi)}_{\eta_i \text{ risk price}} \frac{\overline{v}}{K} \sigma_i^2 \right] & \text{for } i = 1, \ldots, K \\ RRA \cdot \left[\beta_i (1-\phi+\phi\psi_f) \sigma_f^2 + \zeta_i (1-\phi) \zeta_m \sigma_g^2 \right], & \text{for } i > K, \end{cases} \tag{11}$$

so each stock's risk premium depends on its covariances with the systematic factors, as well as the granular residual η (reflecting the fact that some large stocks' idiosyncratic returns directly contribution to aggregate wealth fluctuations). The VMF risk premium is given by

$$\mathbb{E}[r_m] = \underbrace{RRA \cdot (1-\phi+\phi\psi_f) \sigma_f^2}_{\text{DMF risk premium}=E[f]} + \underbrace{RRA \left[\zeta_m^2 (1-\phi) \ \sigma_g^2 + (1-\phi) \sigma_\eta^2 \right]}_{\text{IFF risk premium}=E[\epsilon]}, \tag{12}$$

where ϕ is the labor share of wealth. So, realized market returns are exposed to f, g, and η, and the market risk premium contains components related to compensation for exposure to each of these terms. As the labor share of wealth gets large (i.e., $\phi \to 1$), only covariances with labor income drive risk premia. In that case $E[f] = RRA \cdot \psi_f \sigma_f^2$ and $E[r_i] = \beta_i E[f]$ for all stocks, including large stocks. Since $Cov(r_h, \epsilon) = 0$, the risk premium on IFF vanishes – i.e., $E[\epsilon] = 0$.

Maintaining the above assumptions and supposing $\phi > 0$, the model delivers three sets of testable predictions for which we find empirical support:

1. **Attenuation of small cap betas:** For a portfolio of small-cap stocks, its true beta with respect to f, $\beta_S \equiv 1/(N-K) \sum_{i=K+1}^{N} \beta_i$, is smaller than its beta with respect to the value-weighted market (as long as $\beta_S > 0$),

2. **Predictions about VMF alphas:** The alpha of a portfolio of small-cap stocks is positive with respect to a CAPM model that uses r_m as the market factor. An analogous CAPM alpha estimate is negative for the IFF.

3. **Predictions about DMF alphas:** When f is used rather than r_m as the market factor, alphas on the small stock portfolio are zero regardless of the value of ϕ and IFF alphas are zero in the limit as $\phi \to 1$.

As we discuss further in Online Appendix OA.7, the data suggest that the US stock market comprises a relatively small fraction of total wealth, even for high net worth investors who are particularly likely to participate in stock markets. This fact, in conjunction with our evidence in section 3.1.5 suggests that factors like g and, particularly, η likely have substantially lower risk prices and contribute little to risk premia, which is quite different than what would be the case in an economy entirely comprised of US stock market wealth (i.e., $\phi \to 0$). Importantly, even if g and η are priced in intermediate cases when $0 < \phi < 1$, the associated risk prices are distinct (and likely smaller in magnitude) from the level factor in the model, f, that drives both sectors of the economy and is the primary driver of changes in investor wealth.

4.2 Risk Premia on DMF and VMF in a Dynamic Model

In this section, we outline a set of assumptions we use to characterize risk premia in our empirical tests. We allow for three primary differences from the above theoretical framework: 1) we allow for time varying volatilities, 2) we allow g to be vector-valued, and 3) we characterize risk-prices via a reduced-form SDF rather than an explicit structural model.

We assume there are no arbitrage opportunities, which implies the existence of a stochastic discount factor (SDF), m_{t+1}, satisfying the law of one price: $\mathbb{E}_t\left[m_{t+1} r_{i,t+1}\right] = 0$ for $\forall i$ in the economy. We further assume the SDF takes the form

$$m_{t+1} - \mathbb{E}_t[m_{t+1}] = -\gamma \cdot (f_{t+1} - \mathbb{E}_t[f_{t+1}])\mathbb{E}_t[m_{t+1}] - \Lambda(g_{t+1} - \mathbb{E}_t g_{t+1})\mathbb{E}_t[m_{t+1}]. \quad (13)$$

In addition to the fairly standard assumption of constant risk prices, the other key restriction implied by equation (13) is that the granular residual, η_{t+1}, is not priced (as is the case in the model in section 4.1 when $\phi \approx 1$). Later in the paper (in section 6.2.3), we conduct overidentification tests for our assumptions about risk prices by formally including ϵ as an additional pricing factor. In most of these specifications, we fail to reject the null hypothesis that it is not priced. Moreover, when these estimates are borderline significant, the associated risk price is negative (the opposite sign from our large, positive, and significant estimate of γ).[29]

[29]In our model, this could be the case if, for example, shocks that improve the profits of the largest firms in the economy are associated with inefficiencies (e.g., due to negative externalities from weaker competition) which actually reduce investor welfare and thus are associated with a higher marginal value of wealth.

Given the SDF in equation (13), asset i's risk premium satisfies

$$\mathbb{E}_t \left[r_{i,t+1} \right] = \gamma \cdot \beta_i \cdot \sigma_{f,t}^2 + \Lambda' \Sigma_{g,t}' \Sigma_{g,t} \zeta_i, \tag{14}$$

where $\sigma_{f,t}^2$ and $\Sigma_{g,t}' \Sigma_{g,t}$ capture the conditional variance of f and variance-covariance matrix of the g factors, respectively. An asset's risk premium is the product of risk prices γ and Λ, risk quantities, and the asset's factor exposures β_i and ζ_i. The conditional risk premia on the DMF and the VMF are then[30]

$$\mathbb{E}_t[DMF_{t+1}] = \mathbb{E}_t[f_{t+1}] = \gamma \cdot \sigma_{f,t}^2 \text{ and} \tag{15}$$

$$\mathbb{E}_t \left[VMF_{t+1} \right] = \gamma \cdot \sigma_{f,t}^2 + \gamma \cdot (\beta_{m,t} - 1) \cdot \sigma_{f,t}^2 + \Lambda' \Sigma_{g,t}' \Sigma_{g,t} \zeta_{m,t}, \tag{16}$$

repectively. So, given our assumed factor structure and the assumption of no arbitrage, the evidence presented in subsection 3.1.4 regarding time-varying VMF loadings on g_{t+1} implies that DMF and VMF will command different risk premia.

5 Reassessing the risk-return relationship in the time series

In this section, we investigate how the existence of the IFF affects the risk-return tradeoff implied by the VMF and the DMF in the time series. We first characterize econometric biases introduced by an unpriced IFF when estimating the risk-return tradeoff using the VMF versus the DMF. We find evidence the IFF generates an attenuation bias that yields a relatively weak intertemporal risk-return relationship for the VMF, whereas the risk-return relationship is more robust for the DMF.

5.1 Econometric intuition

Suppose a researcher is trying to test the implications of equation (15), but only has access to the VMF (and not the DMF). The existence of the IFF in the VMF leads to upward-biased variance estimates and with potentially different time series dynamics relative to the DMF variance dynamics. To see this formally, notice that

[30] These expressions combine the expression in equation (14) with the facts that $DMF_{t+1} \equiv f_{t+1}$ and $\beta_{DMF} = 1$.

the difference between the VMF and DMF variances is given by:

$$Var_t[VMF_{t+1}] - \sigma^2_{f,t+1} = (\beta^2_{m,t} - 1)\sigma^2_{f,t+1} + \zeta'_{m,t}\Sigma'_{g,t+1}\Sigma_{g,t+1}\ \zeta_{m,t} + Var_t[\eta_{t+1}], \quad (17)$$

where we see the role of the three separate components: time-invariant exposures to g, time-varying exposures to all factors, and the granular residual. Other than the potentially ambiguous role of time-varying $\beta_{m,t}$ (which we argued in Figure 3 is likely to be small), the last two terms unambiguously increase the variance of the VMF to be higher than that of the VMF.

Next, we consider estimates of the risk-return tradeoff for the VMF. To focus on issues with the largest importance for our estimates, we make three additional assumptions here for expositional simplicity.[31] First, we assume that $\beta_{m,t} = 1$, which is approximately satisfied in the data. Second, consistent with the model in section 4.1 as $\phi \to 1$ and the low correlations between VMF's projection onto higher PCs and macroeconomic variables (Figure 4), we assume the IFF isn't priced – i.e., $E_t[\epsilon_{t+1}] = 0$. Third, even though conventional estimates of VMF variances are likely to be misspecified, let us abstract away from issues about estimating variances correctly and suppose that the researcher has access to a "perfect" conditional variance estimate for the VMF, $\sigma^2_{VMF,t}$.[32]

Under the above assumptions, if there is no IFF, we recover the true market price of risk γ via an OLS regression of VMF_{t+1} on a constant and $\sigma^2_{VMF,t}$. However, when the IFF is present, the OLS regression coefficient will be biased, most likely towards zero. As the simplest way to see this, the ratio of the true VMF risk premium ($\gamma\ \sigma^2_{f,t+1}$) to the VMF variance (given by equation 17) is smaller than γ due to the presence of additional unpriced sources of VMF risk. Since the regression includes a constant, one can still correctly recover the correct coefficient as long as the difference between VMF and DMF variances is constant – i.e., the IFF is homoskedastic. When the IFF changes the variance dynamics of VMF relative to DMF, for which we provide empirical evidence in subsection 5.2, OLS estimates of the slope coefficient on $\sigma^2_{m,t+1}$

[31]We can relax each assumption (which are not imposed in our estimation procedure), but doing so makes expressions much more complex without adding much intuition.

[32]Notice that the VMF variance is a nonlinear function of time-varying loadings (whose dynamics depend on shifts in the size distribution over time), multiple factor variance components, and the variance of the granular residual. Thus, conditional variance dynamics of VMF are unlikely to be well-described by a low-dimensional GARCH process even if the individual factor variance components do follow such dynamics.

will not be consistent for γ, instead converging to

$$\operatorname*{plim}_{T\to\infty}\widehat{\gamma} = \frac{Cov\left[\sigma^2_{VMF,t}, VMF_{t+1}\right]}{Var\left[\sigma^2_{VMF,t}\right]} = \gamma \cdot \frac{Var\left[\sigma^2_{f,t}\right] + Cov\left[\sigma^2_{f,t}, \sigma^2_{\epsilon,t}\right]}{Var\left[\sigma^2_{f,t}\right] + Var\left[\sigma^2_{\epsilon,t}\right] + 2 \cdot Cov\left[\sigma^2_{f,t}, \sigma^2_{\epsilon,t}\right]}, \quad (18)$$

where we use $\sigma^2_{f,t}$ ($\sigma^2_{\epsilon,t}$) to denote the conditional variance of the DMF (IFF) rather than $\sigma^2_{DMF,t}$ ($\sigma^2_{IFF,t}$) for brevity. The use of a mis-measured proxy for the DMF is associated with a downward multiplicative bias provided that $Var\left[\sigma^2_{\epsilon,t}\right] \geq -Cov\left[\sigma^2_{f,t}, \sigma^2_{\epsilon,t}\right]$.[33] In this empirically relevant case, the researcher would obtain a downward-biased coefficient estimate for γ even when using the true VMF variance as a predictor, for which we provide evidence in the next subsection.

5.2 Risk-return tradeoff estimates

In this subsection, we investigate how accounting for the IFF affects estimates of the market index intertemporal risk-return tradeoff relationship using GARCH volatility forecasting models. To demonstrate the effects of the IFF, we estimate GARCH models using either the VMF or the DMF. Our empirical tests estimate the following specification

$$r_{t+1} = \phi_0 + \phi_1 \cdot r_t + \gamma \cdot \sigma^2_t + u_{t+1}, \quad (19)$$

$$u_{t+1} = \sigma_t v_{t+1}, \quad (20)$$

$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 u^2_t + \alpha_2 \sigma^2_t, \quad (21)$$

where r_{t+1} is either VMF_{t+1} or DMF_{t+1}, and u_t is a forecasting error from the previous period and v_{t+1} has unit variance. Here the conditional variance forecasts (σ^2_t) following the GARCH(1,1) process can be viewed as a weighted average of the squared forecasting error and the conditional variance forecast from the previous period, with corresponding weights being α_1 and α_2, respectively.

While we adopt the GARCH(1,1) as our baseline volatility forecasting model, there are numerous extensions of GARCH models aimed at better capturing pervasive features in the data including volatility persistence, asymmetry, and the leptokurtic distribution of financial asset returns. Accordingly, we complement our baseline GARCH(1,1) with estimates from two additional nonlinear GARCH models: the

[33] This condition is guaranteed if $\sigma^2_{f,t}$ and $\sigma^2_{\epsilon,t}$ are positively correlated or if $Var\left[\sigma^2_{\epsilon,t}\right] > Var\left[\sigma^2_{f,t}\right]$.

threshold GARCH (GJR-GARCH(1,1)) from Glosten et al. (1993), and the exponential GARCH (E-GARCH(1,1)) from Nelson (1991). These two models, whose specifications for volatility are detailed in Online Appendix OA.9 (along with other details about our GARCH implementation), are widely adopted in the empirical literature for the risk-return trade-off given their ability to capture "leverage effects", a known feature of financial asset volatility.[34]

We further investigate the risk-return tradeoff relationship under an alternative model specification. Rather than using conditional variance as a proxy for time-varying risk as in equation (19), we consider using conditional volatility as studied by Baillie and DeGennaro (1990) and French et al. (1987) according to:

$$r_{t+1} = \phi_1 + \phi_2 \cdot r_t + \gamma \cdot \sigma_t + u_{t+1} \tag{22}$$

We provide estimates of the conditional volatility from the baseline GARCH specification in Online Appendix Figure OA.6 for visualization. As expected from the unconditional analysis in Table 2, the conditional volatility is typically lower for the DMF compared to the VMF. Additionally, the difference exhibits substantial variation at both high and low frequencies. This nontrivial difference lends empirical support to our conjecture that the IFF adds noise to the lagged forecasting errors used to generate the conditional variance forecasts within the GARCH forecasting models.

Table 3 reports γ values estimated according to the above specifications using our 1963-2021 sample.[35] Panel A provides results from the baseline GARCH specification, whereas Panels B and C provide results for the GJR-GARCH and E-GARCH specifications, respectively. Each column reports results using a different market index proxy. The first column uses the VMF. Column 2 uses the DMF constructed by projecting the VMF onto the first PC. Column 3 reports results when the DMF is constructed by projecting the VMF onto the CRSP equal-weighted index for comparison. The γ differences reported in the table take the difference between that estimated from the VMF and each respective version of the DMF. Finally, the first

[34]For reviews of this literature, see French et al. (1987), Schwert (1990), Franses and Van Dijk (1996), Poon and Granger (2003), and Brownlees et al. (2012).

[35]We find similar results in a robustness check that performs the analysis using data from 1927-2021, which we discuss in more detail in Online Appendix OA.1.2. We present results for this robustness check in Table OA.3.

(second) set of results in the rows of each panel defines risk according to the variance (volatility) specification in equation (19) (equation (22)).

We find positive and statistically significant estimates for γ from the DMF across all specifications, which contrasts with the positive but always statistically insignificant γ estimates from the VMF. Furthermore, consistent with the attenuation bias discussed earlier, our estimate for γ from the VMF is always smaller than that from the DMF. In other words, adjusting the market index to purge it of the IFF yields a stronger risk-return tradeoff in the time series. The larger estimates of risk aversion that obtain when using the DMF are more consistent with those estimated in other economic settings such as through structural implementations of the intertemporal capital asset pricing model (ICAPM) (Campbell et al., 2018), factor model implementations of the ICAPM (Chabi-Yo et al., 2022), and in a setting that estimates the conditional market risk premium using option prices (Chabi-Yo and Loudis, 2020).

In the last column of Table 3, we estimate risk-return tradeoff coefficients when using excess returns on the CRSP equal-weighted index as an instrument. This version of the DMF provides similar the risk-return tradeoff estimates as in our main DMF specification. Given that the number of stocks in the equal weighted index is sufficiently large for the law of large numbers to hold almost exactly, the granular residual in this index is approximately zero. However, like the value-weighted market index and as Figure 3 shows, the equal-weighted market index features time-varying loadings on higher PCs and is unlikely to purge the time-varying loading components associated with the IFF (which is one reason we do not use this as our main DMF specification).

6 Explaining the cross-section of average returns in the presence of IFF

In this section, we analyze how the presence of the IFF affects standard cross-sectional asset pricing tests. In particular, the IFF can generate substantial differences between estimates of stock or portfolio-level market risk (β_p) when using either the VMF or DMF as the market proxy. This difference has implications for standard asset pricing tests, which we highlight in this section. In subsection 6.1, we provide some analytical results on how the IFF influences tests of the cross-sectional predictions of the model outlined in section 4, then present our empirical results in section 6.2.

6.1 How does IFF impact estimates of market betas?

When the factors f_{t+1} and g_{t+1} are tradable portfolios, which is the case for our proxies, the law of iterated expectations and (14) imply the expected excess return on portfolio p is given by:

$$\mathbb{E}[r_{p,t+1}] = \beta_p \cdot \gamma \cdot \mathbb{E}[\sigma^2_{f,t+1}] + \mathbb{E}[\Lambda' \Omega_t \zeta_p] = \beta_p \mathbb{E}[f_{t+1}] + \zeta_p' \mathbb{E}[g_{t+1}]. \tag{23}$$

Thus, the main empirical prediction of the theory is that, after controlling for exposures to other orthogonal priced factors, expected returns line up with β_p, where the slope of the relationship is equal to the market risk premium.[36]

In order to estimate the market risk premium and discount rate from exposure to market risk for a portfolio, we use a standard two-stage regression. In the first stage, we estimate loadings from the following time series regression:

$$r_{p,t+1} = \alpha_p + \beta_p \cdot f_{t+1} + \zeta_p' \cdot g_{t+1} + \eta_{p,t+1}, \tag{24}$$

where $r_{p,t+1}$ is the excess portfolio return. When using the DMF as a proxy for f_{t+1}, we simply use higher-order PCs as proxies for g_{t+1} in equation (24) because they are orthogonal to the DMF by construction. However, the PCs are not orthogonal to the VMF, so in analyses where we use the VMF as a proxy for f_{t+1}, we use higher-order PCs that have themselves been orthogonalized with respect to the VMF as proxies for g_{t+1}.[37] We then construct portfolio returns that are hedged of their exposure to g_{t+1} as follows

$$r^{hedged}_{p,t+1} \equiv r_{p,t+1} - \widehat{\zeta}_p' \cdot g_{t+1}. \tag{25}$$

Finally, we estimate λ_1 in the second-stage regression

$$\frac{1}{T} \sum_{t=1}^{T} r^{hedged}_{p,t+1} = \lambda_0 + \lambda_1 \cdot \widehat{\beta}_p + a_p. \tag{26}$$

[36]For ease of exposition, we slightly abuse the term "market risk premium" to refer to the risk premium on either the VMF or the DMF, depending on context. Empirically, this distinction is small because we cannot reject that time series averages of VMF and DMF returns are the same.

[37]Ensuring the g_{t+1} proxies in equation (24) are orthogonal to the f_{t+1} proxy both conforms to our assumptions about g_{t+1} and also makes it so that including addition PCs as proxies for g_{t+1} does not affect the estimate of β_p. See Online Appendix section OA.10 for more details on the orthogonalization procedure.

While it is more common to run a multivariate regression and freely estimate λ_2, the risk premia on g_{t+1}, the algebra is simpler – and estimates are more precise under the null that portfolios were priced properly – if we instead impose the theoretical restriction that $\lambda_2 = E[g_{t+1}]$ by using $r_{p,t+1}^{hedged}$ in the second-stage regression.[38] Equation (23) implies $\lambda_0 = 0$ and $\lambda_1 = \mathbb{E}[f_{t+1}]$. Two issues arise when estimating λ_1 and testing the latter condition. First, one faces the standard generated regressor problem when using estimated β_p's to estimate the parameters (λ_0 and λ_1) in equation (26). We deal with this issue by estimating associated standard errors in our tests according to the bootstrap procedure described in Online Appendix OA.11. Second, in the presence of IFF, we only have access to a noisy proxy for f_{t+1} (i.e., the VMF) which causes attenuation bias in β_p^{VMF} estimates relative to the true β_p's. We discuss this issue next.

When using the VMF as a proxy for f_{t+1} in the presence of IFF, the OLS estimate of β_p in equation (24) converges to[39]

$$\underset{T\to\infty}{\text{plim}}\widehat{\beta}_p^{VMF} = \beta_p \cdot \left\{ \frac{Var\left[f_{t+1}\right]}{Var\left[f_{t+1} + \epsilon_{t+1}\right]} \right\} + \frac{Cov\left[\epsilon_{t+1}, \eta_{p,t+1}\right]}{Var\left[f_{t+1} + \epsilon_{t+1}\right]} \equiv \beta_p + \kappa_p, \qquad (27)$$

where κ_p denotes the bias in $\widehat{\beta}_p^{VMF}$ relative to β_p. $\widehat{\beta}_p^{VMF}$ is affected by the IFF in two ways: (1) an attenuation effect represented by the term in the brackets, and (2) direct effect represented by the direct correlation between the IFF, ϵ_{t+1}, and the idiosyncratic error term of the portfolio, $\eta_{p,t+1}$. In practice, we expect the direct effect (the second term) to be larger for large stocks, given that the IFF naturally inherits the factor exposures associated with larger stocks. Regarding the first effect, the term in brackets is substantially less than 1, leading $\widehat{\beta}_p^{VMF}$ is likely to be attenuated downwards for most individual stocks. However, when using value-weighted portfolios as test assets, the direction of the bias is not immediately clear (note that these two forces exactly offset for the VMF). We expect downward bias for portfolios of small-cap stocks, and we expect this downward multiplicative bias to be partially nullified or even made positive by direct IFF exposures. In practice, we use the DMF as a

[38]We could also impose this constraint by running GLS and including the factors as test assets in our second-stage regressions. However, we choose to use OLS for these regressions because it 1) is more conventional, 2) does not necessitate the estimation of the residual covariance matrix (which would introduce a large amount of noise given that we use 372 test assets), and 3) produces the same asymptotic results as GLS under the model null (see Cochrane, 2005, Ch. 12.2).

[39]Due to the orthogonality assumption on g_{t+1}, the univariate and multivariate estimates of β_p both converge to the same value.

proxy for f_{t+1}, and we simply refer to the difference $\beta_p^{VMF} - \beta_p^{DMF}$ as the "beta gap" rather than calling it a "bias."

6.2 Empirical results with PC-based pricing factors

In this subsection, we use our empirical proxy for f_{t+1}, the DMF, to implement standard cross-sectional asset pricing tests and provide evidence the IFF induces differences between β_p^{DMF} and β_p^{VMF} (beta gaps) that are large enough to affect inferences from such tests. We also show that the beta gap behaves consistently with the analytical results developed in the previous subsection.

Our asset pricing tests in this section use the same 372 characteristics-sorted portfolios (or subsets of them) that were used to construct our PCA as test assets, with one key distinction. To be more consistent with standard practice, we use value-weighted versions of these portfolios in our tests (though we note that results are similar if we instead use equal-weighted test assets).[40] All analyses are at the monthly frequency and span the 1963-2021 sample period unless otherwise specified.

6.2.1 The beta gap is often large, especially for small-cap portfolios

Equation (27) implies the beta gap is likely more severe for assets with smaller market capitalization due to the multiplicative bias term. We begin by confirming that β_p's estimated using the VMF are generally lower than those estimated using the DMF in Figure 5, which plots kernel density estimates of the β_p^{VMF} and β_p^{DMF} estimated across all 372 portfolios. The blue (red) curve represents the β_p^{DMF} (β_p^{VMF}) density. The table beneath the figure displays percentiles of the beta distributions, as well as some summary statistics for the beta gap distribution (i.e., $\beta_p^{VMF} - \beta_p^{DMF}$). The β_p^{DMF} density has fatter tails, especially on the right, and is shifted rightwards relative to the β_p^{VMF} density. β_p^{DMF} percentiles are mostly larger than the β_p^{VMF} percentiles, with differences being particularly large in the right tail. While the average beta gap (-0.13) might not seem large relative to the average β_p^{VMF} (1.09), it is quite substantial relative to the cross-sectional standard deviation of β_p^{VMF} (0.19). Likewise, the standard deviation of the beta gap (0.14) and the standard deviation of β_p^{DMF} (0.28) are quite large relative to standard deviation of β_p^{VMF}.

[40]See Online Appendix Table OA.1 for a summary of these portfolios. These portfolios are constructed in a standardized manner and, conveniently, are also tabulated to include portfolio market capitalization, which will be useful for testing IFF predications related to asset size. This will allow us to test the above predictions that relate beta gap (i.e., κ_p) with portfolio size.

Figure 6 plots the beta gap for the following subsets of portfolios as a function of each portfolio's log-size: 1) 10 size-sorted portfolios, 2) 10 market-beta-sorted portfolios, 3) 25 size-by-beta-sorted portfolios, and 4) all 372 portfolios in our sample.[41] OLS best-fit lines are also plotted in red for visualization. Consistent with the intuition from Equation (27), the beta gap is larger in magnitude for portfolios with smaller market capitalization. Furthermore, the average log market capitalization of each portfolio explains a substantial fraction of the variation in the beta gap.

Finally, Table 4 reports estimates of the beta gap from size-by-beta-sorted portfolios over our 1963-2021 sample,[42] and offers perhaps the most direct test of our predictions about beta gap behavior. Recall from equation (27) that true betas are attenuated downward by a multiplicative factor implying that, holding the direct covariance between ϵ_{t+1} and $\eta_{p,t+1}$ (which depends mostly on size) constant, the attenuation bias is larger for portfolios with higher exposures to f_{t+1}. Within size quintiles, our estimates suggest that β_p^{VMF} accurately reflect the ordering of β_p^{DMF}, though a gap exists due to the presence of the IFF. Thus, we closely approximate this comparative static by moving down columns of Table 4 (i.e., increasing beta, holding size fixed). Indeed, beta gaps become more negative as we move down *every* column in the table. Likewise, holding β_p approximately constant by moving across columns within the same row, β_p^{VMF} increases with size, which is also consistent with intuition from equation (27). These patterns are uniformly monotonic, holding for every cell in the table.

6.2.2 Accounting for IFF increases the importance of market risk in explaining the cross-section

In this section, we run time series and cross-sectional regressions for each portfolio set in our universe of portfolios using either the VMF or the DMF as a proxy for f_{t+1}, and study a number of related statistics including: 1) the standard deviation of β_p, 2) the market risk premium (λ_1), 3) the cross-sectional standard deviation of the market-exposure-implied discount rate across portfolios in each portfolio set ($std(\lambda_1 \cdot \beta_p)$), 4) cross-sectional R-squared values (i.e., the fraction of variation in average hedged

[41] Due to their historical significance, we also provide results for the 25 size-by-book-to-market (BM) sorted portfolios in Online Appendix Figure OA.7, Panel 1. Results are similar to those from the other portfolio sorts presented in Figure 6.

[42] We also show these results are robust when using a 1993-2021 sample in Online Appendix Table OA.4.

portfolio excess returns explained by DMF or VMF betas), and 5) the magnitude of time series alphas. Although we use the 1963-2021 sample and value-weighted test assets, our results are robust to using a more modern sample from 1993-2021 and using equal-weighted test assets.[43]

We present our main results in Figures 7-8 and in Table 5, which we describe briefly here before discussing more details below. Figure 7 plots SML lines constructed from two-stage regressions that use betas estimated from either the VMF or the DMF for each of the four portfolio sets from Figure 6.[44] The left panels plot results when average raw excess portfolio returns are used as left-hand side variables in equation (26), while the right panels plot results when the average of r_{t+1}^{hedged} from equation (25) controlling for five PCs are used as left-hand side variables.[45,46] Blue (red) numbers plot (β_p, $\mathbb{E}[r_{p,t}]$) pairs associated with the DMF (VMF). Analogously, the blue (red) line represents the corresponding OLS-estimated SML. The theory-implied SML is represented by the black line and is constructed using the average excess CRSP value-weighted market return as its slope.

Figure 8 provides a visual summary of statistics related to these SML plots. In particular, we provide the cross-sectional $std(\beta_p)$ (Panel 1), the market risk premium, λ_1 (Panel 2), the cross-sectional $std(\lambda_1 \cdot \beta_p)$ (Panel 3), and the cross-sectional R-squared values (Panel 4). All statistics are estimated separately for the 16 distinct portfolio sets in our sample of 372 portfolios (and for the set of all portfolios), and returns are hedged of exposure to the first five PCs according to equation (25).[47]

[43] See Online Appendix Tables OA.5 and OA.6, respectively.

[44] Due to their historical significance, we also provide results for the 25 size-by-book-to-market (BM) sorted portfolios in Online Appendix Figure OA.7, Panel 2. Results are similar to those from the other portfolio sorts presented in Figure 7.

[45] We provide robustness checks when estimating the all-portfolio SML with varying numbers of hedging PCs in Online Appendix Figure OA.8. These results show that results do not change much qualitatively or quantitatively when five or more PCs are used as controls.

[46] A few points should be noted when reviewing this figure. First, because the PC proxies for g_{t+1} are orthogonalized with respect to the VMF when generating the VMF-based SML (and the PC proxies need not be orthogonalized with respect to DMF because the PCs are already orthogonal by construction), the β_p^{DMF} and β_p^{VMF} values are the same in the left and right panels. One subtlety is that we must drop the first PC from the set of g_{t+1} controls when evaluating the DMF because it is perfectly collinear with the DMF. We find similar results (unreported) if we drop the first PC as a control when evaluating the VMF. Second, although the β_p estimates are the same whether we control for g_{t+1} or not, controlling for g_{t+1} exposure does affect average returns. So, whereas using either the DMF or the VMF can shift β_p locations on the x-axis within each plot due to the beta gap, using r_{t+1}^{hedged} in the right panel can shift average returns on the y-axis relative to plots in the left panel.

[47] Results are qualitatively and quantitatively similar when we impose the normalization restric-

Finally, Table 5 presents detailed summary statistics when all portfolios are used as test assets and when using different numbers of PCs (up to 15) for the hedging procedure to show robustness to this choice.

Std($\beta_{\mathbf{p}}$) is uniformly higher across test asset sets when using the DMF

Figure 7 illustrates how using the DMF generates larger β_p spreads than the VMF across all portfolio sets analyzed in the figure. Focusing on the size-sorted portfolio results in Panels 1a-b, all the β_p^{DMF} values are larger than the corresponding β_p^{VMF} values (other than the largest size portfolio), which implies the attenuation term in equation (27) dominates the additive term in all but the largest size portfolio. Similarly, for the beta-sorted portfolio results in Panels 2a-b, the beta gap is typically negative for high-beta portfolios, indicating that the attenuation term in equation (27) dominates in these cases. The opposite is true for low-beta portfolios indicating the additive bias term in equation (27) dominates for these portfolios. These results are consistent with those presented in Figure 6 and Table 4. Panel 1 in Figure 8 shows that $std(\beta_p^{DMF})$ is uniformly higher than $std(\beta_p^{VMF})$ among all portfolio sets we investigate. Finally, the first column of Panel B in Table 5 shows that the difference $std(\beta_p^{DMF})-std(\beta_p^{VMF}) = 0.10$ is statistically significant with a standard error of 0.01 when all portfolios are used as test assets.

The DMF yields modestly larger, more stable market risk premium estimates that are closer to theoretical predictions

The slopes of the blue and red lines in Figure 7 represent market risk premium estimates using either the DMF or the VMF, respectively, for each portfolio set in the different panels. Those estimated using the DMF are typically closer to the theoretical SML slope in black and higher than those estimated using the DMF. One point to note is that both market risk premium estimates become better approximations of the theoretical market risk premium when using hedged returns that control for exposure to other factors in the right panel.[48] Focusing on the size-sorted portfolio results in

tions on PCs from subsection 4.2 such that $E[\sum_{i=1}^{N_t} w_{i,t}\beta_i] = 1$ and $E[\frac{1}{N_t}\sum_{i=1}^{N_t}\zeta_i] = 0$. This is not surprising because, as we discuss in Online Appendix section OA.3, the DMF (without these restrictions) is highly correlated with the alternative DMF when imposing these restrictions. We provide analogous results under this normalization restriction in Online Appendix Figure OA.9 for comparison.

[48]This result is related to a point made by Giglio and Xiu (2021). Namely, we can estimate consistent risk premia on a factor of interest when we properly control for exposure to all other factors that explain average returns in the cross section.

Panel 1, the more accurate slope estimates under the DMF specification result in lower pricing errors with respect to the theoretical SML (vertical distances between expected returns and the black line). The pricing errors are almost uniformly smaller in this case relative to the VMF specification, implying accounting for the IFF weakens the size anomaly.

We now focus on the all-portfolio case in Panels 4a-b. When we use no controls (Panel 4a) we get the well-known result that the VMF-based SML is flat. Using the DMF improves this slightly, yielding a small but positive slope. Importantly, using the DMF yields a larger spread in betas than is achieved using the VMF. When we hedge for exposure to the first five PCs (Panel 4b), both SML lines conform more to the theoretical SML line, with the SML based on the DMF being closer.

Echoing the conclusions for the representative portfolio sets considered in Panels 1-3 of Figure 7, Panel 2 in Figure 8 shows that the market risk premium estimates are typically higher and more stable across all portfolio sets when using the DMF relative to when using the VMF as the market proxy. Panel A in Table 5 reports market risk premium (λ_1) estimates using all 372 portfolios as test assets and using a range of PC controls. λ_1 point estimates based on the DMF are typically higher than those based on the VMF and closer to their theory-based targets (i.e., time series averages of the DMF and the VMF).[49] However, these differences are not statistically significant. This result may seem surprising, but is not inconsistent with the IFF theory we develop, provided that other PCs absorb additional sources of confounding variation in priced risk.

The DMF explains a larger fraction of average return variation

Given our findings that SML slopes are similar (or modestly larger) but more stable, and that dispersion in market betas increases when we use the DMF as our market proxy, it follows that the dispersion in risk premia ($\lambda_1 \cdot \beta_p$) explained by market exposures increases. Once again, this conclusion turns out to be remarkably stable across different test portfolios. Panel 3 in Figure 8 shows that estimates of $std(\lambda_1 \cdot \beta_p)$ are *uniformly higher* across portfolio sets when the DMF is used, implying that a larger fraction of the overall variation in average returns can be explained by exposure to the DMF relative to that which can be explained by exposure to the VMF. Panel B in Table 5 shows that this result is robust across the number of PC hedging

[49]In unreported results, we find that we cannot reject the null hypothesis that the time series averages of the DMF and the VMF are the same.

controls we consider, and that the differences are statistically significant whenever five or more PCs are used as controls. In other words, as long as we control for enough other sources of priced risk, exposure to the DMF explains a higher fraction of cross-sectional variation in average returns relative to exposure to the VMF in a statistically significant way.

Panel 4 in Figure 8 displays cross-sectional R-squared values when different portfolio sets are used as test assets, and Panel C in Table 5 reports cross-sectional R-squared values when all portfolios are used as test assets but varying the number of PC hedging controls. In both cases, R-squared values generated using the DMF are uniformly higher than those generated using the VMF.[50]

The DMF and VMF yield similar time-series alphas

Panel D in Table 5 reports the mean absolute pricing error (MAPE) across all portfolios from time series regressions, which are similar when using either the DMF or the VMF. Although using the DMF does not improve these alphas relative to using the VMF, the IFF theory does not imply an improvement in alphas. However, it is reassuring that we can replace the standard VMF with our DMF without significantly affecting the alphas.

6.2.3 IFF is not priced in bivariate overidentification tests

Table 6 provides cross-sectional regression results from a bivariate model that includes both the DMF and the IFF. It also provides statistics from the univariate DMF model (repeated from Table 5) for comparison. Introducing the IFF as a control increases the market risk premium estimate relative to that under the univariate model so that it is closer to the theoretical (time series average) target (see Panel A). The IFF does not generate a statistically or economically significant risk premium when using hedged portfolio returns. Interestingly, when using unhedged returns, the IFF risk premium is actually statistically and economically significant at -2.73% (annualized). The fact that the IFF commands a negative risk premium under this model implies that it is priced quite differently by investors than the market return, providing more evidence that our empirical measure of the IFF is distinct from the market itself. Panel B shows that the bivariate model generates more variation in

[50] A model's R-squared value is a common measure of performance in cross-sectional regressions. To avoid issues associated with the R-squared distributions (see Kan et al., 2013), we test for differences in R-squared values using the bootstrap procedure described in Online Appendix OA.11.

market-exposure-implied discount rates than under the univariate model due to the relatively high market risk premium estimates under the former model. Panels C and D display cross-sectional R-squared values and time series alphas, respectively, implied by each model, which are similar across models. These results imply that controlling for the IFF does not actually help the model's ability to price test assets.

7 Incorporating characteristics-based factors

As noted by Kozak et al. (2018), PCs that explain a significant amount of time series variation in returns may still play no role in pricing. To explore robustness to this possible shortcoming, in this section we conduct additional analyses using characteristics-based factors that are frequently used in the literature. In particular, we study factors from the following seven models (listed in order of publication date): 1) the Fama and French (1993) three-factor model ("FF3"), 2) the Carhart (1997) four-factor model ("FF3C"), 3) the Fama and French (2015) five-factor model ("FF5"), 4) the Hou et al. (2015) four-factor model ("q4"), 5) the Stambaugh and Yuan (2017) four-factor model ("SY"), 6) the Daniel et al. (2020) three-factor model ("DHS"), and 7) the Hou et al. (2021) five-factor model ("q5"). Additionally, we explore whether insights from the IFF theory that we develop affects inference in these factor models with respect to the importance of the size anomaly.

7.1 Conclusions from Section 6 are similar when controlling for standard factors

To begin, Table 7 reproduces results from Table 5, but hedges for exposure to factors in each model described above in addition to five PCs. Results are similar to those in Table 5, and we conclude that controlling for these factors does not affect our interpretation of results from the previous section. We also provide analogous results without hedging the first five PCs in Online Appendix Table OA.7. In this case, the market risk premium estimate is less volatile, the $std(\lambda_1 \cdot \beta_p)$ statistic is typically larger, and the cross-sectional R-squared is typically larger when using the DMF compared to using the VMF.

7.2 The size premium disappears when accounting for IFF

All of these factor models include a size factor (except the DHS model). The bias in β_p highlighted in equation (27) implies the measured market risk will be too low on small-cap portfolios when using the VMF as a market proxy. We therefore expect the size premium embodied by size factors in these models to be reduced when using the DMF as a market proxy compared to using the VMF, which is also consistent with implications from our theoretical model. To document the reduced role of size factors in these models when the IFF is accounted for, we run factor spanning tests for the size factors in each model.

Table 8 reports size factor alphas from regressions of size factors in each model on the other model factors when using either the DMF or the VMF as the market factor. If the resulting size factor alphas are small and statistically insignificant, this implies adding the size factors to these models provides little improvement in a model's ability to price test assets (Barillas and Shanken, 2017). Panel A provides results over the 1963-2021 sample. In this case, when we use the DMF, the size factors in all models are spanned by the other factors (i.e., the size factor alphas are both economically and statistically insignificant). However, when we use the VMF, the size factor alphas are both statistically and economically significant in all models except for the FF3 and FF3C models. We also report βs of the size factors with respect to both the DMF and the VMF. In line with our theory, size factors have higher loadings on the DMF than the VMF, which allows for models with the DMF to span size factors. Results are similar in the 1963-1992 (Panel B) and 1993-2021 (Panel C) sub-samples. So, using a market factor that is adjusted to account for the IFF effectively allows us to remove the size factors from all of these models without impacting their pricing implications.

7.3 IFF has a negative alpha and is largely explained by size factors

To further explore how the IFF is related to size factors in these models, we run spanning tests for the IFF either including or excluding the size factors. As implied by the previous subsection, we expect the IFF to be closely related to size factors in each of these models. In particular, because the IFF contains information associated with large cap stocks, we expect the IFF to be negatively correlated with size factors.

Table 9 presents results from these spanning tests where we regress the IFF on model factors using either the VMF (Panel A) or the DMF (Panel B) as the market factor. Within each panel, we also run tests where we either include or exclude the size factor in each model as a right-hand-side variable. Starting with Panel A, the IFF alphas are always negative and typically statistically significant with the magnitude increasing when size factors are omitted. On the other hand, when we use the DMF as the market proxy, we cannot reject the null that the IFF alphas are equal to zero whether we include size factors or not.[51] Overall, these results are consistent with the IFF not being priced when using the DMF. However, when using the VMF (on which the IFF has a mechanical positive loading), including size factors helps improve IFF pricing because of the IFF's negative loading on size.

7.4 VMF-based results are sensitive to including size factors but DMF-based results are not

Finally, to investigate how the IFF and size factors interact in the cross-section, we repeat the analysis from Figure 8 when hedging test portfolio returns of exposure to model factors either including or excluding size factors as part of the controls and without using any PC controls. We drop the PC controls here to avoid including any controls that may correlate with size factors and present results for this analysis in Figure 9.[52] The dark (light) red bars represent statistics that obtain when hedging for all factors (all factors excluding the size factor) in a given model and when the VMF is used. The dark (light) blue bars are analogous, but use the DMF.

Panel 1 displays market risk premium estimates, λ_1, for these various specifications. When using the VMF, including size factor results in a more volatile (across models) risk premium estimate that is often far too high (and always higher than the market risk premium estimate from models that include the size factor controls). Contrarily, when using the DMF, market risk premium estimates are reasonable in magnitude and similar whether size factor controls are included or not. Simi-

[51] This is with the exception that alphas under the q4 and q5 models are marginally significant, which is due to the fact that the IFF has strong negative loadings on factors with positive risk premia in these models.

[52] We also provide an analogous figure when hedging the first five PCs in addition to factors from the models in Online Appendix Figure OA.10. However, results are very similar between models that control for the size factors versus those that do not, which is an indication that the first five PCs span the size factors in these models. Furthermore, we provide detailed summary statistics related to this figure in Online Appendix Tables OA.8 and OA.9.

lar patterns obtain in Panel 2, which displays the cross-sectional standard deviation in market-exposure-implied risk premia associated with these models, $std(\lambda_1 \cdot \beta_p)$. Panel 3 displays R-squared values from these regressions. In all cases, R-squared values are higher when using the DMF rather than the VMF whether size factors are included or not. Additionally, aside from the FF3 and FF3C model cases, the highest R-squared values correspond with cases that use the DMF and exclude the size factors. Panel 4 displays the mean absolute pricing errors (MAPE) from time series regressions. When the VMF is used (red bars), the MAPE is always markedly higher when the size factor is excluded as a control. On the other hand, when the DMF is used (blue bars), the MAPE is comparable between models that include or exclude the size factors. Importantly, when excluding the size factors, the DMF always generates MAPEs that are lower than when using the VMF.[53] All-in-all, results indicate these statistics of interest are sensitive to controlling for size factors when using the VMF, but are relatively insensitive to this choice when using the DMF. Furthermore, the DMF generally performs better than the VMF. These results further support our claim that purging the market factor of the IFF reduces the role of the size-related risk in explaining average returns in the cross section.

8 Conclusion

In this paper, we provide evidence that the standard value-weighted market factor is not well diversified and, while most its fluctuations are explained by a dominant market factor (DMF), it also is driven by nontrivial, time-varying exposures on factors to which an average firm and macroeconomic variables are not exposed. As we argue through analytical calculations and empirical results, the presence of this IFF can lead a researcher to be more likely to reject predictions of standard asset pricing theory. Quantitatively, after eliminating the IFF, we find stronger evidence of a link between the conditional mean and volatility of the market portfolio. In addition, we demonstrate that estimated market exposures can often be substantially biased in the cross-section and provide evidence that correcting for the presence of the IFF outperforms the standard value-weighted index across a number of standard asset

[53]The astute reader will notice that dark red bars in Panel 4 of Figure 9 are often slightly lower than their dark blue DMF counterparts, though differences in magnitudes are small. Online Appendix Figure OA.11 shows that this pattern reverses (i.e., MAPEs are lower with the DMF as the market proxy) if we use equal-weighted portfolios as test assets.

pricing statistics. Finally, we consider the implications of the IFF for standard factor models, and we find that using our DMF in place of the value weighted market portfolio eliminates the need to include size factors across all models we investigate.

While value-weighting has many advantages (e.g., lower rebalancing, theoretical justification from CAPM), we uncover one potential disadvantage. We focus on empirical asset pricing tests related to the market return. However, we have not explored implications for value-weighted portfolios that are often used as test assets, which we leave for future research. Indeed, we find the IFF loads on other standard pricing factors. Further, our mechanism may help to make sense of potentially "puzzling" results such as the observation that the CAPM works better on announcement days (Savor and Wilson, 2014), or that the CAPM works well for jumps in market returns but not well for diffusive shocks (Bollerslev et al., 2016).[54]

Similarly, the fat-tailed size distribution likely implies that other aggregate measures commonly used in empirical asset pricing may be affected by instability coming from the changing composition of large firms. For instance, making similar corrections to the aggregate price-dividend ratio could generate stronger return/cash flow predictability. Changing large firm composition can contaminate aggregate valuation ratios with terms unrelated to the key source of predictability of interest, which is related to the more standard argument about confounding effects of cash flow predictability as discussed by Kelly and Pruitt (2013).[55] The main point we highlight in is that value weighted indices may reflect many sources of risk which are not important for the average firm and that stripping out these factors can lead to appreciably different conclusions when testing related theories.

References

Acharya, V. V. and L. H. Pedersen (2005): "Asset pricing with liquidity risk," *Journal of Financial Economics*, 77, 375–410.

[54]While ours is not the only potential explanation, the basic intuition for these types of results within our framework is that the relative variance of the DMF relative to the IFF could be larger in these cases, causing standard measures to more accurately reflect the quantity of risk during such isolated time periods.

[55]Indeed, Kelly and Pruitt (2013)'s approach of extracting a common factor from many portfolios may help to address these composition effects.

Amihud, Y. and H. Mendelson (1986): "Asset pricing and the bid-ask spread," *Journal of Financial Economics*, 17, 223–249.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009): "Real-time measurement of business conditions," *Journal of Business and Economic Statistics*, 27, 417–427.

Asness, C., A. Frazzini, R. Israel, T. J. Moskowitz, and L. H. Pedersen (2018): "Size matters, if you control your junk," *Journal of Financial Economics*, 129, 479–509.

Bai, J. and S. Ng (2006): "Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions," *Econometrica*, 74, 1133–1150.

Baillie, R. T. and R. P. DeGennaro (1990): "Stock returns and volatility," *Journal of Financial and Quantitative Analysis*, 25, 203–214.

Banz, R. W. (1981): "The relationship between return and market value of common stocks," *Journal of financial economics*, 9, 3–18.

Barillas, F. and J. Shanken (2017): "Which Alpha?" *Review of Financial Studies*, 30, 1316–1338.

Berk, J. B. (1995): "A Critique of Size-Related Anomalies," *Review of Financial Studies*, 8, 275–286.

Black, F. (1972): "Capital market equilibrium with restricted borrowing," *Journal of Business*, 45, 444–455.

——— (1976): "Studies in stock price volatility changes, Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section," *American Statistical Association*, 177–181.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988): "A capital asset pricing model with time-varying covariances," *Journal of political Economy*, 96, 116–131.

Bollerslev, T., S. Z. Li, and V. Todorov (2016): "Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns," *Journal of Financial Economics*, 120, 464–490.

Brandt, M. W. and Q. Kang (2004): "On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach," *Journal of Financial Economics*, 72, 217–257.

Brownlees, C., R. Engle, and B. Kelly (2012): "A practical guide to volatility forecasting through calm and storm," *Journal of Risk*, 14, 3.

Bryzgalova, S. (2015): "Spurious factors in linear asset pricing models," *LSE manuscript*, 1.

Bryzgalova, S., J. Huang, and C. Julliard (2022): "Bayesian Solutions for the Factor Zoo: We Just Ran Two Quadrillion Models," *Journal of Finance, forthcoming.*

Byun, S. (2016): "The usefulness of cross-sectional dispersion for forecasting aggregate stock price volatility," *Journal of Empirical Finance*, 36, 162–180.

Campbell, J. Y. (1996): "Understanding risk and return," *Journal of Political economy*, 104, 298–345.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2018): "An intertemporal CAPM with stochastic volatility," *Journal of Financial Economics*, 128, 207–233.

Campbell, J. Y. and L. Hentschel (1992): "No news is good news: An asymmetric model of changing volatility in stock returns," *Journal of Financial Economics*, 31, 281–318.

Campbell, S. D., S. Delikouras, D. Jiang, and G. M. Korniotis (2016): "The human capital that matters: Expected returns and high-income households," *The Review of Financial Studies*, 29, 2523–2563.

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Cederburg, S. and M. S. O'Doherty (2019): "Understanding the Risk-Return Relation: The Aggregate Wealth Proxy Actually Matters," *Journal of Business and Economic Statistics*, 37, 721–735.

Chabi-Yo, F., A. Goncalves, and J. Loudis (2022): "An Intertemporal Risk Factor Model," *Working paper.*

Chabi-Yo, F. and J. Loudis (2020): "The conditional expected market return," *Journal of Financial Economics*, 137, 752–786.

Chamberlain, G. and M. Rothschild (1983): "Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51, 1281–1304.

Chou, R. Y. (1988): "Volatility persistence and stock valuations: Some empirical evidence using GARCH," *Journal of Applied Econometrics*, 3, 279–294.

Chou, R. Y., R. F. Engle, and A. Kane (1992): "Measuring risk aversion from excess returns on a stock index," *Journal of Econometrics*, 52, 201–224.

Christie, A. A. (1982): "The stochastic behavior of common stock variances: Value, leverage and interest rate effects," *Journal of Financial Economics*, 10, 407–432.

Cochrane, J. H. (2005): *Asset Pricing*, Princeton, New Jersey: Princeton University Press.

Connor, G. and R. A. Korajczyk (1986): "Performance measurement with the arbitrage pricing theory: A new framework for analysis," *Journal of Financial Economics*, 15, 373–394.

——— (1988): "Risk and return in an equilibrium APT: Application of a new test methodology," *Journal of Financial Economics*, 21, 255–289.

Cremers, M., A. Petajisto, and E. Zitzewitz (2013): "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation," *Critical Finance Review*, 2, 1–48.

Daniel, K., L. Mota, S. Rottke, and T. Santos (2020): "The Cross-Section of Risk and Return," *Review of Financial Studies*, 33, 1927–1979.

Davis, S. J., J. Haltiwanger, R. Jarmin, J. Miranda, C. Foote, and E. Nagypal (2006): "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms [with Comments and Discussion]," *NBER Macroeconomics Annual*, 21, 107–179.

Fama, E. F. and K. R. French (1992): "The cross-section of expected stock returns," *Journal of Finance*, 47, 427–465.

——— (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

——— (1996): "Multifactor explanations of asset pricing anomalies," *Journal of Finance*, 51, 55–84.

——— (2015): "A Five-Factor Asset Pricing Model," *Journal of Financial Economics*, 116, 1–22.

Franses, P. H. and D. Van Dijk (1996): "Forecasting stock market volatility using (nonlinear) GARCH models," *Journal of Forecasting*, 229–235.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987): "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3–29.

Freyberger, J., A. Neuhierl, and M. Weber (2020): "Dissecting characteristics nonparametrically," *Review of Financial Studies*, 33, 2326–2377.

Gabaix, X. (2011): "The granular origins of aggregate fluctuations," *Econometrica*, 79, 733–772.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005): "There is a risk-return trade-off after all," *Journal of Financial Economics*, 76, 509–548.

Giglio, S. and D. Xiu (2021): "Asset Pricing with Omitted Factors," *Journal of Political Economy*, 129, 1947–1990.

Giglio, S., D. Xiu, and D. Zhange (2021): "Test assets and weak factors," *Available at SSRN 3768081.*

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993): "On the relation between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779–1801.

Gomez, M. (2017): "Asset Prices and Wealth Inequality," *Working paper*.

Goyal, A. and P. Santa-Clara (2003): "Idiosyncratic risk matters!" *Journal of Finance*, 58, 975–1007.

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical Asset Pricing via Machine Learning," *Review of Financial Studies*, 33, 2223–2273.

Hamilton, J. D. (1994): *Time Series Analysis*, Princeton, New Jersey: Princeton University Press.

Harvey, C. R. (1989): "Time-varying conditional covariances in tests of asset pricing models," *Journal of Financial Economics*, 24, 289–317.

Harvey, C. R. and Y. Liu (2021): "Lucky factors," *Journal of Financial Economics*, 141, 413–435.

Heaton, J. and D. Lucas (2000): "Portfolio choice and asset prices: The importance of entrepreneurial risk," *Journal of Finance*, 55, 1163–1198.

Hedegaard, E. and R. J. Hodrick (2016): "Estimating the risk-return trade-off with overlapping data inference," *Journal of Banking & Finance*, 67, 135–145.

Hou, K., C. Xue, and L. Zhang (2015): "Digesting Anomalies: An Investment Approach," *Review of Financial Studies*, 28, 650–705.

——— (2021): "An Augmented q-Factor Model with Expected Growth," *Review of Finance*, 25, 1–41.

Jagannathan, R. and Z. Wang (1996): "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3–53.

Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American Economic Review*, 95, 161–182.

Kan, R., C. Robotti, and J. Shanken (2013): "Pricing Model Performance and the Two-Pass Cross-sectional Regression Methodology," *Journal of Finance*, 68, 2617–2648.

Kan, R. and C. Zhang (1999): "Two-pass tests of asset pricing models with useless factors," *Journal of Finance*, 54, 203–235.

Kelly, B. and S. Pruitt (2013): "Market Expectations in the Cross-Section of Present Values," *Journal of Finance*, 68, 1721–1756.

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

Kleibergen, F. (2009): "Tests of risk premia in linear factor models," *Journal of Econometrics*, 149, 149–173.

Koopman, S. J. and E. Hol Uspensky (2002): "The stochastic volatility in mean model: empirical evidence from international stock markets," *Journal of Applied Econometrics*, 17, 667–689.

Kothari, S. P., J. Shanken, and R. G. Sloan (1995): "Another Look at the Cross-Section of Expected Stock Returns," *Journal of Finance*, 50, 185–224.

Kozak, S., S. Nagel, and S. Santosh (2018): "Interpreting factor models," *Journal of Finance*, 73, 1183–1223.

——— (2020): "Shrinking the cross-section," *Journal of Financial Economics*, 135, 271–292.

Lettau, M. and M. Pelger (2020a): "Estimating latent asset-pricing factors," *Journal of Econometrics*, 218, 1–31.

——— (2020b): "Factors That Fit the Time Series and Cross-Section of Stock Returns," *Review of Financial Studies*, 33, 2274–2325.

Lewis, R. and S. Santosh (2021): "Investor betas," *Available at SSRN 3739424*.

Lintner, J. (1965): "The valuation of risk assets and the selection of risky investments in stock portfolios and risky budgets," *Review of Economics and Statistics*, 47, 13–37.

Ludvigson, S. C. and S. Ng (2007): "The empirical risk–return relation: A factor analysis approach," *Journal of Financial Economics*, 83, 171–222.

Lundblad, C. (2007): "The risk return tradeoff in the long run: 1836–2003," *Journal of Financial Economics*, 85, 123–150.

Lustig, H., S. Van Nieuwerburgh, and A. Verdelhan (2013): "The wealth-consumption ratio," *The Review of Asset Pricing Studies*, 3, 38–94.

Malevergne, Y., P. Santa-Clara, and D. Sornette (2009): "Professor ZIPF goes to wall street," Tech. rep., National Bureau of Economic Research.

Malloy, C. J., T. J. Moskowitz, and A. Vissing-Jørgensen (2009): "Long-run stockholder consumption risk and asset returns," *Journal of Finance*, 64, 2427–2479.

Mayers, D. (1973): "Nonmarketable assets and the determination of capital asset prices in the absence of a riskless asset," *The Journal of Business*, 46, 258–267.

Nagel, S. (2013): "Empirical cross-sectional asset pricing," *Annu. Rev. Financ. Econ.*, 5, 167–199.

Nelson, D. B. (1991): "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, 347–370.

Pagan, A. R. and Y. Hong (1991): "Nonparametric estimation and the risk premium," *Nonparametric and Semiparametric Methods in Econometrics and Statistics, W. Barnett, J. Powellm and GE Tauchen (eds)*, 51–75.

Piketty, T., E. Saez, and G. Zucman (2017): "Distributional national accounts: methods and estimates for the United States," *The Quarterly Journal of Economics*, 133, 553–609.

Poon, S. H. and C. W. Granger (2003): "Forecasting volatility in financial markets: A review," *Journal of Economic Literature*, 41, 478–539.

Pukthuanthong, K., R. Roll, and A. Subrahmanyam (2019): "A protocol for factor identification," *Review of Financial Studies*, 32, 1573–1607.

Roll, R. (1977): "A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory," *Journal of financial economics*, 4, 129–176.

Ross, S. A. (1976): "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13, 341–360.

——— (1977): "The capital asset pricing model (capm), short-sale restrictions and related issues," *Journal of Finance*, 32, 177–183.

Savor, P. and M. Wilson (2014): "Asset pricing: A tale of two days," *Journal of Financial Economics*, 113, 171–201.

Schmidt, L. (2016): "Climbing and Falling Off the Ladder: Asset Pricing Implications of Labor Market Event Risk," Working paper.

Schwert, G. W. (1990): "Stock returns and real activity: A century of evidence," *Journal of Finance*, 45, 1237–1257.

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

Stambaugh, R. F. (1982): "On the Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis," *Journal of Financial Economics*, 10, 237–268.

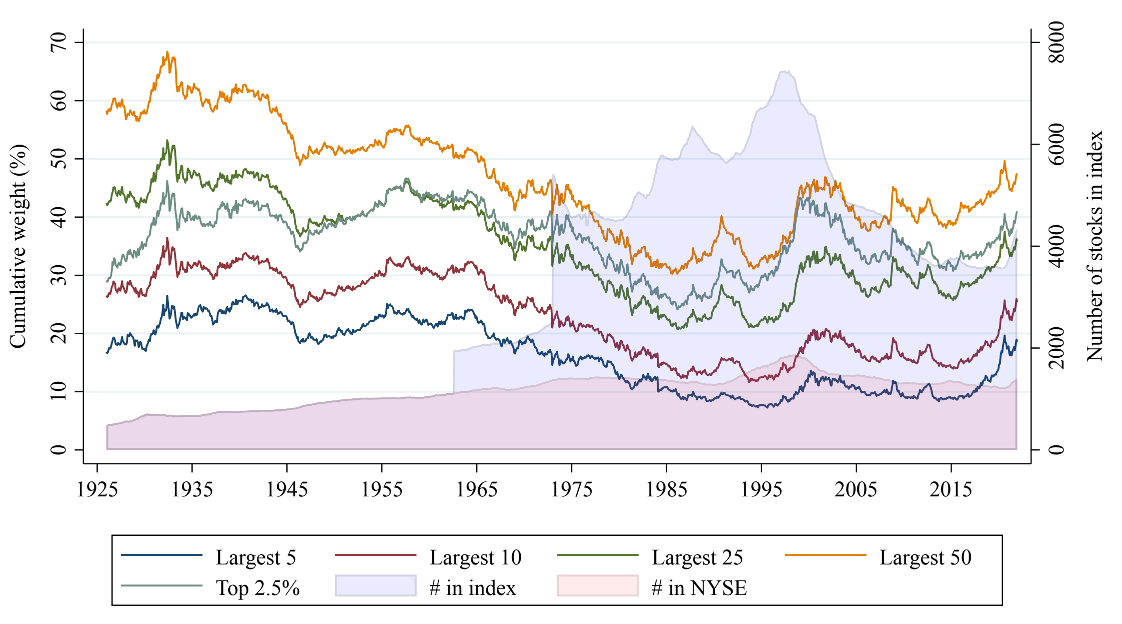
Stambaugh, R. F. and Y. Yuan (2017): "Mispricing Factors," *Review of Financial Studies*, 30, 1270–1315.

Stock, J. H. and M. W. Watson (2002): "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

van Binsbergen, J. H. and C. C. Opp (2019): "Real Anomalies," *Journal of Finance*, 74, 1659–1706.

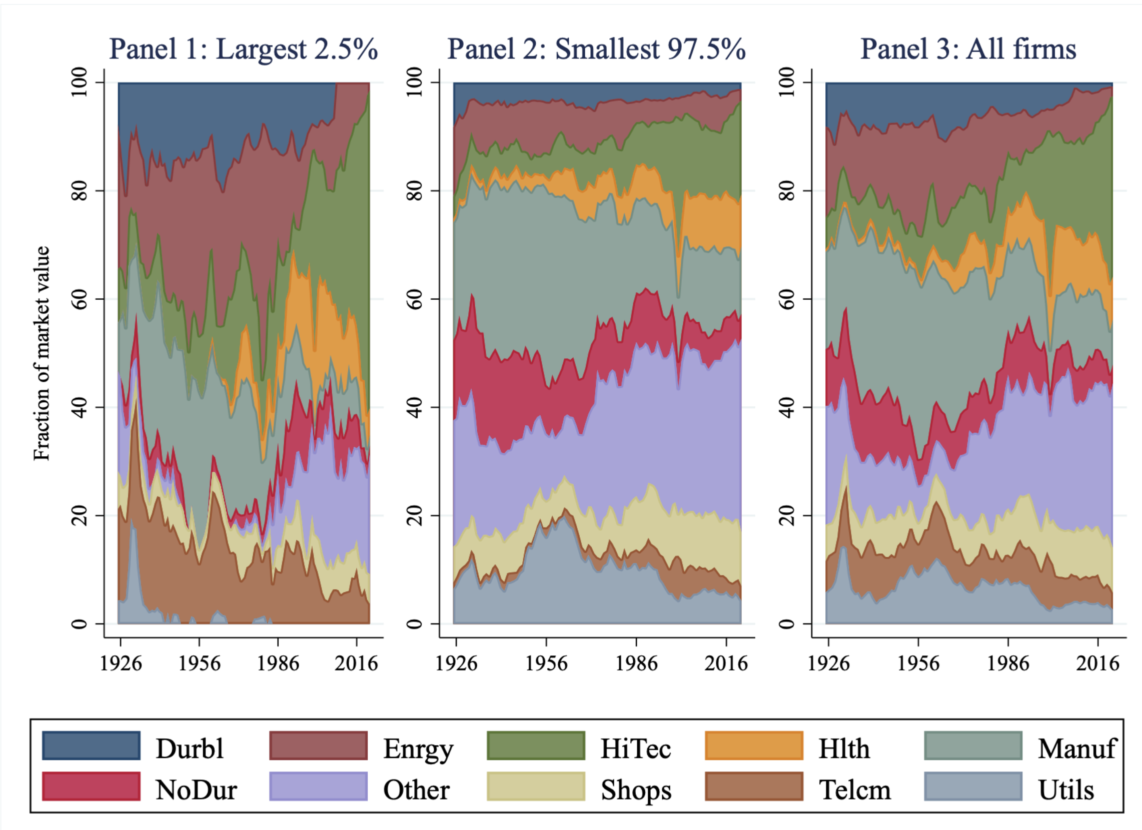
Whitelaw, R. F. (1994): "Time variations and covariations in the expectation and volatility of stock market returns," *Journal of Finance*, 49, 515–541.

Figure 1: Combined weights of largest stocks in the CRSP value-weighted index



Note. This figure, on the left axis, plots combined weights of individual stocks for the 5, 10, 25, and 50 largest stocks in the CRSP value-weighted index respectively. The top 2.5% line, also on the left axis, plots the cumulative market value associated with stocks whose rank in the cross-sectional market value distribution is less than 2.5% times the number of stocks in the NYSE (using the NYSE to compute this breakpoint avoids generating mechanical increases in the series as many small AMEX and NASDAQ firms are added to the database). The right axis plots the total number of stocks included in the CRSP index as well as the NYSE only, where the former includes two large jumps as the CRSP sample coverage expands. One can visually confirm that combined weights from a very small number of large firms dominate the rest of firms in the CRSP value-weighted index.

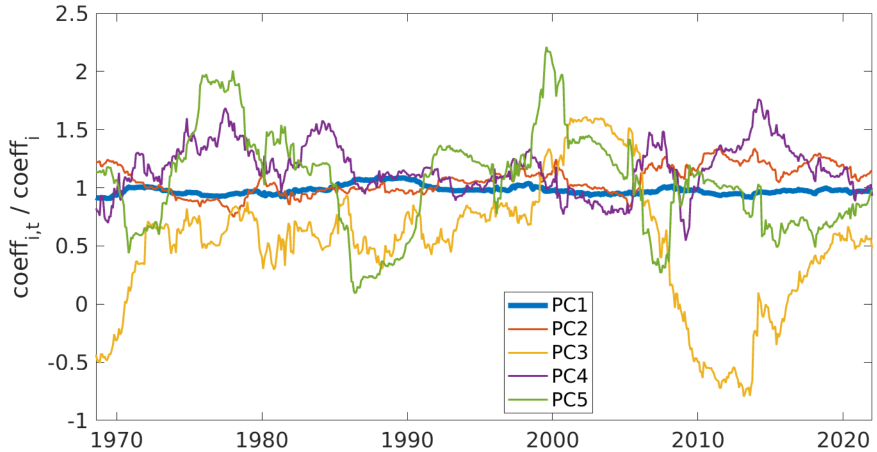
Figure 2: Industry composition of megacaps vs other stocks in CRSP value-weighted index



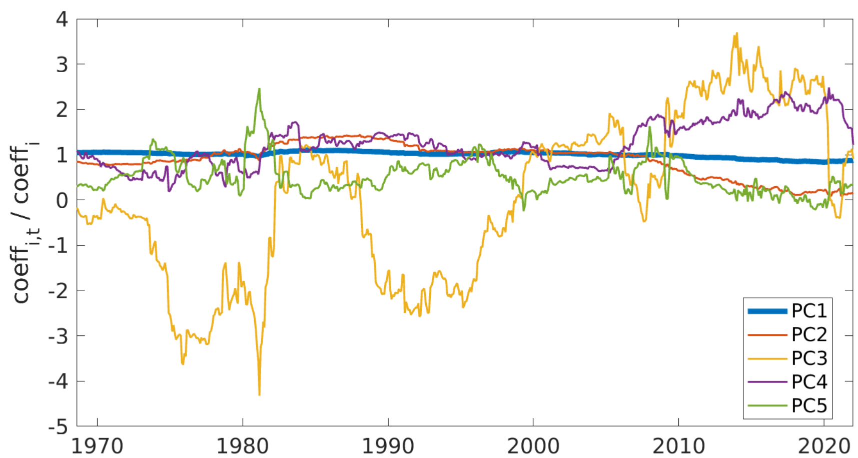
Note: This figure plots the share of total market value within each industry, by year, of stocks in the top 2.5% of the market value distribution (Panel 1), that for all other stocks (Panel 2), and that for all stocks (Panel 3) in the CRSP universe. Industries are defined using the Fama-French 10 industry classification using firm SIC codes, see Kenneth French's website for detailed definitions. We take the SIC code from Compustat when available; otherwise, we use the SIC code from CRSP. We use the number of stocks in the NYSE to classify stocks as belonging to the top 2.5% of the market value distribution to avoid generating mechanical increases in the composition of firms in the top category as many small AMEX and NASDAQ firms are added to the database. The Other category includes mining, construction, building materials, transportation, hotels, business services, entertainment, and finance.

Figure 3: Rolling coefficients of market index on the first three cross-sectional PCs

Panel 1: VW market on EW PCs

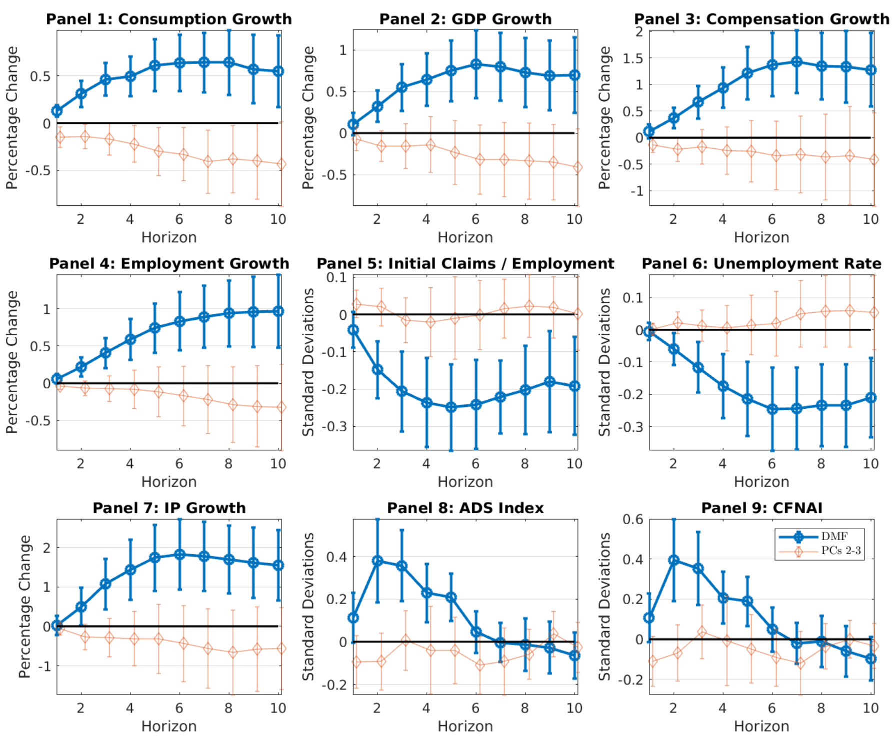


Panel 2: EW market on EW PCs



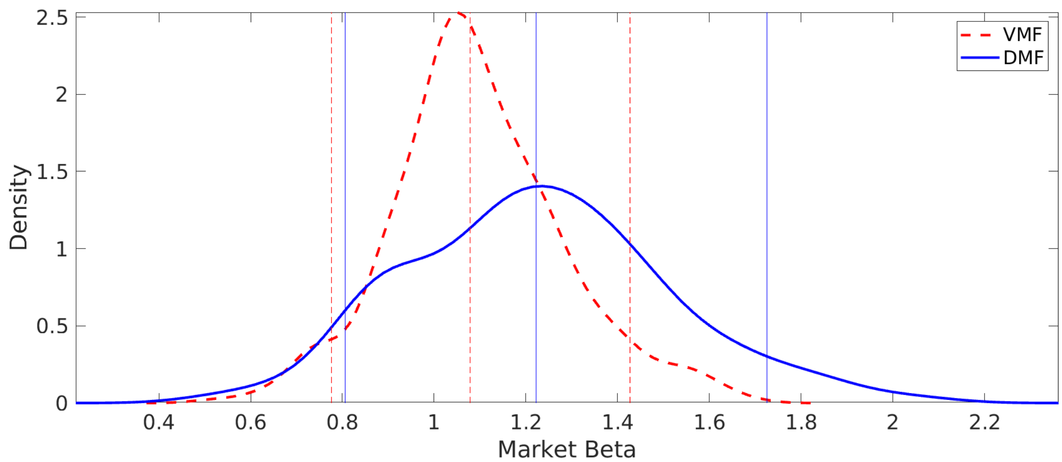
Note. This figure plots normalized coefficients from rolling regressions of the prior 60 months of excess market returns on five principal components constructed from a large cross-section of equal-weighted, characteristics-sorted portfolios. For ease of interpretation, coefficients are scaled by their full sample, OLS counterparts. Panel 1 plots estimates when using the CRSP value-weighted market return, and Panel 2 plots estimates when using the CRSP equal-weighted market return. All data are from the 1963-2021 sample.

Figure 4: Impulse response functions to PCs



Note. This figure displays impulse response functions of selected macroeconomic aggregates to scaled versions of the DMF and to a linear projection of VMF onto PCs two and three where the weights are obtained from a full-sample multivariate linear regression of the CRSP value-weighted market excess returns on PCs one through three as described by equation (OA.5). Monthly returns are compounded to the quarterly frequency to match that of the macroeconomic aggregates. The point estimates represent the impact of a standard deviation increase in the predictor variable on each macroeconomic aggregate. The initial point estimates (i.e., when the horizon is equal to one) correspond to the contemporaneous response of each macroeconomic aggregate to each predictor variable. The error bars indicate 95% confidence intervals computed using Newey-West standard errors with 10 lags. The descriptions of the macroeconomic variables can be found in the main text. All data are from the 1963-2019 sample. Additional details on the estimation procedure can be found in Online Appendix OA.4.

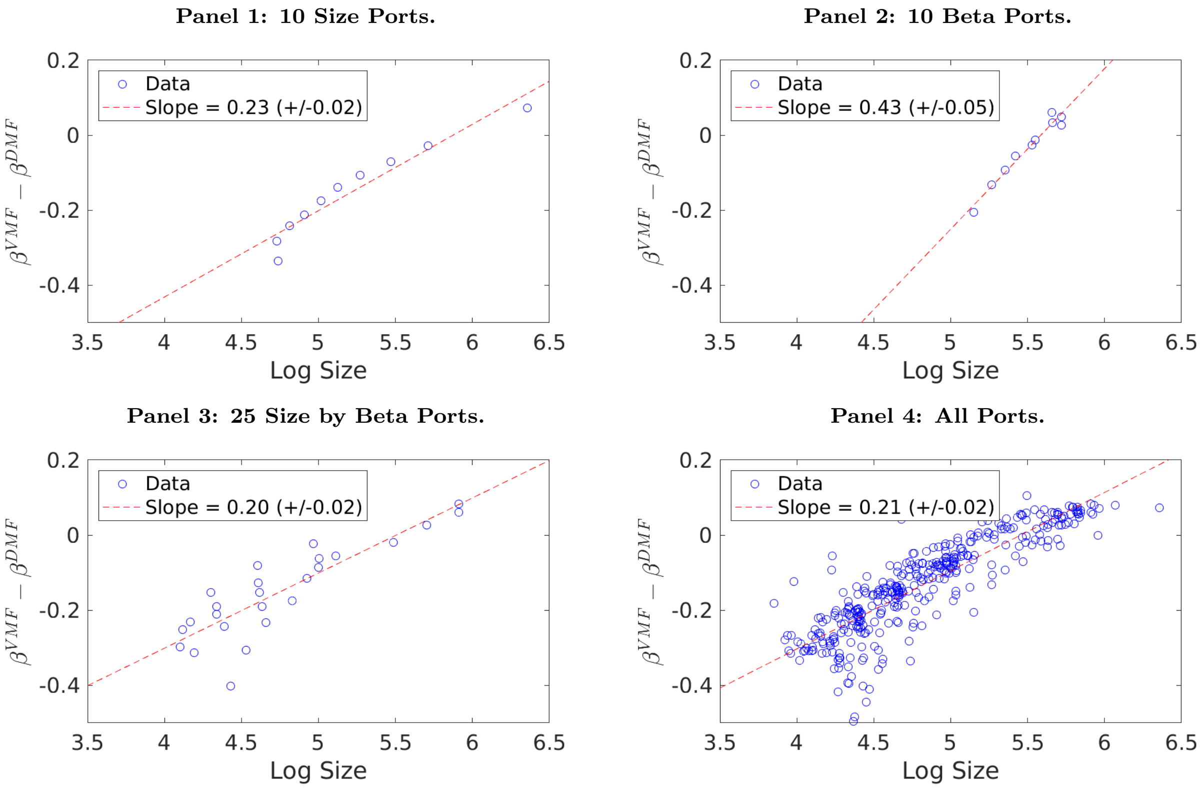
Figure 5: Distribution of β_p^{VMF} and β_p^{DMF} for a large number of test assets



	Mean	Stdev.	Market Beta Percentile 1	5	10	25	50	75	90	95	99
β_p^{DMF}	1.22	0.28	0.59	0.81	0.87	1.02	1.22	1.40	1.60	1.73	1.94
β_p^{VMF}	1.09	0.19	0.67	0.78	0.88	0.99	1.08	1.21	1.34	1.43	1.60
$\beta_p^{VMF} - \beta_p^{DMF}$	-0.13	0.13	0.08	0.06	0.05	-0.03	-0.13	-0.22	-0.30	-0.33	-0.42

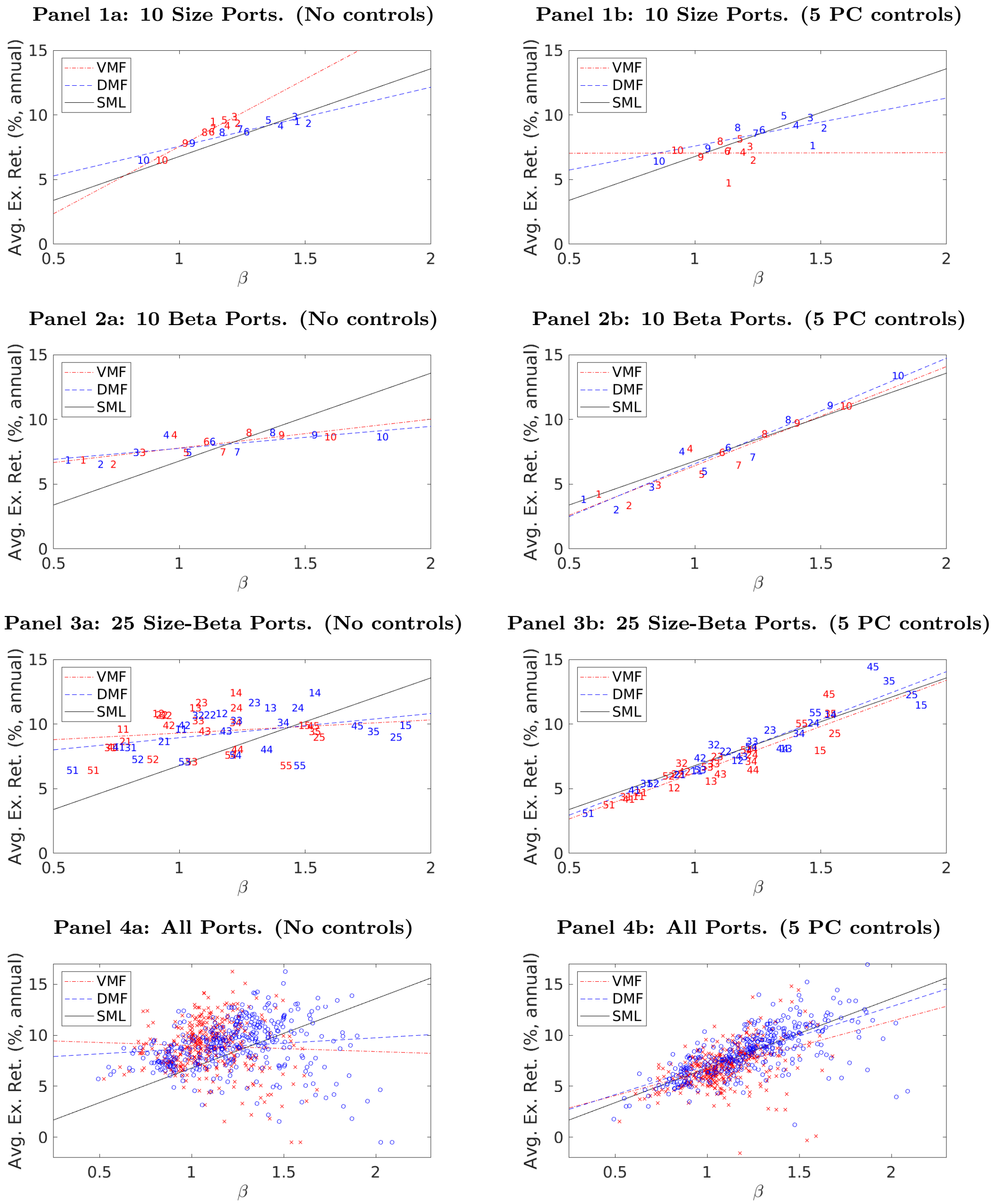
Note. This figure plots kernel density estimates of market betas estimated using the VMF and DMF. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Color-coded vertical lines indicate the 5th, 50th, and 95th percentiles of the two distributions. We also tabulate a number of univariate summary statistics for each distribution. All data are from the 1963-2021 sample.

Figure 6: Beta gap relationship for various portfolio sets of interest



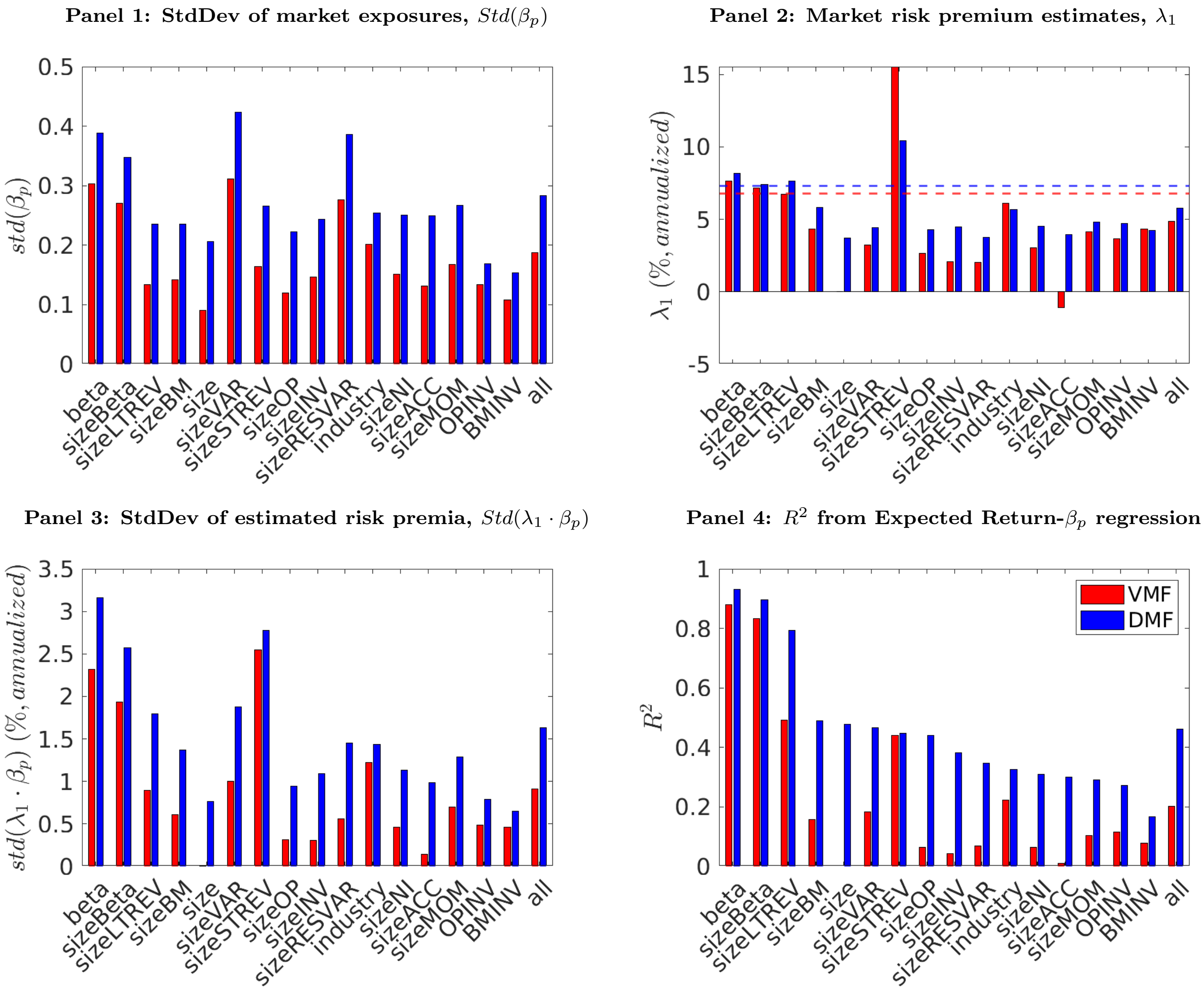
Note. This figure compares the beta gap ($\beta_p^{VMF} - \beta_p^{DMF}$) as a function of each portfolio's average log market equity. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. The title of each plot is the name of the portfolio sort used to construct the colored SMLs. The red line in each plot is the OLS best fit line. Log-size is computed as the average of each portfolio's monthly log market capitalization. All data are from the 1963-2021 sample.

Figure 7: Expected return-beta relationship for various portfolio sets of interest



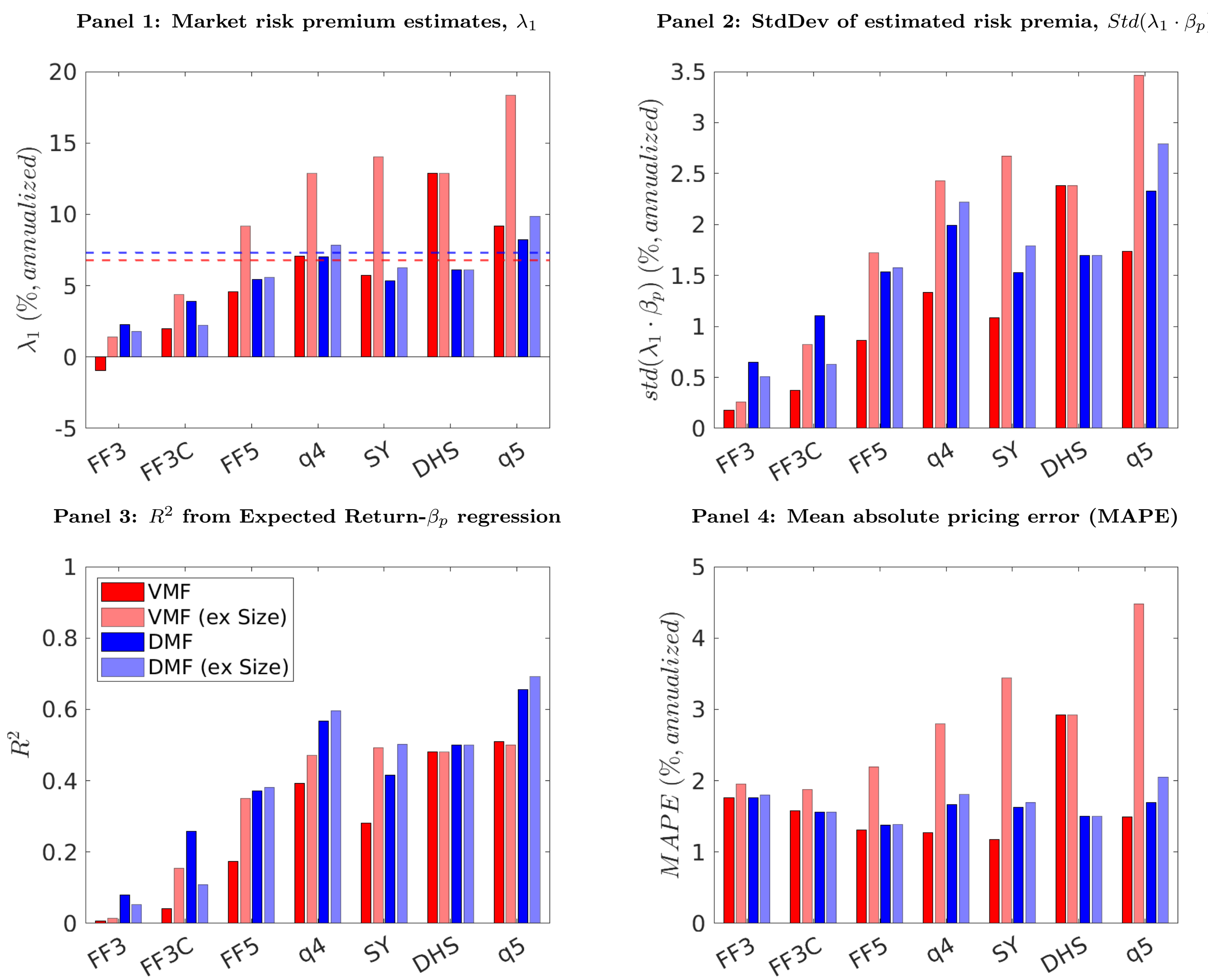
Note. This figure displays SML estimates for various portfolio sets, with and without the hedging procedure described by equation (25). The left column uses unhedged portfolio returns, while the right uses returns that are hedged of exposure to the first five PCs. The title of each plot is the name of the portfolio sort used to construct the colored SMLs. The red (blue) line is constructed from full-sample market betas estimated using the VMF (DMF). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. The black line is the SML implied by the average of VMF. In Panels 1-3, each number refers to the portfolio number in the sort. For the 25 portfolio sorts, the first number refers to the first sorting variable (size), and the second number refers to the second sorting variable (beta). For example, the size-by-beta-sorted portfolios, portfolio "15" is "small-high-beta" and portfolio "51" is "large-low-beta." In Panel 4, we use all 372 portfolios from the PCA as test assets in this case, which are described in Online Appendix Table OA.1. In this case, we denote each portfolio using a circle marker to avoid cluttering the graphs with numbers. All data are from the 1963-2021 sample.

Figure 8: Asset pricing statistics by portfolio set



Note. This figure plots comparisons of statistics related to SMLs implied by the VMF and DMF using specific sets of test assets (described on the x axes) to construct various respective SMLs. All portfolio returns are hedged with respect to five PCs using the procedure described by equation (25). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. The test asset groups are based on portfolio sets are based portfolios obtained from Kenneth French's webpage and described in Online Appendix Table OA.1. The "all" group uses all portfolios as test assets. Panel 1 plots the standard deviation of market betas of each portfolio group using each respective market proxy. Panel 2 plots the cross-sectional-regression-implied market risk premium from each set of test assets. Panel 3 plots the standard deviation of the risk premia for portfolios in each test asset set implied by each model. Panel 4 plots R-squared values implied by each cross-sectional regression. All data are from the 1963-2021 sample.

Figure 9: Asset pricing statistics by factor model



Note. This figure plots comparisons of asset pricing statistics implied by the VMF and the DMF. In this case we use portfolio returns that are hedged with respect to factors from popular factor models rather than PCs as in Figure 8, but still according to equation (25). In all cases, we remove the market factor from each model and replace it with the VMF or the DMF. In results "ex Size" we remove the size factor from each model (except for the DHS model, which does not include a size factor). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. We use all 372 value-weighted portfolios as test assets, which were obtained from Kenneth French's webpage and described in Online Appendix Table OA.1. Panel 1 plots the cross-sectional-regression-implied market risk premium from each factor model. Panel 2 plots the standard deviation of the risk premia for portfolios in each test asset set implied by each factor model. Panel 3 plots r-squared values implied by each cross-sectional regression. Panel 4 plots the mean absolute pricing error (α) from time series regressions of portfolio returns on all factors using each respective market proxy for each factor model. All data are from the 1963-2021 sample.

Table 1: Loadings of value- and equal-weighted market indexes on cross-sectional PCs

Variable	Sharpe ratio	(1)	(2)	(3)	(4)	(5)	$\beta_{PC_i} \cdot \mathbb{E}[PC_i]$
Panel A: VW CRSP index excess returns (VMF)							
VMF	0.44***						6.79***
	(0.14)						(2.00)
PC1	0.52***	4.07***	4.07***	4.07***	4.07***	4.07***	7.31***
	(0.14)	(0.10)	(0.07)	(0.06)	(0.06)	(0.03)	(1.79)
PC2	0.04		1.29***	1.29***	1.29***	1.29***	0.17
	(0.13)		(0.07)	(0.06)	(0.06)	(0.04)	(0.60)
PC3	0.80***			-0.58***	-0.58***	-0.58***	-1.62**
	(0.13)			(0.06)	(0.06)	(0.05)	(0.73)
PC4	0.17				0.47***	0.47***	0.27
	(0.14)				(0.06)	(0.05)	(0.29)
PC5	0.65***					0.52***	1.17
	(0.15)					(0.04)	(0.76)
α		-0.53	-0.70	0.93*	0.65	-0.52	
		(0.77)	(0.55)	(0.51)	(0.48)	(0.37)	
R^2 (univariate)		0.843	0.085	0.017	0.011	0.014	
R^2 (multivariate)		0.843	0.928	0.945	0.957	0.970	
Panel B: EW CRSP index excess returns							
EW index	0.48***						9.36***
	(0.13)						(2.51)
PC1	0.52***	5.47***	5.47***	5.47***	5.47***	5.47***	9.82***
	(0.14)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)	(2.46)
PC2	0.04		-0.91***	-0.91***	-0.91***	-0.91***	-0.12
	(0.13)		(0.04)	(0.04)	(0.05)	(0.05)	(0.41)
PC3	0.80***			-0.12***	-0.12***	-0.12***	-0.34
	(0.13)			(0.03)	(0.03)	(0.03)	(1.02)
PC4	0.17				0.22***	0.22***	0.13
	(0.14)				(0.03)	(0.03)	(0.22)
PC5	0.65***					0.19***	0.44
	(0.15)					(0.03)	(1.02)
α		-0.46	-0.34	0.00	-0.13	-0.56*	
		(0.48)	(0.31)	(0.31)	(0.29)	(0.29)	
R^2 (univariate)		0.958	0.027	0.000	0.002	0.001	
R^2 (multivariate)		0.958	0.984	0.985	0.986	0.987	

Note: This table provides results from regressions from the unadjusted index excess return on PCs. PCs are signed to have positive average returns and scaled to have unit standard deviation. Panel A provides results when using the CRSP value-weighted market and Panel B provides results when using the CRSP equal-weighted market. The first column provides annualized Sharpe ratios for each variable. Columns 2-6 provide results from multivariate regressions of monthly excess market returns on the PCs. Coefficients are normalized such that they represent the percent increase in the excess market return induced by a one standard deviation increase in the respective PCs. The last column provides the contribution of each PC to the average market return (annualized, %). The first row in the last column provides the average excess market return for comparison (annualized, %). The third to last rows in each panel provide the alphas from each regression (annualized, %). The last two rows provide r-squared values from either a univariate or multivariate regression of the excess market return on PCs. Standard errors are in parenthesis and are computed using the bootstrap procedure described in Online Appendix OA.11. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table 2: Summary statistics of excess index returns and IFFs

Variable	Mean	Stdev.	Skew.	Kurt.	Median	IQR	Rsq
Panel A: 1963-2021 sample							
VMF	6.79	15.36	-0.56	5.07	11.12	18.70	
DMF (1PC)	7.31	14.10	-0.35	5.83	9.69	15.98	0.84
IFF (1PC)	-0.53	6.09	-0.37	5.22	0.52	6.85	
DMF (EW)	6.43	13.29	-0.23	6.10	8.78	14.71	0.75
IFF (EW)	0.36	7.70	-0.33	6.00	1.42	8.23	
Panel B: 1963-1992 sample							
VMF	4.81	15.59	-0.40	5.41	6.60	19.35	
DMF (1PC)	6.32	14.34	-0.30	5.99	7.92	16.71	0.85
IFF (1PC)	-1.51	6.14	-0.24	5.72	-0.94	6.03	
DMF (EW)	5.57	13.34	-0.14	6.23	6.92	14.79	0.73
IFF (EW)	-0.77	8.08	0.02	5.76	-0.82	8.64	
Panel C: 1993-2021 sample							
VMF	8.80	15.12	-0.73	4.73	14.74	18.22	
DMF (1PC)	8.26	13.86	-0.39	5.67	13.25	14.59	0.84
IFF (1PC)	0.54	6.05	-0.52	4.76	2.26	7.30	
DMF (EW)	7.42	13.28	-0.34	5.74	10.81	14.24	0.77
IFF (EW)	1.39	7.23	-0.76	7.18	2.53	8.01	

Note: This table provides summary statistics for monthly excess returns on the VMF, DMF variants constructed using either the first PC (constructed from equal-weighted portfolios described in Online Appendix Table OA.1) or the CRSP equal-weighted index, respectively, and the respective IFFs. The VMF is the CRSP value-weighted market index and the DMF variants are obtained by projecting the VMF onto the either the first PC or the CRSP equal-weighted market index, as indicated in the table. The PCs are constructed from equal-weighted portfolios described in Online Appendix Table OA.1. The IFFs are obtained by subtracting the respective DMF variants from the VMF. All returns are in excess of the risk-free rate. We report the mean, standard deviation, skewness, kurtosis for these objects using the 1963-2021 sample (Panel A), the 1963-1992 sub-sample (Panel B), and the 1993-2021 sub-sample (Panel C). Given the high skewness and excess kurtosis of these distributions, we also report median and interquartile range (IQR) as robust measures of location and scale. We also report r-squared values, which are obtained from the regression of the VMF on each DMF proxy over each sample. All statistics are annualized and in percent. We annualize the Mean and Median (Standard Deviation and IQR) statistics by multiplying monthly frequency values by 12 ($\sqrt{12}$).

Table 3: Risk-return tradeoff coefficient (γ) across GARCH models (1963-2021)

Risk defn.	Parameter	VMF	DMF (1PC)	DMF (EW)
Panel A: GARCH				
Variance	γ	2.47	7.98***	7.01***
(σ_t^2)		(2.01)	(2.61)	(2.54)
	$\gamma^{VMF} - \gamma^{DMF}$		-5.51	-4.53
Volatility	γ	0.19	0.68***	0.56***
(σ_t)		(0.18)	(0.22)	(0.20)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.49	-0.38
Panel B: GJR-GARCH				
Variance	γ	1.01	4.17**	5.29***
(σ_t^2)		(1.88)	(1.67)	(1.94)
	$\gamma^{VMF} - \gamma^{DMF}$		-3.16	-4.28
Volatility	γ	0.15	0.44***	0.49***
(σ_t)		(0.19)	(0.16)	(0.17)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.29	-0.34
Panel C: E-GARCH				
Variance	γ	1.61	6.17***	5.78**
(σ_t^2)		(1.88)	(2.08)	(2.26)
	$\gamma^{VMF} - \gamma^{DMF}$		-4.56	-4.17
Volatility	γ	0.18	0.46***	0.39**
(σ_t)		(0.17)	(0.17)	(0.17)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.28	-0.22

Note: This table compares γ estimates and standard errors (in parentheses) from the VMF with those from DMF variants. The first column corresponds with results using the VMF, which is the CRSP value-weighted index. Columns two and three correspond to results from DMF variants that are constructed using either the first PC from the PCA on equal-weighted portfolios described in Online Appendix Table OA.1 or the CRSP equal-weighted market index. All returns are in excess of the risk-free rate. We consider three popular GARCH volatility forecasting models such as a GARCH (1,1) (Panel A), a Threshold GARCH model of Glosten et al. (1993) (Panel B) and an exponential GARCH model of Nelson (1991) (Panel C). Within each panel, the first (second) set of results is based on the conditional variance (volatility) as a proxy for risk. Rows labeled $\gamma^{VMF} - \gamma^{DMF}$ report the difference between γ's estimated using the VMF and the respective DMF variants. We report outer product standard errors (Hamilton, Time Series Analysis, 1994, pp. 133-148) in parentheses. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table 4: Difference between VMF and DMF betas for size × beta-sorted portfolios

Market beta quintile	Size quintile Small	2	3	4	Large	Small - Large
Low	-0.23	-0.15	-0.08	-0.02	0.08	-0.31
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
2	-0.25	-0.19	-0.13	-0.06	0.06	-0.31
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
3	-0.30	-0.21	-0.15	-0.09	0.03	-0.32
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)
4	-0.31	-0.24	-0.19	-0.12	-0.02	-0.29
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
High	-0.40	-0.31	-0.23	-0.17	-0.05	-0.35
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
High-Low	-0.17	-0.15	-0.15	-0.15	-0.14	
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	

Note: This table reports estimates of the bias induced in market betas associated with the presence of the IFF in cross-sectional asset pricing tests, which tests the predictions of equation (27). In particular, this table reports $\beta^{VMF} - \beta^{DMF}$. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Our model predicts this difference is increasing in β and decreasing in size, consistent with the results in the table. Estimates are from monthly returns on 25 size-by-beta-sorted portfolios from Kenneth French's data library. Standard errors are in parentheses and are computed using the bootstrap procedure described in Online Appendix OA.11. All data are from the 1963-2021 sample.

Table 5: Comparing VMF and DMF asset pricing statistics

	Number of orthogonalization PCs					
	0	3	5	7	10	15
Panel A: Market risk premium estimates (λ_1)						
DMF	1.05	3.74*	5.75***	5.60***	5.40***	5.64***
$[\mathbb{E}\left[r_{t+1}^{DMF}\right] = 7.31$, se $= 1.83]$	(2.34)	(2.04)	(1.91)	(1.88)	(1.90)	(1.88)
VMF	-0.59	4.10*	4.86**	4.67**	4.60**	4.80**
$[\mathbb{E}\left[r_{t+1}^{VMF}\right] = 6.79$, se $= 2.00]$	(3.20)	(2.28)	(2.22)	(2.17)	(2.11)	(2.09)
DMF - VMF	1.64	-0.36	0.89	0.93	0.80	0.84
$[\mathbb{E}\left[r_{t+1}^{DMF} - r_{t+1}^{VMF}\right] = 0.53$, se $= 0.79]$	(1.12)	(1.38)	(1.02)	(1.01)	(0.98)	(0.95)
$p\,(DMF < VMF)$	0.06	0.62	0.22	0.21	0.24	0.21
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)						
DMF	0.30	1.06*	1.63***	1.59***	1.53***	1.60***
$[std\left(\beta_p^{DMF}\right) = 0.28$, se $= 0.01]$	(0.43)	(0.56)	(0.56)	(0.55)	(0.55)	(0.55)
VMF	0.11	0.77*	0.91**	0.88**	0.86**	0.90**
$[std\left(\beta_p^{VMF}\right) = 0.19$, se $= 0.01]$	(0.38)	(0.41)	(0.41)	(0.40)	(0.39)	(0.39)
DMF - VMF	0.19	0.29	0.72**	0.71**	0.66**	0.70**
$[std\left(\beta_p^{DMF}\right) - std\left(\beta_p^{VMF}\right) = 0.10$, se $= 0.01]$	(0.39)	(0.34)	(0.28)	(0.28)	(0.28)	(0.27)
$p\,(DMF < VMF)$	0.45	0.22	0.01	0.01	0.01	0.01
Panel C: SML R-squared						
DMF	0.01	0.21	0.46***	0.43***	0.48***	0.52***
	(0.08)	(0.14)	(0.16)	(0.16)	(0.17)	(0.17)
VMF	0.00	0.12	0.20*	0.18	0.21*	0.23*
	(0.07)	(0.09)	(0.12)	(0.11)	(0.12)	(0.13)
DMF - VMF	0.01	0.09	0.26**	0.26**	0.27**	0.29**
	(0.07)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)
$p\,(DMF < VMF)$	0.45	0.23	0.01	0.01	0.01	0.01
Panel D: Mean absolute pricing error (α)						
DMF	2.06***	1.55***	1.30***	1.43***	1.28***	1.23***
	(0.25)	(0.14)	(0.09)	(0.09)	(0.09)	(0.08)
VMF	2.61***	1.40***	1.29***	1.41***	1.28***	1.25***
	(0.41)	(0.10)	(0.08)	(0.08)	(0.09)	(0.07)
DMF - VMF	-0.55	0.16*	0.01	0.02	0.00	-0.02
	(0.36)	(0.11)	(0.03)	(0.04)	(0.02)	(0.02)
$p\,(DMF < VMF)$	0.90	0.03	0.33	0.23	0.46	0.77

Note: This table presents asset pricing statistics when using all value-weighted portfolios as test assets and hedging their returns with respect to different numbers of PCs using the hedging procedure described by equation (25). We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports time series averages of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$), the VMF ($\mathbb{E}[r_{t+1}^{VMF}]$), and their difference. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. The first column in Panel B also reports the standard deviation of portfolio exposures to the DMF ($std(\beta_p^{DMF})$), the VMF ($std(\beta_p^{VMF})$), and their difference. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that use either the DMF or VMF, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported DMF statistics are less than the VMF statistics based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table 6: Comparing DMF asset pricing statistics when controlling for IFF

	Number of orthogonalization PCs					
	0	3	5	7	10	15
Panel A: Risk premium estimates (λ)						
DMF (univariate)	1.05	3.74*	5.75***	5.60***	5.40***	5.64***
$\left[\mathbb{E}\left[r_{t+1}^{DMF}\right] = 7.31, \text{se} = 1.83\right]$	(2.34)	(2.04)	(1.91)	(1.88)	(1.90)	(1.88)
DMF (bivariate)	-2.78	6.23***	6.83***	6.85***	6.63***	6.98***
	(2.38)	(1.93)	(1.96)	(1.92)	(1.86)	(1.85)
IFF	-2.73***	0.42	0.48	0.67	0.47	0.71
	(0.99)	(0.87)	(0.84)	(0.83)	(0.80)	(0.79)
Univar. - Bivar.	3.83**	-2.49***	-1.08**	-1.25***	-1.23***	-1.34***
	(1.39)	(0.80)	(0.48)	(0.48)	(0.38)	(0.36)
$p\,(Univar. < Bivar.)$	0.01	1.00	0.99	1.00	1.00	1.00
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)						
DMF (univariate)	0.30	1.06*	1.63***	1.59***	1.53***	1.60***
$\left[std\left(\beta_p^{Uni.}\right) = 0.28, \text{se} = 0.01\right]$	(0.43)	(0.56)	(0.56)	(0.55)	(0.55)	(0.55)
DMF (bivariate)	0.79	1.77***	1.93***	1.94***	1.88***	1.98***
$\left[std\left(\beta_p^{Bivar.}\right) = 0.28, \text{se} = 0.01\right]$	(0.57)	(0.55)	(0.56)	(0.55)	(0.53)	(0.53)
Univar. - Bivar.	-0.49	-0.70***	-0.31**	-0.36**	-0.35**	-0.38***
	(0.77)	(0.24)	(0.15)	(0.15)	(0.13)	(0.12)
$p\,(Univar. < Bivar.)$	0.66	1.00	0.98	0.99	0.99	1.00
Panel C: SML R-squared						
DMF (univariate)	0.01	0.21	0.46***	0.43***	0.48***	0.52***
	(0.08)	(0.14)	(0.16)	(0.16)	(0.17)	(0.17)
DMF (bivariate)	0.23**	0.41***	0.51***	0.48***	0.55***	0.57***
	(0.12)	(0.15)	(0.15)	(0.15)	(0.16)	(0.16)
Univar. - Bivar.	-0.22***	-0.19***	-0.05	-0.05	-0.06	-0.05
	(0.10)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)
$p\,(Univar. < Bivar.)$	1.00	1.00	0.81	0.83	0.87	0.82
Panel D: Mean absolute pricing error (α)						
DMF (univariate)	2.06***	1.55***	1.30***	1.43***	1.28***	1.23***
	(0.25)	(0.14)	(0.09)	(0.09)	(0.09)	(0.08)
DMF (bivariate)	2.14***	1.40***	1.29***	1.41***	1.28***	1.25***
	(0.25)	(0.10)	(0.08)	(0.08)	(0.09)	(0.07)
Univar. - Bivar.	-0.08	0.16*	0.01	0.02	0.00	-0.02
	(0.11)	(0.11)	(0.03)	(0.04)	(0.02)	(0.02)
$p\,(Univar. < Bivar.)$	0.69	0.03	0.33	0.23	0.46	0.77

Note: This table presents asset pricing statistics controlling for the IFF and using all value-weighted portfolios as test assets and hedging their returns with respect to different numbers of PCs using the hedging procedure described by equation (25). We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Rows labeled as "Univariate" include only the DMF as a factor (constructed using one PC), whereas rows labeled "Bivariate" include the IFF as a control. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports time series averages of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$). Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. The first column in Panel B also reports the standard deviation of portfolio exposures to the DMF in the univariate model ($std(\beta_p^{Univar.})$) and that from the bivariate model ($std(\beta_p^{Bivar.})$). Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that use either the univariate or the bivariate model, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported univariate model statistics are less than the bivariate model statistics based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table 7: Comparing VMF and DMF asset pricing statistics (with factor model controls)

	Factor model control							
	None	FF3	FF3C	FF5	q4	SY	DHS	q5
Panel A: Market risk premium estimates (λ_1)								
DMF	5.75***	5.98***	5.87***	6.54***	6.27***	5.34***	5.87***	7.02***
$[\mathbb{E}\left[r_{t+1}^{DMF}\right] = 7.31$, se $= 1.83]$	(1.91)	(1.86)	(1.88)	(1.84)	(1.94)	(2.02)	(2.08)	(1.93)
VMF	4.86**	5.23**	4.91**	6.13***	5.66**	3.98*	5.20**	6.45***
$[\mathbb{E}\left[r_{t+1}^{VMF}\right] = 6.79$, se $= 2.00]$	(2.22)	(2.15)	(2.16)	(2.10)	(2.22)	(2.32)	(2.38)	(2.21)
DMF - VMF	0.89	0.74	0.96	0.41	0.61	1.36	0.67	0.57
$[\mathbb{E}\left[r_{t+1}^{DMF} - r_{t+1}^{VMF}\right] = 0.53$, se $= 0.79]$	(1.02)	(0.93)	(0.92)	(0.86)	(0.95)	(1.01)	(1.10)	(0.94)
$p\,(DMF < VMF)$	0.22	0.24	0.17	0.33	0.27	0.10	0.29	0.28
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)								
DMF	1.63***	1.69***	1.66***	1.85***	1.77***	1.51***	1.67***	1.98***
$[std\left(\beta_p^{DMF}\right) = 0.28$, se $= 0.01]$	(0.56)	(0.55)	(0.55)	(0.54)	(0.57)	(0.59)	(0.61)	(0.57)
VMF	0.91**	0.98**	0.92**	1.15***	1.07**	0.75*	0.98**	1.21***
$[std\left(\beta_p^{VMF}\right) = 0.19$, se $= 0.01]$	(0.41)	(0.40)	(0.40)	(0.40)	(0.42)	(0.41)	(0.45)	(0.42)
DMF - VMF	0.72**	0.71***	0.74***	0.70***	0.71**	0.76**	0.69**	0.77***
$[std\left(\beta_p^{DMF}\right) - std\left(\beta_p^{VMF}\right) = 0.10$, se $= 0.01]$	(0.28)	(0.26)	(0.26)	(0.24)	(0.27)	(0.30)	(0.30)	(0.26)
$p\,(DMF < VMF)$	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00
Panel C: SML R-squared								
DMF	0.46***	0.48***	0.51***	0.51***	0.49***	0.41**	0.47***	0.54***
	(0.16)	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)	(0.14)
VMF	0.20*	0.22*	0.23*	0.28**	0.25**	0.14	0.22*	0.30**
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)
DMF - VMF	0.26**	0.25***	0.28***	0.23***	0.24**	0.27**	0.24**	0.24***
	(0.08)	(0.07)	(0.08)	(0.06)	(0.07)	(0.09)	(0.08)	(0.06)
$p\,(DMF < VMF)$	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00
Panel D: Mean absolute pricing error (α)								
DMF	1.30***	1.29***	1.20***	1.36***	1.33***	1.39***	1.32***	1.31***
	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)	(0.11)	(0.10)	(0.11)
VMF	1.29***	1.29***	1.20***	1.29***	1.31***	1.37***	1.32***	1.28***
	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.10)	(0.09)	(0.10)
DMF - VMF	0.01	0.01	0.00	0.07	0.02	0.02	0.01	0.04
	(0.03)	(0.04)	(0.04)	(0.06)	(0.05)	(0.03)	(0.04)	(0.06)
$p\,(DMF < VMF)$	0.33	0.40	0.46	0.11	0.21	0.17	0.40	0.14

Note: This table presents asset pricing statistics when using all value-weighted portfolios as test assets and hedging their returns with respect to five PCs as well as factors from prominent models using the hedging procedure described by equation (25). We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Each column corresponds to controlling for a factors in different models (where the acronyms are defined in the text) in addition to the five PCs. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports time series averages of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$), the VMF ($\mathbb{E}[r_{t+1}^{VMF}]$), and their difference. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. The first column in Panel B also reports the standard deviation of portfolio exposures to the DMF ($std(\beta_p^{DMF})$), the VMF ($std(\beta_p^{VMF})$), and their difference. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that use either the DMF or VMF, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported DMF statistics are less than the VMF statistics based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table 8: Size factor spanning tests

	Factor model					
	FF3	FF3C	FF5	q4	SY	q5
Panel A: 1963-2021 sample						
α^{DMF}	-0.41	-1.38	0.14	0.78	0.85	1.60
	[-0.37]	[-1.29]	[0.12]	[0.53]	[0.68]	[1.03]
β^{DMF}_{Size}	0.43***	0.46***	0.42***	0.41***	0.46***	0.39***
	[18.27]	[17.38]	[14.91]	[12.72]	[13.86]	[11.04]
α^{VMF}	1.44	1.47	3.26**	4.56***	6.47***	7.39***
	[1.08]	[1.11]	[2.45]	[2.79]	[4.40]	[4.23]
β^{VMF}_{Size}	0.19***	0.19***	0.15***	0.14***	0.09**	0.09**
	[6.16]	[6.02]	[4.32]	[4.07]	[2.32]	[2.41]
Panel B: 1963-1992 sample						
α^{DMF}	-0.24	0.27	1.12	3.43*	2.57	3.41*
	[-0.17]	[0.18]	[0.71]	[1.92]	[1.48]	[1.73]
β^{DMF}_{Size}	0.49***	0.49***	0.47***	0.43***	0.48***	0.43***
	[17.43]	[17.63]	[13.87]	[11.34]	[12.50]	[10.68]
α^{VMF}	1.85	3.08	5.31***	8.34***	8.73***	10.06***
	[1.06]	[1.54]	[2.79]	[3.69]	[4.21]	[4.18]
β^{VMF}_{Size}	0.25***	0.24***	0.18***	0.18***	0.11**	0.16***
	[5.54]	[5.38]	[3.73]	[3.79]	[1.96]	[3.08]
Panel C: 1993-2021 sample						
α^{DMF}	-1.61	-3.16*	-0.44	-0.46	-0.89	0.97
	[-0.92]	[-1.94]	[-0.26]	[-0.22]	[-0.49]	[0.44]
β^{DMF}_{Size}	0.41***	0.49***	0.34***	0.39***	0.46***	0.35***
	[11.40]	[10.31]	[8.34]	[7.04]	[8.67]	[5.81]
α^{VMF}	0.25	-0.11	3.06*	3.44	4.84**	6.35***
	[0.13]	[-0.06]	[1.68]	[1.53]	[2.38]	[2.74]
β^{VMF}_{Size}	0.17***	0.18***	0.06	0.08	0.06	0.02
	[4.29]	[4.12]	[1.29]	[1.41]	[1.42]	[0.37]

Note: This table presents factor spanning tests for the size factors in each factor model when either the VMF or the DMF is used as the market factor in each model. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. α^{DMF} (α^{VMF}) represents the alpha when the DMF (VMF) is used, and is annualized and in percent. β^{DMF}_{Size} (β^{VMF}_{Size}) represents the size-factor market βs when the DMF (VMF) is used. Panels A, B, and C display results from the 1963-2021, 1963-1992, and 1993-2021 samples, respectively. T-statistics are reported in brackets and are based on the bootstrap procedure described in Online Appendix OA.11. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively.

Table 9: IFF spanning tests

	Factor model					
	FF3	FF3C	FF5	q4	SY	q5
Panel A: VMF as the market factor						
α (w/ Size)	-0.31	-1.34***	-0.51	-1.40***	-1.24**	-2.06***
	[-0.99]	[-4.50]	[-1.37]	[-2.74]	[-2.55]	[-4.03]
β_{IFF}^{VMF} (w/ Size)	0.23***	0.25***	0.24***	0.25***	0.28***	0.26***
	[22.48]	[29.18]	[24.85]	[20.75]	[28.52]	[22.36]
β_{IFF}^{Size} (w/ Size)	-0.48***	-0.48***	-0.48***	-0.43***	-0.48***	-0.42***
	[-17.54]	[-21.43]	[-25.73]	[-14.19]	[-25.11]	[-13.57]
α (w/o Size)	-1.00	-2.04***	-2.09***	-3.37***	-4.33***	-5.18***
	[-1.44]	[-2.74]	[-2.95]	[-4.29]	[-4.97]	[-6.23]
β_{IFF}^{VMF} (w/o Size)	0.14***	0.16***	0.17***	0.19***	0.24***	0.22***
	[7.74]	[9.33]	[9.46]	[11.30]	[10.44]	[12.54]
Panel B: DMF as the market factor						
α (w/ Size)	-0.06	-1.39***	-0.16	-1.08	-0.58	-1.49**
	[-0.16]	[-3.63]	[-0.35]	[-1.59]	[-0.90]	[-2.21]
β_{IFF}^{DMF} (w/ Size)	0.24***	0.28***	0.25***	0.25***	0.28***	0.26***
	[11.16]	[17.07]	[12.64]	[9.28]	[15.06]	[9.63]
β_{IFF}^{Size} (w/ Size)	-0.57***	-0.60***	-0.59***	-0.52***	-0.60***	-0.52***
	[-14.24]	[-17.09]	[-20.00]	[-10.77]	[-21.91]	[-10.69]
α (w/o Size)	0.17	-0.57	-0.24	-1.48*	-1.09	-2.32**
	[0.23]	[-0.71]	[-0.32]	[-1.71]	[-1.09]	[-2.51]
β_{IFF}^{VMF} (w/o Size)	-0.01	0.01	0.00	0.04*	0.01	0.06**
	[-0.60]	[0.46]	[0.05]	[1.87]	[0.25]	[2.56]

Note: This table presents factor spanning tests for the IFF using each factor model when either the VMF or the DMF is used as the market factor in each model. Panel 1 (Panel 2) presents results when the VMF (DMF) is used as the market factor. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Within each panel, we run test using models that include their size factors (designated as "w/ Size") and that exclude their size factors (designated as "w/o Size"). α is annualized and in percent. β_{IFF}^{VMF} (β_{IFF}^{DMF}) represents IFF market βs when the VMF (DMF) is used as the market proxy. β_{IFF}^{Size} represents IFF exposure to size factors (in models where we include the size factors). T-statistics are reported in brackets and are based on the bootstrap procedure described in Online Appendix OA.11. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Online Appendix (For Online Publication Only)

"Idiosyncratic financial risk and a reevaluation of the risk-return tradeoff"

Sung Je Byun, Johnathan A. Loudis, and Lawrence D.W. Schmidt

OA.1 Summary of additional results and robustness

In this section, we summarize figures and tables in the Online Appendix related to additional results and robustness checks referenced in the main text. Please see references to these figures and tables in the main text for more context, explanation, and interpretation.

OA.1.1 Figures

- Figure OA.1 presents IRF results similar to those in Figure 4 from the main text, except in this case data spans 1963-2021 (i.e., including the COVID pandemic) rather than 1963-2019 as in the main text.

- Figure OA.2 presents IRF results similar to those in Figure 4 from the main text, except in this case we provide results when the CRSP value-weighted market is projected on up to the first 10 PCs.

- Figure OA.3 presents IRF results similar to those in Figure 4 from the main text, except in this case we provide results when the CRSP value-weighted market is projected on up to the first 10 PCs and the projection is done in a rolling fashion.

- Figure OA.4 presents IRF results similar to those in Figure 4 from the main text, except in this case we provide results when the CRSP value-weighted market is projected on the first five PCs, the projection is done in a rolling fashion, and we provide IRF estimates for the GR (which is estimated as the residual from these projections).

- Figure OA.5 compares portfolio betas estimated using different proxies for the market portfolio to those estimated using the VMF. The main point is that the differences in betas computed using the DMF relative to those using the VMF is driven by the first PC, and that a projection of the VMF onto PCs yields beta estimates similar to those using the VMF itself (as could be inferred from results in Table 1 where we show that the VMF is explained well by a small number of PCs). Additionally, market risk (as measured with respect to the VMF) can largely be explained by a projection of VMF onto PCs rather than by the residual from those projections, which represents an estimate of the granular residual. The intuition here is that the granular residual is "small" (see Table OA.2) and therefore does not affect estimates of VMF-based market risk much.

- Figure OA.6 (Panel 1) provides a visualization of results from Table 2. We plot kernel density estimates of the VMF and the IFF as estimated from our baseline specification with one PC as well as when using the equal-weighted index as an instrument. Panels 2 and 3 in Figure OA.6 plot the log of the estimated volatilities and their differences when using either the VMF or the DMF under the GARCH(1,1) specification. As expected from the unconditional analysis in Table 2, the conditional volatility is typically lower for the DMF compared to the VMF. As is clear from Panel 3, the difference exhibits substantial variation at both high and low frequencies. This nontrivial difference lends empirical support to our conjecture that the IFF adds noise to the lagged forecasting errors used to generate the conditional variance forecasts within the GARCH forecasting models.

- Figure OA.7 presents beta gap and SML plots similar to those in Figures 6 and 7, respectively, from the main text, except in this case we provide results for the 25 size-by-book-to-market-sorted portfolios. Panel 1 shows the beta gap plot, which demonstrates results similar to those from Figure 6 when using size-by-book-to-market-sorted portfolios. In Panel 2a, which shows SMLs constructed without hedging returns of their exposure to g_{t+1}, the VMF-based SML has a negative slope indicating a negative risk-return relationship. Contrarily, the DMF-based SML has a positive slope and portfolios' expected returns move closer to the theory-implied SML.[56] Once again, this happens because most expected returns lie above the theoretical SML and most of the β_p^{DMF} are noticeably larger than their β_p^{VMF} counterparts. All β_p^{DMF} shift to the right relative to their corresponding β_p^{VMF} (aside from a few from the largest size category) and the shift is usually larger for the smaller portfolios, consistent with our theory. Hedging for exposure to g_{t+1} in Panel 2b makes the red and blue lines more similar to each other and to the theory-implied SML, as with results in Figure 7 from the main text.

- Figure OA.8 presents SML plots similar to those in Figure 7 from the main text (for the all-portfolio case), except in this case we provide results using varying numbers of PCs as hedging controls. Each panel uses the same set of all 372 portfolios as test assets, but uses progressively more PCs for constructing the g_{t+1} controls ranging from zero (Panel 1) to 15 (Panel 6) to demonstrate robustness to this choice. In the case where we use no controls (Panel 1), we get a well-known result. Namely, the VMF-based SML is flat. Using the DMF improves this slightly, yielding a small but positive slope. Importantly, using the DMF yields a larger spread in betas than is achieved using the VMF. As we add progressively more PCs to the g_{t+1} construction, the SMLs produced by both the DMF and the VMF conform more with the theory-implied SML. In fact, there are very few qualitative changes in results once five or more PCs are used in the analysis. In all cases, the DMF-based SML maintains a higher slope (market risk premium estimate) than the VMF-based SML despite having a larger spread in betas. This implies that, after controlling for exposure to other factors, g_{t+1}, using the DMF implies that the market factor explains more of the spread in portfolio average returns than when using the VMF. In other words, using the DMF results in the market factor inducing more spread in implied discount rates. This is what allows the DMF-based SML to fit the theory-implied SML despite inducing a higher spread in betas relative to those from the VMF. This need not have been the case, and implies that the DMF is able to explain more variation in average returns compared to the VMF.

[56] This is with the exception of the small-growth portfolio, whose expected returns are notoriously hard to explain (see, e.g., Fama and French (1993))

- Figure OA.9 presents asset pricing statistics similar to those in Figure 8 from the main text, except in this case we construct DMF using the dominant factor from a set of factors based on the first five PCs that have been rotated to satisfy the assumed restrictions in equations (3) and (4).

- Figure OA.10 presents asset pricing statistics related to controlling for factors from prominent factor models similar to those in Figure 9 from the main text, except in this case we also control for the first five PCs (in addition to the factors from the models) when constructing hedged portfolio returns used in the tests.

- Figure OA.11 presents asset pricing statistics related to controlling for factors from prominent factor models similar to those in Figure 9 from the main text, except in this case we use equal-weighted portfolios as test assets.

OA.1.2 Tables

- Table OA.3 presents time series risk-return tradeoff estimates similar to those in Table 3 from the main text, except in this case data spans 1927-2021 rather than 1963-2019 as in the main text. One downside to this estimation is that we have fewer portfolios for the PCA analysis and hence may lose some important contributors to the first principal component.[57] On the other hand, the longer time series allows us to check robustness with respect to dropping some portfolios as well as to test whether our results are robust to the explosion of new stocks that entered the CRSP value-weighted index in the 1960s (see Figure 1). Results for this robustness are similar to those from the 1963-2021 sample reported in Table 3, with a few slight differences. As before, γ values estimated using various the DMF are always larger than those estimated using the VMF. However, the estimated γ values are smaller in magnitude to those estimated from the 1963 sample, which is likely a consequence of high volatility and low realized returns during the Great Depression period. All-in-all, results

[57] In this sample, we estimate the PCA using 103 equal-weighted anomaly portfolios returns that are available in 1927 according to our portfolio summary in Online Appendix Table OA.1. Despite this, the correlation between the first PC constructed from the 1927-2021 sample and that from the 1963-2021 sample (over the latter period) is 0.9995.

from the 1927 sample support the conclusion that using the DMF helps recover a stronger risk-return tradeoff than using the VMF.

- Table OA.4 presents differences between betas estimated using VMF and DMF across size-by-beta-sorted portfolios similar to those in Table 4 from the main text, except in this case data spans 1993-2021 rather than 1963-2019 as in the main text.
- Table OA.5 presents asset pricing statistics similar to those in Table 5 from the main text, except in this case data spans 1993-2021 rather than 1963-2019 as in the main text.
- Table OA.6 presents asset pricing statistics similar to those in Table 5 from the main text, except in this case we use equal-weighted portfolios as test assets rather than value-weighted portfolios as in the main text.
- Table OA.7 presents asset pricing statistics similar to those in Table 7 from the main text, except in this case we only hedge test portfolio of exposure to factors from the models as opposed to both factors from the models and PCs as in the main text.
- Tables OA.8 and OA.9 present asset pricing statistics related to either including or excluding size factors from each model as controls when evaluating using the VMF or the DMF, respectively. These tables provide more details on statistics summarized in Figure 9 from the main text.

OA.2 How large is the granular residual?

In this section, we estimate the GR "size" relative to the VMF and the IFF. Specifically, we estimate the ratios of the GR variance and standard deviation to those from the VMF and the IFF.

Equation (7) provides a framework for estimating the GR from the IFF and our PCs. The IFF can be decomposed into three components according to the equation: 1) time-varying loadings on the DMF ($\tilde{f}_{t+1}$), 2) time-varying loadings on other factors ($\tilde{g}_{t+1}$), and 3) the GR (η_{t+1}). Since, according to our maintained assumptions, $\tilde{f}_{t+1}$ is spanned by the first PC and $\tilde{g}_{t+1}$ are spanned by the remaining PCs, we can estimate

η_{t+1} by projecting the IFF onto PCs and treating the residual as a proxy for η_{t+1}. Using this framework, we estimate the GR in two ways. First, we assume IFF loadings on PCs are constant and run full-sample regressions of the IFF on various numbers of PCs. The residuals are then estimates of the GR. Second, we assume IFF loadings on PCs are time varying and run centered rolling regressions with a 60-month window to estimate the GR each month, t, as the residual from the rolling regression in that month.[58]

Table OA.2 provides GR summary statistics for the full-sample analysis (Panel A) and the rolling analysis (Panel B). Focusing first on the full-sample analysis in Panel A, when only one PC is used in the regression the estimate of the GR is equivalent to the IFF because the IFF is orthogonal to the first PC. In this case, the GR has the same volatility as the IFF (6.09%, annualized), the variance ratio $\sigma^2_{GR}/\sigma^2_{IFF}$ is 100%, and $\sigma^2_{GR}/\sigma^2_{VMF}$ is the same as the $\sigma^2_{IFF}/\sigma^2_{VMF}$ ratio (15.71%). As we increase the number of PCs used in the regressions, the GR size shrinks as more variation in the IFF is attributed to constant loadings on the PCs. To avoid overfitting issues, the maximum number of PCs we use is 15. Results from the full-sample regressions with between 5 and 15 PCs indicate that the GR standard deviation is approximately 2.21-1.65% (annualized), and that its variance (standard deviation) is approximately 2-3% (13-19%) of the VMF variance (standard deviation) and 14-17% (36-43%) of the IFF variance (standard deviation).

Focusing now on the rolling analysis in Panel B, results using one PC to estimate the GR now do not mechanically reproduce the IFF. This is because we are now accounting for the effect of possible time-varying IFF loading on $\tilde{f}_{t+1}$. The fact that $\sigma^2_{GR}/\sigma^2_{IFF} < 100\%$ in this case (it is 89.63%) implies the IFF has some time-varying loading on $\tilde{f}_{t+1}$. Allowing for time-varying coefficients allows us to better estimate the potentially time-varying loadings of the IFF on these PCs, but increases the possibility of overfitting so again limit the maximum number of PCs in the regressions to 15. Results from the rolling regressions with between 5 and 15 PCs indicate that the GR standard deviation is approximately 1.61-2.21% (annualized), and that its variance (standard deviation) is approximately 1-2% (7-13%) of the VMF variance (standard deviation) and 10-14% (27-36%) of the IFF variance (standard deviation).

[58]In any 60-month window, if month t is the month of interest for estimating the GR, the regression is estimated using data from months $t-30$ to $t+29$.

OA.3 PCA: Additional details

In this section, we discuss some additional details related to our PCA described in section 2.3 and the use of PCA more generally as it applies to the study of financial economics.

OA.3.1 PCA details

It is well-known that equity returns follow a strong factor structure (see, for example, Kozak et al., 2018). When asset returns follow a factor structure, it is also well-known that average returns can be described by a linear relationship with factor loadings (Ross 1976, 1977). Given these insights, Chamberlain and Rothschild (1983) and Connor and Korajczyk (1986, 1988) pioneered work related to using PCs to explain average asset returns. Using PCs as factors can be thought of as a statistical and complementary method to what has become the *de facto* method of constructing factors from portfolios sorted on stock characteristics.[59] Indeed, Kozak et al. (2018) conclude that there is nothing special about such factor models and their ability to explain average returns and note that a model with a small number of PCs does about as well explaining average returns as any of the extant factor models they study. Most importantly, asymptotic properties of these estimators are well-understood and one can obtain consistent estimates of common factors with a large number of portfolios and time periods.[60]

[59]Some notable recent extensions of the basic PCA framework include Kelly et al. (2019), who develop an instrumented principal components analysis that helps provide a link between factors constructed from characteristics to those constructed from PCs. Lettau and Pelger (2020b) show how imposing restrictions via their "risk premium PCA" allows their constructed factors to better fit the cross-section of expected returns in addition to explaining time series variation in returns as with a regular PCA. Giglio and Xiu (2021) show how PCs can be used to consistently estimate risk premia on both traded and non-traded factors.

[60]Bai and Ng (2006) proved that errors associated with estimating factors do not affect the limiting distribution of factor-augmented vector autoregression (FAVAR) estimators, which nests our model as a special case. In other words, estimates of factors are consistent and subsequent regression estimates which use estimated principal components have the same asymptotic properties as if the true unobserved factors were used instead. This property holds when $\frac{T}{N} \to 0$ as $N, T \to \infty$. Sample sizes which are used to construct estimates of factors and loadings usually involve a number of stocks which is an order of magnitude larger than the number of time periods, so this condition approximately holds. Even when this condition fails, estimated factors can be treated as data and $\sqrt{T}$-consistent estimates of second stage regression parameters obtain when $\frac{\sqrt{T}}{N} \to 0$, but standard errors need to be corrected to reflect the fact that a generated regressor is used. Stock and Watson (2002) provide conditions under which factors are consistently estimable even when loadings exhibit moderate time-variation.

We choose the 372 portfolios represented by the portfolio sorts described in Online Appendix Table OA.1 for our PCA to balance a few tradeoffs. First, we would like to get as broad of a cross-section of portfolios as possible. Second, we need both equal-and value-weighted versions of the portfolios. The equal-weighted portfolios are used for the PCA, and the value-weighted portfolios are used in our main asset pricing tests (as per standard convention). Third, we would like the portfolios to be as standard and easily accessible as possible. Most portfolios are based on 5x5 double-sorts. However, we include two sets of single-sorted portfolios (size and beta) because our theory has special implications for the characteristics of these portfolios, which are discussed in subsection 6.1. Note that these are a superset of the equity portfolios used for the PCA analysis in Giglio and Xiu (2021).

Strictly speaking, our estimated PCs are a rotation of f_{t+1} and g_{t+1}, which we use for parsimony. However, if we rotate the PCs so as to impose the restrictions that $E[\sum_{i=1}^{N_t} w_{i,t}\beta_i] = 1$ and $E[\frac{1}{N_t}\sum_{i=1}^{N_t} \zeta_i] = 0$ and furthermore assume the true f_{t+1} can be constructed by the first five or fewer PCs, the correlations between these restricted f_{t+1} values and the original first PC range between 0.989 and 0.990. Furthermore, we find the average loadings across all portfolios on higher PCs are close to zero. Additionally, by construction, the first PC is orthogonal to higher PCs, consistent with our assumption that g_{t+1} are orthogonal to f_{t+1}. In other words, the PCs themselves approximately conform to our assumptions about f_{t+1} and g_{t+1}. We provide evidence in Online Appendix Figure OA.12 that the PCs approximately conform to the assumptions about loadings on f_{t+1} and g_{t+1} in equations (3) and (4). Namely, the value-weighted loading of value-weighted portfolios on the DMF are 0.993, and the average loadings of these portfolios on higher PCs are approximately zero.

OA.3.2 Equal-weighted portfolios have a stronger factor structure than value-weighted portfolios

There is a strong factor structure in the returns of characteristics-sorted portfolios, with one dominant factor that can be interpreted as a level factor. The factor structure is even stronger in equal-weighted portfolios relative to value-weighted portfolios.

To help get a sense for the strength of the factor structure in our data, we provide plots of the explained variation in returns as a function of the number of PCs from our main 1963-2021 sample when using equal-weighted portfolios in the PCA in Figure

OA.13 (Panel 1). For a comparison, the plot also shows results from a PCA that uses value-weighted portfolios. The cross-section of equal-weighted portfolios has a stronger factor structure than that of value-weighted portfolios.[61] For instance, in the case of the equal-weighted (value-weighted) portfolio PCA, the first PC explains about 85% (82%) of the variation in excess returns. This result is not surprising given our results from the previous subsections. First, to the extent value-weighted portfolios are plagued by granular residuals, we expect these portfolios to have a weaker factor structure as long as the granular residuals themselves do not have a strong factor structure. Second, if value-weighted portfolios have time-varying loadings on low-order PCs from the equal-weighted portfolio PCA, these would manifest as variation that is unexplained by the low-order PCs. The first PC in both cases places positive weights that are similar in magnitude on all portfolios, which is a desirable property when viewing this as a (scaled) level factor that proxies for the DMF.

Panel 2 provides the fraction of variation in the value-weighted portfolios that can be explained by PCs constructed from the equal-weighted portfolios as a function of the number of PCs used. PCs constructed from equal-weighted portfolios explain a large fraction of variation in the value-weighted portfolios with the first five PCs explaining about 90% of the variation in the value-weighted portfolio returns. However, even all 372 PCs constructed from equal-weighted portfolios cannot explain all the variation in their value-weighted counterparts, which is expected if the value-weighted portfolios contain granular residuals but the equal-weighted portfolios do not.

OA.3.3 Identifying DMF and IFF empirically: Additional details

We construct the empirical DMF by projecting the VMF onto the first PC, which ensures that the DMF satisfies our restriction in equation (3). In other words, we run the following regression

$$VMF_{t+1} = a + b \cdot PC_{1,t+1} + \varepsilon_{VMF,t+1}, \tag{OA.1}$$

[61] That is, for a given number of PCs, the equal-weighted PCA is always able to explain more variation in equal-weighted portfolios than the value-weighted PCA is able to explain for value-weighted portfolios. In unreported results, we find this to be a robust result across many different sets of base portfolios used in the PCA.

where $PC_{1,t+1}$ is the first PC. We then define the DMF as

$$DMF_{t+1} \equiv b \cdot PC_{1,t+1}. \tag{OA.2}$$

Given our estimate of the DMF, we construct our empirical proxy for the IFF as

$$IFF_{t+1} \equiv VMF_{t+1} - DMF_{t+1}, \tag{OA.3}$$

which is consistent with the definition of the IFF in equation (7).

The first PC effectively serves as an instrumental variable (IV) for identifying f_{t+1} from the observable VMF and higher PCs are used as proxies for g_{t+1}.[62] Instrumenting the VMF with just the first PC from our equal-weighted portfolio PCA solves two issues in the signal extraction problem implied by equation (7). First, the equal-weighted PCA results in an IV estimate of f_{t+1} that is effectively purged of η_{t+1}. Second, instrumenting with just the first PC limits the effect of time-varying exposures to higher PCs represented by $\zeta_{m,t}$, for which we provide evidence in section 3.1.4.

OA.4 Impulse response functions: Additional details

Suppose the DMF goes up by one standard deviation unit. By how much would an econometrician revise her forecast about future macroeconomic conditions on the basis of that information? Suppose instead that the component of market returns explained by higher PCs increases by one standard deviation unit. By how much, if at all, would these forecasts change? To answer these questions, we follow the method in Schmidt (2016), which is itself related to Jordà (2005).[63] While we refer the reader to those papers for further details, the main result is that if we have a variable Y_t, we can estimate a quantity analogous to an impulse response in a VAR to a shock, z_{t+1}

[62]Strictly speaking, our estimated PCs are a rotation of f_{t+1} and g_{t+1}. However, we discuss how the PCs approximately conform to our assumptions about f_{t+1} and g_{t+1} in Online Appendix section OA.3 and in Online Appendix Figure OA.12.

[63]The identifying assumptions proposed in Schmidt (2016) rely on correct specification of the mean of returns and allow for the conditional expectation of the macro variable Y_{t+h} to potentially be misspecified. This is advantageous given that returns are much closer to random walks than many of the macro variables in the consideration set.

at horizon $h = 1, ..., H$ by running the following regression:

$$Y_{t+h} = \alpha_h + \beta_h z_{t+1} + \gamma_h Y_t + \phi_h(L) Y_t + \epsilon_{t+h} \tag{OA.4}$$

where in our context z_{t+1} is the market return component of interest, $\phi(L)$ is a lag polynomial, and ϵ_{t+h} is the forecast error. The IRF is simply $\{\beta_h\}_{h=1}^{H}$. Our choice of Y variables are log industrial production (IP) growth, log employment growth, initial claims divided by lagged employment, log per capita consumption growth, log compensation growth, the ADS business conditions index (Aruoba et al., 2009), the unemployment rate, log GDP growth, and the Chicago Fed National Activity Index (CFNAI).

To construct the two specifications of z_{t+1} described above, we regress the CRSP value-weighted market excess return (i.e., the VMF) on the DMF (i.e., a scaled version of the first PC) and PCs two through N as follows:

$$r_{t+1} = a + \beta \cdot DMF_{t+1} + \zeta'_{2\to N} \cdot PC_{2\to N,t+1} + \eta_{t+1}\,, \tag{OA.5}$$

where $PC_{2\to N,t+1}$ is a vector of PCs two through N. We estimate this regression in the full-sample with three PCs using data from 1963-2019 in our main specification (Figure 4) on a monthly frequency then aggregate the time series within each quarter to match the observation frequency of the macroeconomic aggregates. We show these results are robust to using data from 1963-2021 (Figure OA.1), using more PCs (Figure OA.2), using a rolling estimation window (Figure OA.3), and for the GR when estimated as the residual from the rolling regression with five PCs (Figure OA.4).

OA.5 Motivating theoretical model

Consider a static exchange economy populated by a continuum (with unit mass) of ex-ante identical investors indexed by $j \in [0, 1]$. Investors choose portfolios to maximize expected utility over terminal consumption C_j and have identical preferences with constant absolute risk aversion $A > 0$: $U_j \equiv -E[\exp(-AC_j)]$.

Investors receive identical endowments of N stocks, each of which pays a stochastic and normally-distributed dividend D_i at the end of the period. Let P_i denote the price of stock i and $1 + r_i = D_i/P_i$ denote its realized gross return. Each stock is available in unit net supply, and there is a risk-free asset available in zero net supply.

Therefore, total financial wealth W is given by the total stock market capitalization, $W = \sum_{i=1}^{N} P_i$.

Stocks $i = 1, \dots, K$ are assumed to be "large" stocks that together account for a share $\overline{v} \in (0, 1)$ of the total stock market capitalization, i.e. $\sum_{i=1}^{K} P_i/W = \overline{v}$. Stocks $i = K+1, \dots, N$ are "small" and account for the remaining share $1 - \overline{v}$ of the total stock market value. For simplicity, all large stocks are assumed to have the same price $P_1 = \cdots = P_K = \overline{v}W/K$, and all small stocks are assumed to have the same price $P_{K+1} = \cdots = P_N = (1-\overline{v})W/(N-K)$.[64] Throughout this section we assume that K is much smaller than N. In the limiting case, we will let the total number of traded stocks N increase to infinity while holding fixed both the number of large stocks K and the share $\overline{v}$ that these stocks contribute to the total stock market capitalization.

Excess returns for all stocks $i = 1, \dots, N$ are assumed to follow a two-factor model:[65]

$$r_i - \mathbb{E}[r_i] = \beta_i \left(f - \mathbb{E}[f]\right) + \zeta_i \left(g - \mathbb{E}[g]\right) + \eta_i \tag{OA.6}$$

where $f \sim N(\mathbb{E}[f], \sigma_f^2)$ and $g \sim N(\mathbb{E}[g], \sigma_g^2)$ are common factors, β_i and ζ_i are the loadings of stock i's return on the common factors, and $\eta_i \sim N(0, \sigma_i^2)$ is an idiosyncratic component that satisfies $\mathbb{E}[\eta_i | f,\, g] = 0$. The common factors themselves are assumed to be tradable excess returns on an asset that is available in zero net supply,[66] and their means ($\mathbb{E}[f]$ and $\mathbb{E}[g]$) will be pinned down by equilibrium conditions. Idiosyncratic components are assumed to be uncorrelated across stocks: $\mathbb{E}[\eta_i \eta_{i'} | f,\, g] = 0$ for all $i \neq i'$. Without loss of generality, f is assumed to be scaled so that

$$\frac{\overline{v}}{K} \sum_{i=1}^{K} \beta_i + \frac{1 - \overline{v}}{N - K} \sum_{i=K+1}^{N} \beta_i = 1, \tag{OA.7}$$

which implies that the loading of the market portfolio's return on f is equal to 1. We

[64] This assumption can be easily relaxed to allow for arbitrary weights of individual large stocks in the market index. However, our main focus is to determine how the pricing of risks differs across large and small stocks, not within the set of large stocks. Expected dividends are assumed to be sufficiently large or small to justify all stock valuations.

[65] This can be generalized to include more g_{t+1} factors without changing the main implications of the model, so we limit the model to having two factors for parsimony.

[66] We ignore agents' holdings of this asset in the portfolio choice problem that follows. Market clearing conditions imply that these holdings must be equal to zero for all agents in equilibrium because agents are ex ante identical and thus choose identical portfolios. Alternatively, we can think of these factors as specific portfolios of the individual assets over which the agents are choosing optimal allocations.

also assume that the average small stock has a loading of zero on g so that[67]

$$\frac{1}{N-K}\sum_{i=K+1}^{N}\zeta_i = 0. \tag{OA.8}$$

The market portfolio is then a function of both the common factors and a weighted average of idiosyncratic components, η_i, of individual stocks that may or may not be diversified away as the number of stocks increases:

$$\begin{aligned} r_m &= \frac{\overline{v}}{K}\sum_{i=1}^{K} r_i + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N} r_i \\ &= \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\mathbb{E}[r_i] + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N}\mathbb{E}[r_i]\right)}_{=\mathbb{E}[r_m]} + \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\beta_i + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N}\beta_i\right)}_{=1}\cdot(f-\mathbb{E}[f]) \\ &\quad + \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\zeta_i +\right)}_{\equiv\zeta_m}\cdot(g-\mathbb{E}[g]) + \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\eta_i + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N}\eta_i\right)}_{\equiv\eta} \\ &= \mathbb{E}[r_m] + f - \mathbb{E}[f] + \zeta_m\cdot(g-\mathbb{E}[g]) + \eta. \end{aligned} \tag{OA.9}$$

Note that we have used the restrictions in equations OA.7 and OA.8 to set the market loading on f to 1 and in the definition of ζ_m. So, in this economy the market is a contaminated proxy for f that also contains terms related to the other factor, g, as well as a potentially undiversified aggregated granular residual, η.

In addition to dividend income, each investor receives labor income $Y_j \equiv Y + \varepsilon_j$, where Y is aggregate labor income and ε_j is an idiosyncratic component that is independent and identically distributed across investors j. Aggregate labor income is given by

$$Y = \overline{Y} + \phi_f(f-\mathbb{E}[f]) + \phi_g(g-\mathbb{E}[g]) + \phi_\eta\frac{\overline{v}}{K}\sum_{i=1}^{K}\eta_i + \xi \tag{OA.10}$$

[67]This assumption means that we can interpret g as an equal-weighted long-short portfolio of small stocks. If the number of small stocks is high relative to large stocks, the average stock does not load on g and we can think of this (approximately) as an equal-weighted long-short portfolio of all assets.

where ξ is a mean zero random variable that is independent of $f, g, \eta_1, \ldots, \eta_N$. Aggregate labor income is thus correlated with both the common factors and the idiosyncratic shocks $\eta_1, \ldots, \eta_K$ to the returns of the large stocks. The idiosyncratic labor income component ε_j has zero mean conditional on all common shocks, $\mathbb{E}[\varepsilon_j | f, g, \eta_1, \ldots, \eta_N, \xi] = 0$. However, the distribution of ε_j across investors can depend on the realization of these common shocks. To accomodate flexible relationships between aggregate shocks and the cross-sectional income distribution, we assume that the conditional moment-generating function $M_\varepsilon(t|f, g, \eta_1, \ldots, \eta_N, \xi) \equiv \mathbb{E}[\exp(t\varepsilon_j)|f, g, \eta_1, \ldots, \eta_N, \xi]$ is exponentially affine in the common shocks (excluding the idiosyncratic components of returns for small stocks):

$$M_\varepsilon(t|f, g, \eta_1, \ldots, \eta_N, \xi) = \exp\left(\gamma_0(t) + \gamma_f(t)f + \gamma_g(t)g + \sum_{i=1}^{K} \gamma_i(t)\eta_i + \gamma_\xi(t)\xi\right) \tag{OA.11}$$

where the coefficients $\gamma_0(t), \gamma_1(t), \ldots, \gamma_K(t), \gamma_f(t), \gamma_g(t), \gamma_\xi(t)$ may depend on $t \in \mathbb{R}$. We can then define "idiosyncratic risk-adjusted" labor income as the following affine function of these common shocks:

$$\begin{aligned}
\tilde{Y} &\equiv Y - \frac{1}{A} \log \mathbb{E}[\exp(-A\varepsilon_j)|f, g, \eta_1, \ldots, \eta_N, \xi] \\
&= \left(\overline{Y} + \phi_f(f - \mathbb{E}[f]) + \phi_g(g - \mathbb{E}[g]) + \phi_\eta \frac{1-\overline{v}}{K} \sum_{i=1}^{K} \eta_i + \xi\right) \\
&\quad - \frac{1}{A}\left(\gamma_0(-A) + \gamma_f(-A)f + \gamma_g(-A)g + \sum_{i=1}^{K} \gamma_i(-A)\eta_i + \gamma_\xi(-A)\xi\right) \\
&\equiv \mathbb{E}\left[\tilde{Y}\right] \cdot \left(1 + \psi_f(f - \mathbb{E}[f]) + \psi_g(g - \mathbb{E}[g]) + \psi_\eta \sum_{i=1}^{K} \eta_i + \psi_\xi \xi\right)
\end{aligned} \tag{OA.12}$$

for constants ψ_f, ψ_g, ψ_η, and ψ_ξ.

Let $\theta^j \in \mathbb{R}^N$ denote the fraction of investor j's total wealth W invested in each stock $i = 1, \ldots, N$, with the remaining net wealth share $1 - \sum_{i=1}^N \theta_i^j$ invested (or borrowed) at the risk-free rate r_f. Terminal consumption is then given by the sum of labor income and financial income:

$$C_j = Y_j + W(1 + r_p(\theta^j)) = Y_j + W\left(\sum_{i=1}^{N} \theta_i^j (1 + r_i) + \left[1 - \sum_{i=1}^{N} \theta_i^j\right](1 + r_f)\right) \tag{OA.13}$$

where $r_p(\theta^j)$ denotes the realized net return on investor j's portfolio conditional on the choice of portfolio θ^j. In Online Appendix OA.6 we show that investor j's portfolio choice problem can be rewritten in the following form:

$$\underset{\theta^j \in \mathbb{R}^N}{\operatorname{argmax}} \{-\mathbb{E}[\exp(-AC_j)]\} = \underset{\theta^j \in \mathbb{R}^N}{\operatorname{argmax}} \left\{ \mathbb{E}[r_p(\theta^j)] - A\mathcal{W}\left(\frac{1-\phi}{2} Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j))\right)\right\} \tag{OA.14}$$

where $\phi \equiv \frac{\mathbb{E}[\tilde{Y}]}{\mathbb{E}[\tilde{Y}]+W} \in (0,1)$ is a measure of the importance of human capital in total (human and financial) wealth, $1 + r_h \equiv \tilde{Y}/\mathbb{E}[\tilde{Y}]$ can be interpreted as a "return" on human capital, and $\mathcal{W} \equiv \mathbb{E}[\tilde{Y}] + W$. Online Appendix OA.6 derives the following first-order condition for θ_i^j, the portfolio share of any stock $i = 1, \ldots, N$:

$$\mathbb{E}[r_i] - r_f = A\mathcal{W}\left[(1-\phi)Cov(r_i, r_p(\theta^j)) + \phi Cov(r_i, r_h)\right] \tag{OA.15}$$

This first-order condition can be interpreted as a simple modification of the traditional CAPM expected return-beta relationship to account for human capital risk. In fact, the traditional CAPM holds in the special case where $\phi = 0$ and human capital represents a negligible share of total wealth. Note that $A\mathcal{W}$ is equivalent to relative risk aversion (RRA).

Investors are identical ex-ante, so in equilibrium they all choose the same optimal portfolio. With a unit mass of investors, market clearing implies that this optimal portfolio is the market portfolio: $\theta_1^{j*} = \cdots = \theta_K^{j*} = \overline{v}/K$, $\theta_{K+1}^{j*} = \cdots = \theta_N^{j*} = (1-\overline{v})/(N-K)$, and $r_m = r_p(\theta^j*)$. Plugging in these values (and noting that $\beta_p(\theta^{j*}) = 1$ and $\zeta_p(\theta^{j*}) = \zeta_m$ by market clearing), we obtain closed-form expressions for equilibrium expected excess returns for each stock:

$$\mathbb{E}[r_i]-r_f = \begin{cases} A\mathcal{W}\left[(1-\phi)\left(\beta_i\sigma_f^2 + \zeta_i\zeta_m\sigma_g^2 + \frac{\overline{v}}{K}\sigma_i^2\right) + \phi\left(\beta_i\psi_f\sigma_f^2 + \zeta_i\psi_g\sigma_g^2 + \psi_\eta\frac{\overline{v}}{K}\sigma_i^2\right)\right] & \text{for } i = 1, \ldots, K \\ A\mathcal{W}\left[(1-\phi)\left(\beta_i\sigma_f^2 + \zeta_i\zeta_m\sigma_g^2 + \frac{1-\overline{v}}{N-K}\sigma_i^2\right) + \phi\left(\beta_i\psi_f\sigma_f^2 + \zeta_i\psi_g\sigma_g^2\right)\right] & \text{for } i = K+1, \ldots, N \end{cases} \tag{OA.16}$$

Throughout we have ignored the fact that the common factors are assumed to be tradable excess returns on assets that are available in zero net supply. Therefore, they satisfy the same pricing equation (OA.15). This equation pins down average returns on f, g, and η:

$$\mathbb{E}[f] = A\mathcal{W}\left(1-\phi+\phi\psi_f\right)\sigma_f^2 \tag{OA.17}$$

$$\mathbb{E}[g] = A\mathcal{W}\left((1-\phi)\zeta_m + \phi\psi_g\right)\sigma_g^2 \tag{OA.18}$$

$$\mathbb{E}[r_\eta] = A\mathcal{W}\left((1-\phi)\sigma_\eta^2 + \phi\psi_\eta\sigma_{\eta,K}^2\right), \tag{OA.19}$$

where r_η is an η-mimicking portfolio, σ_η^2 is the variance of the market-level granular residual, η, in equation (OA.9) and $\sigma_{\eta,K}^2$ is the variance of the value-weighted large stock granular residuals. We can also use equation (OA.15) to pin down the expected excess return on the market:

$$\mathbb{E}[r_m] - r_f = E[f] + \zeta_m \cdot \mathbb{E}[g] + \mathbb{E}[r_\eta], \tag{OA.20}$$

We can then express the equilibrium expected returns for individual stocks as

$$\mathbb{E}[r_i] - r_f = \begin{cases} \beta_i \mathbb{E}[f] + \zeta_i \mathbb{E}[g] + A\mathcal{W}(1 - \phi + \phi\psi_\eta)\frac{\overline{v}}{K}\sigma_i^2 & \text{for } i = 1, \ldots, K \\ \beta_i \mathbb{E}[f] + \zeta_i \mathbb{E}[g] + A\mathcal{W}(1 - \phi)\frac{1-\overline{v}}{N-K}\sigma_i^2 & \text{for } i = K+1, \ldots, N \end{cases} \tag{OA.21}$$

Having derived equilibrium risk-return relationships when investors trade a finite number of stocks, we now consider the limiting case where the number of stocks N increases to infinity, while a fixed number K of large stocks still account for a fixed share $\overline{v} > 0$ of the total stock market capitalization.

Proposition 1. (Expected excess returns)
In the limit as the total number of stocks $N \to +\infty$, equilibrium expected excess returns on all stocks are given by

$$\mathbb{E}[r_i] - r_f = \begin{cases} \beta_i \mathbb{E}[f] + \zeta_i \mathbb{E}[g] + A\mathcal{W}\left(1 - \phi + \phi\psi_\eta\right)\frac{\overline{v}}{K}\sigma_i^2 & \textit{for } i = 1, \ldots, K \\ \beta_i \mathbb{E}[f] + \zeta_i \mathbb{E}[g] & \textit{for } i = K+1, K+2, \ldots \end{cases} \tag{OA.22}$$

Furthermore, the market risk premium is given by

$$\mathbb{E}[r_m] - r_f = E[f] + \zeta_m \cdot \mathbb{E}[g] + \mathbb{E}[r_\eta]. \tag{OA.23}$$

If we specialize equation (OA.10) such that labor income does not load on g or the granular residuals, we have that $\phi_g = 0$, $\phi_\eta = 0$, $\psi_g = 0$, and $\psi_\eta = 0$. In this case, the risk premium expressions become in the limit as $N \to +\infty$):

$$\mathbb{E}[f] = A\mathcal{W}\left(1 - \phi + \phi\psi_f\right)\sigma_f^2, \tag{OA.24}$$

$$\mathbb{E}[g] = A\mathcal{W}\left(\left(1 - \phi\right)\zeta_m\right)\sigma_g^2, \tag{OA.25}$$

$$\mathbb{E}[r_\eta] = A\mathcal{W}\left(1 - \phi\right)\sigma_\eta^2, \tag{OA.26}$$

$$\mathbb{E}[r_i] - r_f = \begin{cases} \beta_i \mathbb{E}[f] + \zeta_i \mathbb{E}[g] + A\mathcal{W}(1-\phi)\,\frac{\overline{v}}{K}\sigma_i^2 & \textit{for } i = 1, \ldots, K \\ \beta_i \mathbb{E}[f] + \zeta_i \mathbb{E}[g], & \textit{for } i = K+1, K+2, \ldots \end{cases} \tag{OA.27}$$

and the market risk premium is still as given in equation (OA.23).

The following two propositions provide testable implications implied by the model, and we provide the related derivations in Online Appendix OA.6.

Proposition 2 (Beta gap)
Let β_p^m and β_p^f represent the betas of a portfolio p of stocks. In the limit as $N \to +\infty$, if the portfolio is comprised of all small stocks in the economy (i.e., $i = K+1, K+2 \ldots$), then the beta gap is given by

$$\beta_S^m - \beta_S^f \;\; = \;\; -\left(\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)\beta_S, \tag{OA.28}$$

where σ_ϵ^2 is the variance of the IFF in the model economy (i.e., $\epsilon \equiv \zeta_m g + r_\eta$ and $\sigma_\epsilon^2 \equiv \zeta_m^2 \sigma_g^2 + \sigma_\eta^2$). Furthermore, if $\beta_S > 0$ (i.e., the empirically relevant case), $\beta_S^m - \beta_S^f < 0$. If the portfolio is comprised of all large stocks in the economy (i.e., $i = 1, 2, \ldots K$), then the beta gap is given by

$$\beta_L^m - \beta_L^f \;\; = \;\; \frac{1}{\overline{\nu}}\left[\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2} - \beta_L \frac{\zeta_m^2 \sigma_g^2 - \overline{\nu}\sigma_i^2}{\sigma_f^2 + \sigma_\epsilon^2}\right]. \tag{OA.29}$$

The sign of the large stock portfolio beta gap is ambiguous and depends on a number of model parameters.

Proposition 3 (Small cap portfolio and IFF alphas)
Let α_j^m and α_j^f represent the alphas of an asset j with respect to a capital asset pricing model (i.e., one-factor model) that uses either the value-weighted market excess return, $r_m - r_f$, or f as its single factor. In the limit as $N \to +\infty$ and assuming the labor income does not load on g or the granular residuals (i.e., $\phi_g = \phi_\eta = \psi_g = \psi_\eta = 0$),

then the alphas on the portfolio of all small stocks in the economy are given by

$$\begin{aligned}\alpha_S^m &= \beta_S \mathbb{E}[f] \left[\frac{\phi \psi_f \sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi+\phi\psi_f)\sigma_\epsilon^2} \right] \\ &\overset{\phi \to 1}{\to} \frac{\sigma_\epsilon^2}{\sigma_f^2+\sigma_\epsilon^2} \beta_S \mathbb{E}[f] \text{ , and}\end{aligned} \tag{OA.30}$$

$$\alpha_S^f = 0. \tag{OA.31}$$

Furthermore, assuming $\beta_S > 0$ *and* $\psi_f > 0$ *(i.e., the empirically-relevant cases),* $\alpha_S^m > 0$*. The alphas on the IRF are given by*

$$\begin{aligned}\alpha_\epsilon^m &= A\mathcal{W} \left(\frac{\sigma_f^2}{\sigma_f^2+\sigma_\epsilon^2} \right) \left((1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi)\zeta_m^2 \sigma_g^2 + (1-\phi)\sigma_\eta^2 \right) \\ &\quad - A\mathcal{W}(1-\phi+\phi\psi_f)\sigma_f^2 \\ &\overset{\phi \to 1}{\to} -A\mathcal{W} \left(\frac{\sigma_\epsilon^2}{\sigma_f^2+\sigma_\epsilon^2} \right) \psi_f \sigma_f^2 \text{ , and}\end{aligned} \tag{OA.32}$$

$$\alpha_\epsilon^f = A\mathcal{W}(1-\phi)\zeta_m^2 \sigma_g^2 + A\mathcal{W}(1-\phi)\sigma_\eta^2 \overset{\phi \to 1}{\to} 0. \tag{OA.33}$$

Furthermore, assuming $\psi_f > 0$ *(i.e., the empirically-relevant case),* $\alpha_\epsilon^m < 0$*.*

OA.6 Details of model derivations

Using the definition of $\tilde{Y}$, we can simplify the calculation of investor j's expected utility by applying the law of iterated expectations conditional on the common shocks:

$$\begin{aligned}
& -E[\exp(-AC_j)] \\
&= -E[\exp(-A[Y_j + W(1+r_p(\theta^j))])] \\
&= -E[\exp(-A[Y + \varepsilon_j + W(1+r_p(\theta^j))])] \\
&= -E\left[E[\exp(-A[Y + \varepsilon_j + W(1+r_p(\theta^j))])|f,\eta,\xi]\right] \\
&= -E\left[\underbrace{\exp(-AY)\cdot E[\exp(-A\varepsilon_j)|f,\eta,\xi]}_{=\exp(-A\tilde{Y})}\cdot\exp(-AW(1+r_p(\theta^j)))\right] \\
&= -E\left[\exp(-A[\tilde{Y} + W(1+r_p(\theta^j))])\right]
\end{aligned} \tag{OA.34}$$

In the last line, the term that replaces C_j within the expectation is a function of only common (not idiosyncratic) shocks, and can be computed as a weighted average of returns on aggregate human capital $r_h \equiv \tilde{Y}/E[\tilde{Y}]$ and financial assets $r_p(\theta^j)$:

$$\begin{aligned}
\tilde{Y} + W(1+r_p(\theta^j)) &= \underbrace{\left(E\left[\tilde{Y}\right]+W\right)}_{\equiv\mathcal{W}}\times\left(\underbrace{\frac{E\left[\tilde{Y}\right]}{E\left[\tilde{Y}\right]+W}}_{\equiv\phi}\times\underbrace{\frac{\tilde{Y}}{E\left[\tilde{Y}\right]}}_{\equiv 1+r_h}+\underbrace{\frac{W}{E\left[\tilde{Y}\right]+W}}_{\equiv 1-\phi}\times(1+r_p(\theta^j))\right) \\
&= \mathcal{W}\left(\phi(1+r_h)+(1-\phi)(1+r_p(\theta^j))\right)
\end{aligned} \tag{OA.35}$$

Moreover, r_h and $r_p(\theta^j)$ are both affine functions of independent normally-distributed shocks; therefore, these returns are also normally distributed and expected utility is given by

$$
\begin{aligned}
& - E\left[\exp(-A[\tilde{Y} + W(1 + r_p(\theta^j))])\right] \\
&= -E\left[\exp(-A\mathcal{W}[\phi(1 + r_h) + (1 - \phi)(1 + r_p(\theta^j))])\right] \\
&= -\exp\left(-A\mathcal{W}(1-\phi)\left(1 + E[r_p(\theta^j)] - \frac{A\mathcal{W}}{2}(1-\phi)Var(r_p(\theta^j)) + A\mathcal{W}\phi Cov(r_h, r_p(\theta^j))\right)\right)\dots \\
&\quad \dots \times \exp\left(-A\mathcal{W}\phi(1 + E[r_h]) + \frac{A^2\mathcal{W}^2}{2}\phi^2 Var(r_h)\right)
\end{aligned}
\tag{OA.36}
$$

where the term appearing in the final line does not depend on investor j's portfolio choice θ^j. Therefore, after applying a monotone transformation to investor j's expected utility function and ignoring additive terms that do not depend on θ^j, we obtain equation (OA.14) of the main text:

$$
\underset{\theta^j \in \mathbb{R}^N}{\operatorname{argmax}} \left\{-E[\exp(-AC_j)]\right\} = \underset{\theta^j \in \mathbb{R}^N}{\operatorname{argmax}} \left\{ E[r_p(\theta^j)] - A\mathcal{W}\left(\frac{1-\phi}{2} Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j))\right)\right\}
$$

Next, we compute moments for portfolio returns. The realized return $r_p(\theta^j)$ on investor j's stock portfolio is related to the common shocks by

$$
\begin{aligned}
r_p(\theta^j) &= \sum_{i=1}^{N} \theta_i^j r_i + \left[1 - \sum_{i=1}^{N} \theta_i^j\right] r_f \\
&= r_f + \sum_{i=1}^{N} \theta_i^j (r_i - r_f) \\
&= r_f + \sum_{i=1}^{N} \theta_i^j (E[r_i] - r_f) + \sum_{i=1}^{N} \theta_i^j (r_i - E[r_i]) \\
&= E[r_p(\theta^j)] + \sum_{i=1}^{N} \theta_i^j \left(\beta_i (f - E[f]) + \zeta_i (g - E[g]) + \eta_i\right) \\
&= E[r_p(\theta^j)] + \left(\sum_{i=1}^{N} \theta_i^j \beta_i\right) \cdot (f - E[f]) + \left(\sum_{i=1}^{N} \theta_i^j \zeta_i\right) \cdot (g - E[g]) + \left(\sum_{i=1}^{N} \theta_i^j \eta_i\right) \\
&\equiv E[r_p(\theta^j)] + \beta_p(\theta^j)(f - E[f]) + \zeta_p(\theta^j)(g - E[g]) + \eta_p(\theta^j)
\end{aligned}
\tag{OA.37}
$$

The variance of the realized return is given by

$$Var(r_p(\theta^j)) = \left[\beta_p(\theta^j)\right]^2 \sigma_f^2 + \left[\zeta_p(\theta^j)\right]^2 \sigma_g^2 + \sum_{i=1}^{K} \left[\theta_i^j\right]^2 \sigma_i^2 \tag{OA.38}$$

while the covariance with the return on human capital r_h is given by

$$Cov(r_p(\theta^j), r_h) = \begin{cases} \beta_p(\theta^j)\psi_f\sigma_f^2 + \zeta_p(\theta^j)\psi_g\sigma_g^2 + \frac{\overline{v}}{K}\psi_\eta \sum_{i=1}^{K} \theta_i^j \sigma_i^2 & \text{if } i = 1, \dots, K \\ \beta_p(\theta^j)\psi_f\sigma_f^2 + \zeta_p(\theta^j)\psi_g\sigma_g^2 & \text{if } i = K+1, \dots, N \end{cases} \tag{OA.39}$$

Partial derivatives of these moments with respect to the portfolio weight θ_i^j for stock i are given by

$$\frac{\partial}{\partial \theta_i^j} E[r_p(\theta^j)] = E[r_i] - r_f \tag{OA.40}$$

$$\frac{\partial}{\partial \theta_i^j} Var(r_p(\theta^j)) = 2\left[\beta_p(\theta^j)\beta_i\sigma_f^2 + \zeta_p(\theta^j)\zeta_i\sigma_g^2 + \theta_i^j\sigma_i^2\right] \tag{OA.41}$$

$$\frac{\partial}{\partial \theta_i^j} Cov(r_p(\theta^j), r_h) = \begin{cases} \beta_i\psi_f\sigma_f^2 + \zeta_i\psi_g\sigma_g^2 + \frac{\overline{v}}{K}\psi_\eta\sigma_i^2 & \text{for } i = 1, \dots, K \\ \beta_i\psi_f\sigma_f^2 + \zeta_i\psi_g\sigma_g^2 & \text{for } i = K+1, \dots, N \end{cases} \tag{OA.42}$$

For large stocks $i = 1, \dots, K$, the first-order condition for θ_i^j in the portfolio choice problem (OA.14) is

$$\begin{aligned} 0 &= \frac{\partial}{\partial \theta_i^j}\left[E[r_p(\theta^j)] - A\mathcal{W}\left(\frac{1-\phi}{2}Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j))\right)\right] \\ &= E[r_i] - r_f - A\mathcal{W}\left[(1-\phi)\left(\beta_p(\theta^j)\beta_i\sigma_f^2 + \zeta_p(\theta^j)\zeta_i\sigma_g^2 + \theta_i^j\sigma_i^2\right) + \phi\left(\beta_i\psi_f\sigma_f^2 + \zeta_i\psi_g\sigma_g^2 + \frac{\overline{v}}{K}\psi_\eta\sigma_i^2\right)\right] \\ &= E[r_i] - r_f - A\mathcal{W}\left[(1-\phi)Cov(r_i, r_p(\theta^j)) + \phi Cov(r_i, r_h)\right] \\ &= E[r_i] - r_f - A\mathcal{W}\left[\beta_i\left((1-\phi)\beta_p(\theta^j) + \phi\psi_f\right)\sigma_f^2 + \zeta_i\left((1-\phi)\zeta_p(\theta^j) + \phi\psi_g\right)\sigma_g^2 + \left((1-\phi)\theta_i^j + \phi\frac{\overline{v}}{K}\psi_\eta\right)\sigma_i^2\right] \end{aligned} \tag{OA.43}$$

For small stocks $i = K+1, \dots, N$ the first-order condition for θ_i^j is

$$\begin{aligned}0 &= \frac{\partial}{\partial \theta_i^j}\left[E[r_p(\theta^j)] - A\mathcal{W}\left(\frac{1-\phi}{2}Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j))\right)\right] \\ &= E[r_i] - r_f - A\mathcal{W}\left[(1-\phi)\left(\beta_p(\theta^j)\beta_i\sigma_f^2 + \zeta_p(\theta^j)\zeta_i\sigma_g^2 + \theta_i^j\sigma_i^2\right) + \phi\left(\beta_i\psi_f\sigma_f^2 + \zeta_i\psi_g\sigma_g^2\right)\right] \\ &= E[r_i] - r_f - A\mathcal{W}\left[(1-\phi)Cov(r_i, r_p(\theta^j)) + \phi Cov(r_i, r_h)\right] \\ &= E[r_i] - r_f - A\mathcal{W}\left[\beta_i\left((1-\phi)\beta_p(\theta^j) + \phi\psi_f\right)\sigma_f^2 + \zeta_i\left((1-\phi)\zeta_p(\theta^j) + \phi\psi_g\right)\sigma_g^2 + (1-\phi)\theta_i^j\sigma_i^2\right]\end{aligned} \tag{OA.44}$$

Therefore, the first-order condition (OA.15) holds for all stocks. To obtain the equilibrium risk-return relationship stated in equation (OA.16), we plug in the equilibrium portfolio shares $\theta_1^{j*} = \dots = \theta_K^{j*} = \overline{v}/K$ and $\theta_{K+1}^{j*} = \dots = \theta_N^{j*} = (1-\overline{v})/(N-K)$ into the above first-order conditions. The normalization (OA.7) implies that $\beta(\theta^{j*}) = 1$, and equation (OA.16) then follows immediately.

The proof of Proposition 1 is straightforward. The expected return on large stocks $i = 1, \dots, K$ does not change as N increases, so we only need to prove that the expected excess return on each small stock $i = K+1, K+2, \dots$ is given by $E[r_i] - r_f = \beta_i E[f]$. Taking limits in the expression from equation (OA.21) directly gives the desired result:

$$E[r_i] - r_f = \beta_i E[f] + A\mathcal{W}(1-\phi)\underbrace{\frac{1-\overline{v}}{N-K}\sigma_i^2}_{\to 0 \text{ as } N\to+\infty} \to \beta_i E[f] \tag{OA.45}$$

To get the result for the market risk premium, combine equation (OA.20) with the fact that $\sigma_\eta^2 \to \sigma_{eta,K}^2$ as $N \to \infty$. In this case, we use σ_η^2 and $\sigma_{\eta,K}^2$ interchangeably for notational simplicity.

OA.6.1 Derivations of model-implied beta gaps

In this subsection, we derive the expression for the beta gaps for the portfolio of all large-cap and the portfolio of all small-cap stocks in our model economy. Equation (OA.6) implies the return on a value-weighted portfolio of all large-cap stocks in the economy is given by:

$$r_L - \mathbb{E}[r_L] \quad = \quad \beta_L(f - \mathbb{E}[f]) + \zeta_L(g - \mathbb{E}[g]) + \eta_L, \tag{OA.46}$$

where $\beta_L \equiv \frac{1}{K}\sum_{i=1}^{K}\beta_i$, $\zeta_L \equiv \frac{1}{K}\sum_{i=1}^{K}\zeta_i \equiv \frac{\zeta_m}{\overline{\nu}}$, and $\eta_L \equiv \frac{1}{K}\sum_{i=1}^{K}\eta_i$. Similarly, returns on a value-weighted portfolio of all small-cap stocks in the economy is given by:

$$r_S - \mathbb{E}[r_S] \;=\; \beta_S (f - \mathbb{E}[f]) + \eta_S, \tag{OA.47}$$

where $\beta_S \equiv \frac{1}{N-K}\sum_{i=K+1}^{N}\beta_i$ and $\eta_S \equiv \frac{1}{N-K}\sum_{i=K+1}^{N}\eta_i$. The assumption that $\frac{1}{N-K}\sum_{i=K+1}^{N}\zeta_i = 0$ implies that the small-cap portfolio return does not load on g. The large-cap portfolio beta with respect to the value-weighted market return given in equation (OA.9) is given by:

$$\begin{aligned}
\beta_L^m &= \frac{\mathbb{COV}(r_L - \mathbb{E}[r_L], r_m - \mathbb{E}[r_m])}{\mathbb{VAR}(r_m - \mathbb{E}[r_m])} \\
&= \frac{\mathbb{COV}(\beta_L f + \zeta_L g + \eta_L, f + \zeta_m g + \eta)}{\mathbb{VAR}(f + \zeta_m g + \eta)} \\
&= \frac{\mathbb{COV}\left(\beta_L f + \frac{\zeta_m}{\overline{\nu}} g + \frac{1}{K}\sum_{i=1}^{K}\eta_i, f + \zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i + \frac{1-\overline{\nu}}{N-K}\sum_{i=K+1}^{N}\eta_i\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]} \\
&= \frac{\beta_L\sigma_f^2 + \frac{\zeta_m^2}{\overline{\nu}}\sigma_g^2 + \frac{\overline{\nu}}{K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}.
\end{aligned} \tag{OA.48}$$

The small-cap portfolio beta with respect to the value-weighted market return in equation (OA.9) is given by:

$$\begin{aligned}
\beta_S^m &= \frac{\mathbb{COV}(r_S - \mathbb{E}[r_S], r_m - \mathbb{E}[r_m])}{\mathbb{VAR}(r_m - \mathbb{E}[r_m])} \\
&= \frac{\mathbb{COV}(\beta_S f + \eta_S, f + \zeta_m g + \eta)}{\mathbb{VAR}(f + \zeta_m g + \eta)} \\
&= \frac{\mathbb{COV}\left(\beta_S f + \frac{1}{N-K}\sum_{i=K+1}^{N}\eta_i, f + \zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i + \frac{1-\overline{\nu}}{N-K}\sum_{i=K+1}^{N}\eta_i\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]} \\
&= \frac{\beta_S\sigma_f^2 + \frac{1-\overline{\nu}}{N-K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}.
\end{aligned} \tag{OA.49}$$

It is straighforward to show that the betas of the large- and small-cap portfolio with respect to f are just β_L and β_S, respectively.

We define the beta gap as the difference between the VMF beta and the DMF beta (i.e., for an asset i the beta gap is $\beta_i^{VMF} - \beta_i^{DMF}$ or, in the context of the model, the beta gap is $\beta_i^m - \beta_i^f$). Therefore, the beta gap for the large-cap portfolio is given by:

$$\begin{aligned}
\beta_L^m - \beta_L^f &= \frac{\beta_L \sigma_f^2 + \zeta_m^2 \sigma_g^2 + \frac{\overline{\nu}}{K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]} - \beta_L \\
&= \frac{(1-\beta_L)\,\zeta_m^2\sigma_g^2 + \left(\frac{\overline{\nu}}{K} - \beta_L\right)\overline{\nu}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}\frac{1}{\overline{\nu}} \\
&= \frac{\zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2 - \beta_L\left(\zeta_m^2\sigma_g^2 - \overline{\nu}\sigma_i^2\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}\frac{1}{\overline{\nu}}. \qquad \text{(OA.50)}
\end{aligned}$$

In the limit as $N \to \infty$, the large-cap portfolio beta gap is given by:

$$\lim_{N\to\infty} \beta_L^m - \beta_L^f = \frac{1}{\overline{\nu}}\left[\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2} - \beta_L \frac{\zeta_m^2\sigma_g^2 - \overline{\nu}\sigma_i^2}{\sigma_f^2 + \sigma_\epsilon^2}\right] \qquad \text{(OA.51)}$$

where σ_ϵ^2 is the variance of the IFF in the model economy (i.e., $\epsilon \equiv \zeta_m g + r_\eta$ and $\sigma_\epsilon^2 \equiv \zeta_m^2\sigma_g^2 + \sigma_\eta^2$). So, the sign of the large-cap portfolio beta gap is ambiguous in the context of the model, and depends on a number of parameters (e.g., β_L, ζ_m, σ_g, σ_i, $\overline{\nu}$, etc.). The beta gap for the small-cap portfolio is given by:

$$\begin{aligned}
\beta_S^m - \beta_S^f &= \frac{\beta_S \sigma_f^2 + \frac{1-\overline{\nu}}{N-K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]} - \beta_S \\
&= \frac{\left(1 - \beta_S\left(1-\overline{\nu}\right)\right)\frac{1-\overline{\nu}}{N-K}\sigma_i^2 - \beta_S\left(\sigma_i^2\frac{\overline{\nu}^2}{K} + \zeta_m^2\sigma_g^2\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}. \qquad \text{(OA.52)}
\end{aligned}$$

In the limit as $N \to +\infty$, the small-cap portfolio beta gap is given by:

$$\begin{aligned} \lim_{N\to\infty} \beta_S^m - \beta_S^f &= -\left(\frac{\zeta_m^2\sigma_g^2 + \sigma_i^2\frac{\overline{\nu}^2}{K}}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\frac{\overline{\nu}^2}{K}}\right)\beta_S \\ &\equiv -\left(\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)\beta_S, \end{aligned} \tag{OA.53}$$

So, the small cap portfolio beta gap is unambiguously less than zero as long as there is a large number of small-cap stocks and as long as $\beta_S > 0$ (i.e., the empirically-relevant case). Furthermore, the beta gap is proportional to the true small-cap portfolio beta where the constant of proportionality is the ratio of the IFF variance to that of the value-weighted market portfolio variance.

OA.6.2 Derivations of model-implied small-cap portfolio and IFF CAPM alphas

In this subsection, we derive the CAPM alphas of the small-cap portfolios and the IFF (i.e., $\epsilon \equiv \zeta_m g + r_\eta$) with respect to the CAPM when either the value-weighted market factor, $r_m - r_f$, is used as the market factor in the CAPM, or f is used. We assume we are in the case where $N \to +\infty$, and labor income does not load on g or the granular residuals (i.e., $\phi_g = \phi_\eta = \psi_g = \psi_\eta = 0$). We proceed by first stating the expected excess returns on the small-cap portfolio and the IFF. Next, we state expected returns on each asset under the CAPM when either $r_m - r_f$ or f is used as the market factor. The difference between the true expected excess return and the model-based expected excess returns are the implied alphas under each factor model.

Equation (OA.27) implies the expected excess return on the small-cap portfolio is:

$$\mathbb{E}[r_S] - r_f = \beta_S \mathbb{E}[f]. \tag{OA.54}$$

The assumption that $\frac{1}{N-K}\sum_{i=K+1}^{N}\zeta_i = 0$ implies that the small-cap portfolio expected excess return does not load on g. The expected excess return on the IFF is:

$$\mathbb{E}[\epsilon] = \zeta_m \mathbb{E}[g] + \mathbb{E}[r_\eta] \tag{OA.55}$$

When f is used as the market factor, the measured small-cap portfolio exposure to the market is given by β_S so that the implied alpha is:

$$\begin{aligned}
\alpha_S^f &= \mathbb{E}[r_S] - r_f - \beta_S(\mathbb{E}[f]) \\
&= \beta_S\mathbb{E}[f] - \beta_S\mathbb{E}[f] \\
&= 0
\end{aligned} \tag{OA.56}$$

In other words, f correctly prices the small-cap portfolio, as expected. When r_m is used as the market factor, the measured small-cap portfolio exposure to the market is given in equation (OA.49) so that the implied alpha is:

$$\begin{aligned}
\alpha_S^m &= \mathbb{E}[r_S] - r_f - \beta_S^m(\mathbb{E}[r_m] - r_f) \\
&= \beta_S\mathbb{E}[f] - \beta_S\left(\frac{\sigma_f^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\frac{\overline{\nu}^2}{K}}\right)(\mathbb{E}[f] + \zeta_m\mathbb{E}[g] + \mathbb{E}[r_\eta])
\end{aligned}$$

where the second line makes use of the expression for the expected excess return on r_m in equation (OA.20). Making use of the expressions for $\mathbb{E}[f]$, $\mathbb{E}[g]$, and $\mathbb{E}[r_\eta]$ in equations (OA.24)-(OA.26) and simplifying:

$$\begin{aligned}
\alpha_S^m &= \beta_S\mathbb{E}[f] - \beta_S\left(\frac{\frac{\mathbb{E}[f]}{A\mathcal{W}\left(1-\phi+\phi\psi_f\right)}}{\frac{\mathbb{E}[f]}{A\mathcal{W}\left(1-\phi+\phi\psi_f\right)} + \zeta_m\frac{\mathbb{E}[g]}{A\mathcal{W}(1-\phi)} + \frac{\mathbb{E}[r_\eta]}{A\mathcal{W}(1-\phi)}}\right)(\mathbb{E}[f] + \zeta_m\mathbb{E}[g] + \mathbb{E}[r_\eta]) \\
&= \beta_S\mathbb{E}[f]\left[1 - \frac{\mathbb{E}[f] + \mathbb{E}[\epsilon]}{\mathbb{E}[f] + \mathbb{E}[\epsilon]\frac{\left(1-\phi+\phi\psi_f\right)}{(1-\phi)}}\right] \\
&= \beta_S\mathbb{E}[f]\left[1 - \frac{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi)\,\zeta_m^2\sigma_g^2 + (1-\phi)\,\sigma_\eta^2}{(1-\phi+\phi\psi_f)\,\sigma_f^2 + \left((1-\phi)\,\zeta_m^2\sigma_g^2 + (1-\phi)\,\sigma_\eta^2\right)\frac{\left(1-\phi+\phi\psi_f\right)}{(1-\phi)}}\right]
\end{aligned}$$

Converting back to variances using equations (OA.24)-(OA.26) and noting the definition $\sigma_\epsilon^2 \equiv \zeta_m^2\sigma_g^2 + \sigma_\eta^2$, and simplifying:

$$\begin{aligned}
\alpha_S^m &= \beta_S \mathbb{E}[f]\left[1 - \frac{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi)\,\zeta_m^2\sigma_g^2 + (1-\phi)\,\sigma_\eta^2}{(1-\phi+\phi\psi_f)\,\sigma_f^2 + \left((1-\phi)\,\zeta_m^2\sigma_g^2 + (1-\phi)\,\sigma_\eta^2\right)\frac{(1-\phi+\phi\psi_f)}{(1-\phi)}}\right] \\
&= \beta_S \mathbb{E}[f]\left[1 - \frac{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi)\,\sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi)\,\sigma_\epsilon^2\frac{(1-\phi+\phi\psi_f)}{(1-\phi)}}\right] \\
&= \beta_S \mathbb{E}[f]\left[1 - \frac{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi)\,\sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi+\phi\psi_f)\,\sigma_\epsilon^2}\right] \\
&= \beta_S \mathbb{E}[f]\left[\frac{\phi\psi_f\sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\,\sigma_f^2 + (1-\phi+\phi\psi_f)\,\sigma_\epsilon^2}\right] > 0. \qquad \text{(OA.57)}
\end{aligned}$$

Assuming $\beta_S > 0$ and $\psi_f > 0$ (which are the empirically-relevant cases). As $\phi \to 1$, we have:

$$\lim_{\phi\to 1}\alpha_S^m = \frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\beta_S\mathbb{E}[f] > 0. \qquad \text{(OA.58)}$$

So, as wealth in the economy becomes more concentrated in sources other than financial wealth, the measured α_S^m converges to the (negative) beta gap in equation (OA.53) (multiplied by the expected return on f).

When f is used as the market factor, the measured IFF exposure to the market is given by:

$$\begin{aligned}
\beta_\epsilon^f &= \frac{\mathbb{COV}(\epsilon, f)}{\mathbb{VAR}(f)} \\
&= \frac{\mathbb{COV}(\zeta_m g + r_\eta, f)}{\mathbb{VAR}(f)} \\
&= 0. \qquad \text{(OA.59)}
\end{aligned}$$

When r_m is used as the market factor, the measured IFF exposure to the market is given by:

$$\begin{aligned}
\beta_\epsilon^m &= \frac{\mathbb{COV}\left(\epsilon, r_m\right)}{\mathbb{VAR}\left(r_m\right)} \\
&= \frac{\mathbb{COV}\left(\zeta_m g + r_\eta, r_m\right)}{\mathbb{VAR}\left(r_m\right)} \\
&= \frac{\mathbb{COV}\left(\zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i, f + \zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2} \\
&= \frac{\zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2} \\
&= \frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2} \qquad \text{(OA.60)}
\end{aligned}$$

Given the expression for the expected return on ϵ in equation (OA.55), the expression for β_ϵ^f in equation (OA.59), and the expression for the expected return on f in equation (OA.24), the alpha implied when f is used as the market factor is:

$$\begin{aligned}
\alpha_\epsilon^f &= \mathbb{E}\left[\epsilon\right] - \beta_\epsilon^f \mathbb{E}\left[f\right] \\
&= \zeta_m \mathbb{E}\left[g\right] + \mathbb{E}\left[r_\eta\right] \\
&= A\mathcal{W}\left(1-\phi\right)\zeta_m^2\sigma_g^2 + A\mathcal{W}\left(1-\phi\right)\sigma_\eta^2 \qquad \text{(OA.61)}
\end{aligned}$$

In the limit as $\phi \to 1$, this yields:

$$\lim_{\phi\to 1}\alpha_\epsilon^f = 0. \qquad \text{(OA.62)}$$

Given the expression for the expected return on ϵ in equation (OA.55), the expression for β_ϵ^m in equation (OA.60), and the expression for the expected excess return on r_m

in equation (OA.20), the alpha implied when $r_m - r_f$ is used as the market factor is:

$$
\begin{aligned}
\alpha_\epsilon^m &= \mathbb{E}[\epsilon] - \beta_\epsilon^m (\mathbb{E}[r_m] - r_f) \\
&= \zeta_m \mathbb{E}[g] + \mathbb{E}[r_\eta] - \left(\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)(\mathbb{E}[f] + \zeta_m \mathbb{E}[g] + \mathbb{E}[r_\eta]) \\
&= -\left(\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)\mathbb{E}[f] + \zeta_m \mathbb{E}[g]\left(1 - \frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right) + \mathbb{E}[r_\eta]\left(1 - \frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right) \\
&= \left(\frac{\sigma_f^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)\mathbb{E}[f] + \zeta_m \mathbb{E}[g]\left(\frac{\sigma_f^2}{\sigma_f^2 + \sigma_\epsilon^2}\right) + \mathbb{E}[r_\eta]\left(\frac{\sigma_f^2}{\sigma_f^2 + \sigma_\epsilon^2}\right) - \mathbb{E}[f] \\
&= \left(\frac{\sigma_f^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)(\mathbb{E}[f] + \zeta_m \mathbb{E}[g] + \mathbb{E}[r_\eta]) - \mathbb{E}[f]
\end{aligned}
$$

Making use of the expressions for $\mathbb{E}[f]$, $\mathbb{E}[g]$, and $\mathbb{E}[r_\eta]$ in equations (OA.24)-(OA.26) and simplifying:

$$
\begin{aligned}
\alpha_\epsilon^m &= A\mathcal{W}\left(\frac{\sigma_f^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)\left((1 - \phi + \phi\psi_f)\,\sigma_f^2 + (1-\phi)\,\zeta_m^2 \sigma_g^2 + (1-\phi)\,\sigma_\eta^2\right) \\
&\quad - A\mathcal{W}(1 - \phi + \phi\psi_f)\,\sigma_f^2 < 0. \qquad \text{(OA.63)}
\end{aligned}
$$

In the limit as $\phi \to 1$, this yields:

$$
\lim_{\phi \to 1} \alpha_\epsilon^m = -A\mathcal{W}\left(\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}\right)\psi_f \sigma_f^2 < 0. \qquad \text{(OA.64)}
$$

OA.7 Evidence on the share of non-public-equity market wealth

According to the intuition developed in our theoretical model from the previous section (OA.5), we expect the granular residual to have a relatively small price of risk when the labor share of wealth in the economy (ϕ) is close to one. In the limit, the granular residual's price of risk should go to zero. More generally, this argument applies in the real economy when sources of investor wealth other than public equity dominate the wealth portfolio (assuming the granular residual associated with large cap stocks is uncorrelated with these other sources of wealth). In this section, we highlight some evidence from the US economy that suggests public equity comprises a relatively small fraction of the overall wealth portfolio.

First, public firms are an important part, but not the only part, of the U.S. economy. For example, Davis et al. (2006) estimate using Census data that public firms comprised only about 30% of total employment in the U.S. over their sample. For many investors, news about expected labor and/or entrepreneurial income (and potentially uninsurable risk associated with these factors) may be more important than dividends from investments in public firms. In the US, where stock market participation is particularly high relative to most other developed countries, SCF data indicate that the average household has about one fifth of total financial wealth in the stock market. While equity shares are higher among wealthier households, the increase is purely driven by equity in private businesses in which the survey respondent has a role in managing (see, e.g., Gomez, 2017). Piketty et al. (2017) report that approximately 70% of total national income is labor income, and Lustig et al. (2013) estimate that stock market wealth and total non-human wealth only comprise 1% and 8% of total wealth, respectively. Along these lines, van Binsbergen and Opp (2019) report that that firm net payout has never exceeded 20% of U.S. personal consumption expenditures.

Next, as we show in Figure 2, the industry distribution of "megacap" firms (those in the top 2.5% of the market value distribution, using NYSE market cap breakpoints) is not even representative of the industry distribution of the rest of the public equity market and exhibits much more pronounced and volatile time-variation. It also unlikely that idiosyncratic shocks to these megacap firms are representative of other sources of wealth in the economy, and therefore are not priced by the marginal utility of the representative investor.

Finally, in a world where US public equity is only a small portion of many sources of macroeconomic risk – such as those associated with other traded financial assets, private equity, real estate, uninsurable idiosyncratic risk, and human capital returns – that can enter the stochastic discount factor, there are many other examples (other than our theoretical model from section OA.5) in which changes in the value of public equity induced by idiosyncratic shocks to large public firms do not command high risk premia. See, e.g., Mayers (1973), Fama and French (1996), Campbell (1996), Heaton and Lucas (2000), Malloy et al. (2009), Campbell et al. (2016), and Schmidt

(2016), among many others.

So, given the above evidence that the majority of investor wealth in the US economy is not comprised of public equities (let alone megacaps), it is likely that any granular residual generated by these megacaps has a low price of risk that differs substantially from that of the dominant market factor as in our theoretical model.

OA.8 Additional motivating evidence on granularity

To support our discussion and empirical tests in the main text, here we provide some empirical examples of information releases which were associated with large single-day changes in the CRSP value-weighted index.

Given the significant market concentration among largest stocks, drastic changes in a large company's stock price can singlehandedly pull down entire market returns by nontrivial amounts. Figure OA.14 plots cumulative abnormal stock returns for four megacap companies around the events of interest (blue lines) along with their counterfactual impacts on the value-weighted market index return (red lines). The cumulative abnormal returns are constructed by aggregating residuals obtained from a single market factor model. To do so, we estimate a stock's exposure on the market factor (i.e., beta) from an ordinary least squares regression of excess individual stock returns on the excess CRSP value-weighted index returns, using daily observations from 5 months to 1 month prior to the event. To demonstrate a megacap company's nontrivial contribution to the value-weighted market index return, we estimate its counterfactual impacts as differences in returns of the value-weighted market index with the stock of interest and without the stock. This latter calculation captures the direct impact of each company-specific market return on the index relative to a counterfactual in which it had had identical returns as the value-weighted average of other firms over the same period. The red vertical lines indicate the event date of interest.

Starting from the top-left quadrant and moving clockwise, we have the following events. The first three examples are associated with earnings announcements of Microsoft (positive news), GE (negative news), and Apple (positive earnings, but below expectations, triggering a selloff).[68] Prior to these announcements, their stocks accounted for sizable shares in the CRSP value-weighted index; Microsoft, GE, and Apple had weights of 2.8%, 2.5%, and 3.2% of the index. Each of these announcements was associated with a change in market value of around 15-20%; thus, the difference between the index with and without each stock in each month is 0.4%, -0.3%, and -0.4%, respectively. The last example comes from IBM, against which a Federal court, in favor of Telex, handed down an antitrust ruling.[69] Given the IBM's weight of 5.3% prior to the announcement, its 20% abnormal returns associated with the announcement triggered an additional 1% decline in the index in September 1973.

As discussed above, the distinction between aggregate and idiosyncratic shocks becomes quite blurry when one considers a large shock experienced by one of the largest firms in the economy. Large changes in large firms' valuations (e.g., earnings announcements) may act as bellwethers for other firms. In our examples above, GE's disappointing earnings announcement, which was driven by losses in its financial services division in the spring of 2008, may indeed have provided a negative signal about the fortunes of other companies in the financial services sector prior to the burst of the financial crisis. Likewise, Microsoft's surprisingly positive earnings announcement in 1999 was likely to be perceived as conveying positive news for other tech stocks as well, and perhaps the Telex ruling conveyed some useful information about future antitrust policy which might have aggregate implications.[70] Nonetheless, it seems reasonable to assume that investors likely viewed fluctuations in the market index driven by these shocks in a different way from more systematic drivers of returns.

OA.9 GARCH models: Additional details

We estimate parameters in the GARCH volatility forecasting models by numerical maximum likelihood estimation (MLE). After estimating parameters, we calculate

[68]For more details, see Fisher, "Microsoft's Profit Up 75% in Quarter." New York Times, January 20, 1999; Associated Press, "Shares Sink as G.E. Disappoints." New York Times, April 12, 2008; and Wingfield, "Heady Returns", New York Times, January 24, 2014.

[69]See, e.g., Vartan, "Stocks Rise Again in Heavy Trading." New York Times, September 26, 1973 and Vartan "Prices of Stocks Climb Despite Plunge by I.B.M." New York Times, September 18, 1973.

[70]Note that, to the extent that the news conveyed by these announcements had a large impact on valuations of other stocks, the differences between the index return with and without the stocks in question would be pushed towards zero.

their asymptotic standard errors by approximating the first and second derivatives of the log-likelihood function at maximum likelihood estimates. Reported standard errors are computed using the "outer product" method. Details for numerical MLE estimation and calculation of asymptotic standard errors can be found in Hamilton (Time Series Analysis, 1994, pp. 133-148)).

The non-trivial difference between the DMF- and the VMF-based volatility forecasts supports our conjecture that the IFF adds noise to the lagged forecasting errors used to generate conditional variance forecasts when using the VMF. Additional evidence for this conjecture can be found in Byun (2016), who shows that cross-sectional dispersion in the returns of different stocks, a measure of aggregate idiosyncratic risk, does not help forecast volatility of the S&P 500 index when using the cross-sectional dispersion as an additional explanatory variable in GARCH forecasting models. See Byun (2016) for a detailed explanation and also for an alternative approach through which cross-sectional dispersion measures can improve volatility forecasts indirectly.

OA.9.1 Alternative GARCH models

This section provides more details related to the GJR-GARCH and E-GARCH models. In a study of the relationship between expected excess stock returns and volatility, Glosten et al. (1993) introduced the GJR-GARCH model, which allows for a larger feedback from prior squared negative returns relative to positive returns. Denoting by I_t an indicator variable for a positive forecasting error in period t, the conditional variance is given by

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 u_t^2 + \alpha_2 \sigma_t^2 + \alpha_3 I_t u_t^2, \qquad \text{(OA.65)}$$

where $I_t = 1$ for $u_t > 0$ and 0 otherwise.

Next, Nelson (1991) proposed the EGARCH model, which specifies the conditional variance in logarithmic form as

$$\ln(\sigma_{t+1}^2) = \alpha_0 + \alpha_1 \left(|\widetilde{u}_t| - E\,|\widetilde{u}_t|\right) + \alpha_2 \ln(\sigma_t^2) + \alpha_3 \widetilde{u}_t, \qquad \text{(OA.66)}$$

where $\widetilde{u}_t = u_t/\sigma_t$ is a standardized forecast error in period t and $E\,|\widetilde{u}_t| = 2/\pi$ for a normally distributed $\widetilde{u}_t$. Note that the above asymmetric GARCH models capture the leverage effect through α_3: with $\alpha_3 < 0$, the conditional variance (σ_{t+1}^2) is higher for a negative than for a positive forecasting error (u_t).[71]

OA.10 Constructing and controlling for orthogonal factors, g_{t+1}

Even if we estimate β_p consistently, our second-stage estimates of the market risk premium, λ_1, in equation (26) will be biased unless we control for exposure to other common, priced factors g_{t+1}, or if it happens to be the case that β_p and ζ_p are uncorrelated in the cross-section, or if $\lambda_2 = 0$. To address this issue, we use results from the PCA analysis to hedge test asset exposure to these g_{t+1} factors.

We assume that the PCs adequately span g_{t+1}. This is reasonable given our test assets because the PCs are constructed to span the space of test asset returns by construction, and therefore they will also span the unobservable g_{t+1} in our sample. By assumption and to consistently estimate λ_1, we need g_{t+1} to be orthogonal to f_{t+1} in our tests (i.e., cases where we consider either VMF or DMF as proxies for f_{t+1}). Accordingly, in tests that use the VMF, we orthogonalize PCs to this factor in our full sample (1963-2021) and take these as our empirical proxies for g_{t+1}. In tests that use the DMF, we need not orthogonalize the PCs because they are already orthogonal to the DMF by construction.

Using our PCs as proxies for g_{t+1} has a few advantages. First, as noted above, these PCs span the unobservable g_{t+1} in our sample. Second, controlling for g_{t+1} allows us to consistently estimate the market risk premium, λ_1, for reasons similar to those articulated in Giglio and Xiu (2021). Third, PCs offer a statistical characterization of factors that explain average returns in the cross-section, and this obviates the need for us to take a stand on which extant factor model should be used for such a purpose.

Given the empirical proxies for the factors, g_{t+1}, we would like to construct test asset returns that are hedged of exposure to these factors so that we can use these hedged returns to estimate λ_1 in standard cross-sectional regressions and to avoid the need to estimate risk premia associated with the g_{t+1} (which effectively imposes our model's null hypothesis that risk premia on g_{t+1} are equal to their time series

[71]The recognition of the leverage effect goes back to Black (1976), noting a negative correlation between current returns and future volatility, and is further investigated by Christie (1982). According to the leverage effect, a reduction in the equity value would raise the debt-to-equity ratio, hence raising the riskiness of the firm's equity as manifested by an increase in future volatility.

averages), which reduces noise in the estimated λ_1 (under the null). We do this in a manner such that the hedged returns represent returns on (ex post) tradable portfolios given g_{1+t}. Specifically, for each portfolio we estimate equation (24) using our full data (1963-2021), which yields estimates of β_p and ζ_p. We then construct the month-$t+1$ hedged returns, $r_{p,t+1}^{hedged}$ according to equation (25). We construct separate $r_{p,t+1}^{hedged}$ for both the VMF and the DMF, and use these in the respective tests of each factor.

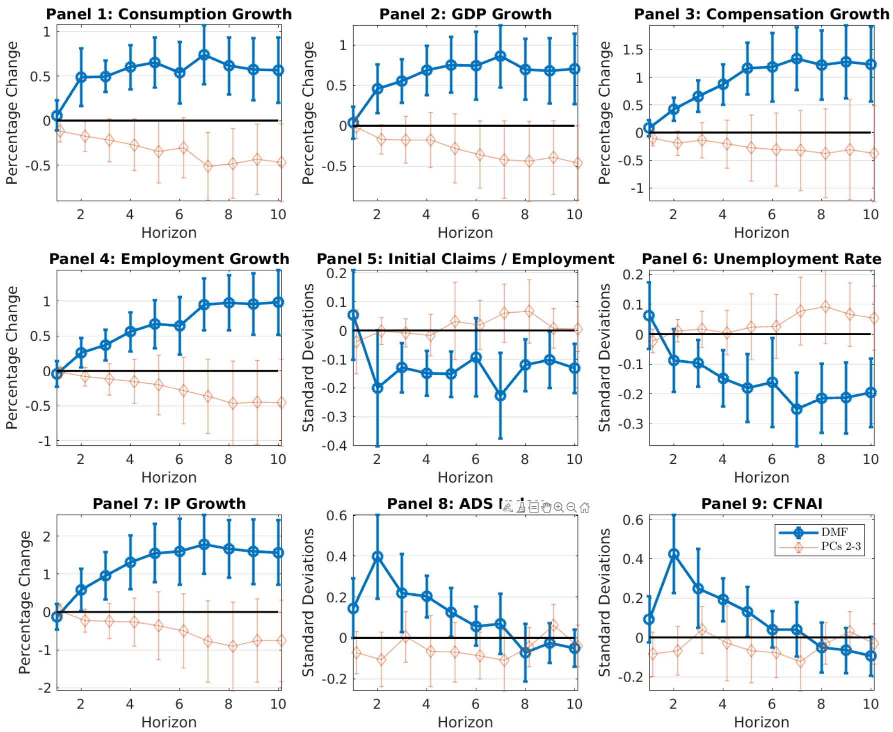
In our main analysis, we choose to use five PCs to construct g_{t+1} for parsimony, although our results are insensitive to this choice.[72] We also provide results without hedging g_{t+1} (i.e., when f_{t+1} is the only factor in equation (24)), and find our main conclusions about the beta gap are insensitive to this choice.

OA.11 Bootstrap procedure for estimating standard errors and p-values

We use the following bootstrap procedure to estimate standard errors and p-values whenever indicated in our tables. We begin by assuming that our PCs, the DMF, and the IFF are fixed within a given sample of data (i.e., we do not sample from portfolio returns then recompute the PCA, the DMF, and the IFF within each simulated sample). We then sample the same number of periods of data as in our sample of interest (with replacement) 10,000 times, recomputing the statistic of interest each time. Reported standard errors are simply the standard deviation of the simulated statistics. P-values are the number of simulated statistics that satisfy whatever testing criterion in which we are interested (as a fraction of the total number of simulations). This bootstrap procedure effectively adjusts for the generated regressor problem when computing standard errors and p-values in our cross-sectional two-stage regressions.

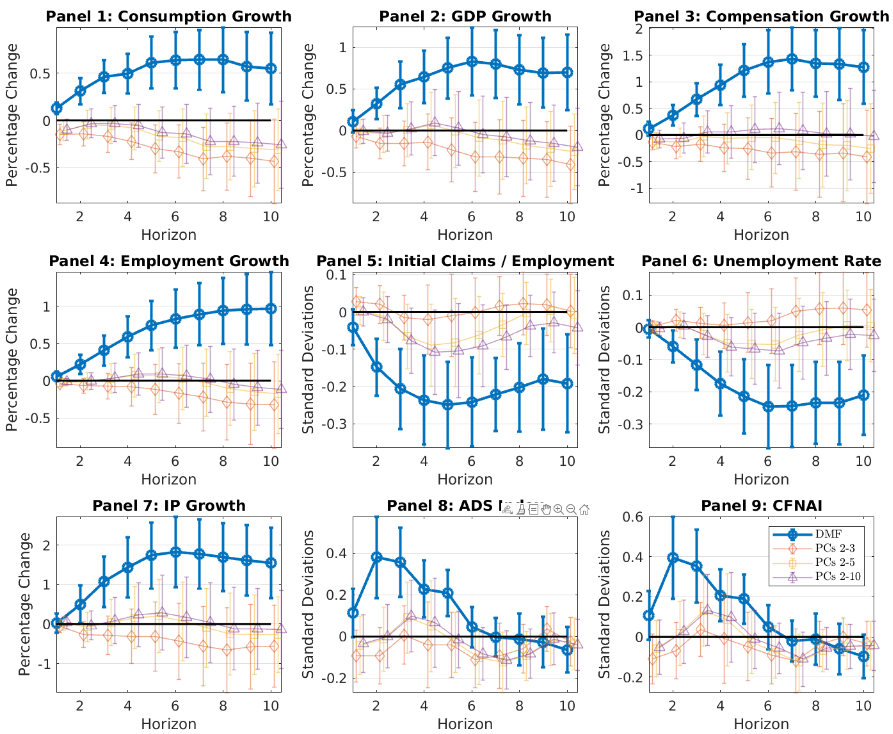
[72]This is true as long as we use approximately three or more PCs so that we can adequately span the return space not spanned by market return proxies.

Figure OA.1: Impulse response functions to PCs (including COVID pandemic)



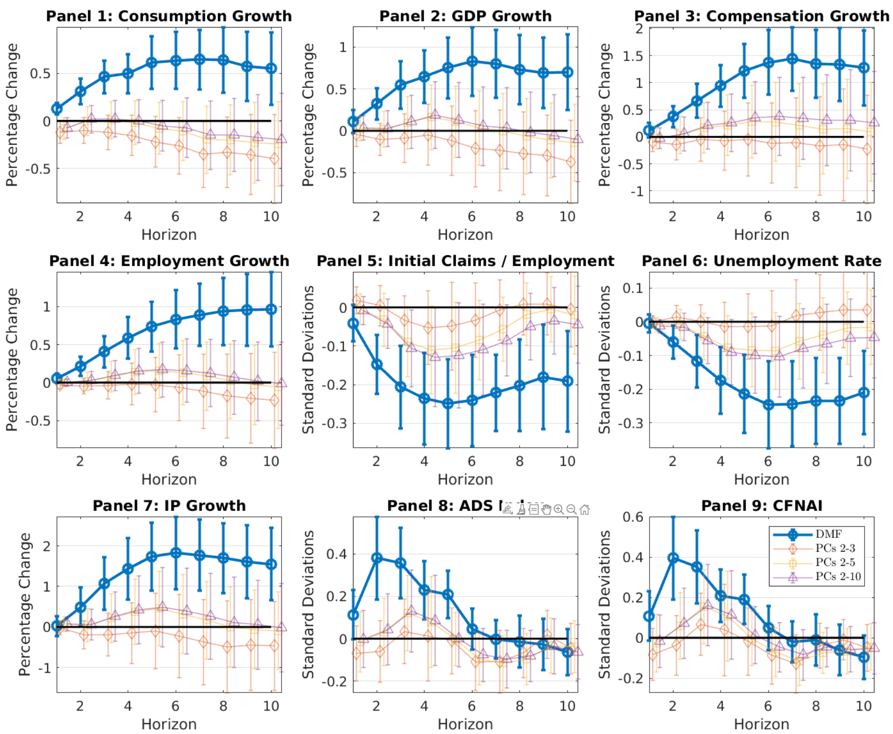
Note. This figure displays impulse response functions of selected macroeconomic aggregates to scaled versions of the DMF and to a linear projection of VMF onto PCs two and three where the weights are obtained from a full-sample multivariate linear regression of the CRSP value-weighted market excess returns on PCs one through three as described by equation (OA.5). Monthly returns are compounded to the quarterly frequency to match that of the macroeconomic aggregates. The point estimates represent the impact of a standard deviation increase in the predictor variable on each macroeconomic aggregate. The initial point estimates (i.e., when the horizon is equal to one) correspond to the contemporaneous response of each macroeconomic aggregate to each predictor variable. The error bars indicate 95% confidence intervals computed using Newey-West standard errors with 10 lags. The descriptions of the macroeconomic variables can be found in the main text. All data are from the 1963-2021 sample. Additional details on the estimation procedure can be found in Online Appendix OA.4.

Figure OA.2: Impulse response functions to PCs (additional PCs)



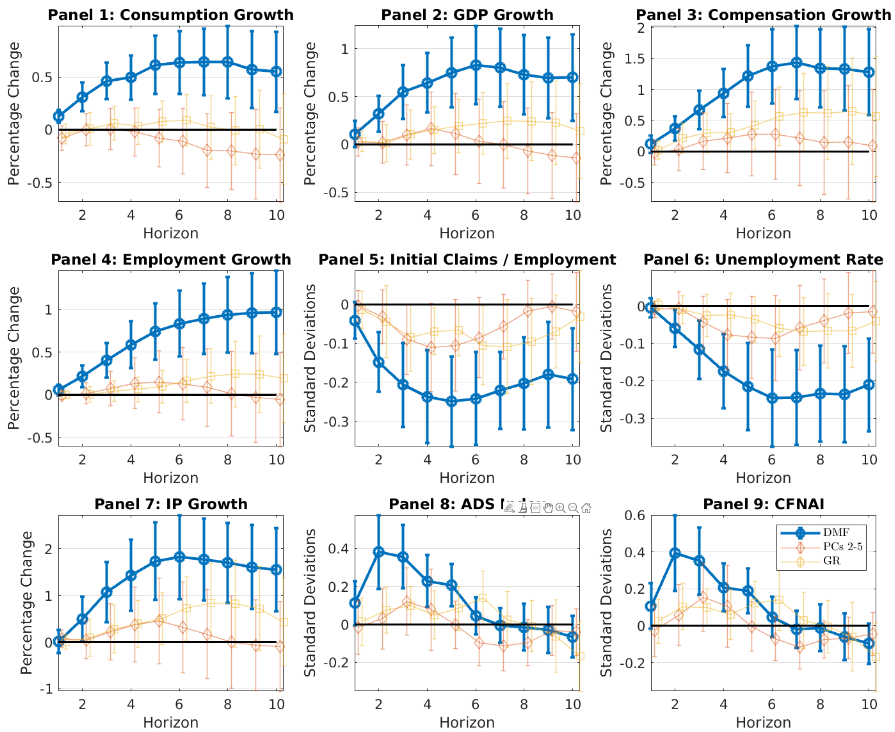
Note. This figure displays impulse response functions of selected macroeconomic aggregates to scaled versions of the DMF and to a linear projection of VMF onto PCs two through three, two through five, and two through ten where the weights are obtained from a full-sample multivariate linear regression of the CRSP value-weighted market excess returns on PCs as described by equation (OA.5). Monthly returns are compounded to the quarterly frequency to match that of the macroeconomic aggregates. The point estimates represent the impact of a standard deviation increase in the predictor variable on each macroeconomic aggregate. The initial point estimates (i.e., when the horizon is equal to one) correspond to the contemporaneous response of each macroeconomic aggregate to each predictor variable. The error bars indicate 95% confidence intervals computed using Newey-West standard errors with 10 lags. The descriptions of the macroeconomic variables can be found in the main text. All data are from the 1963-2019 sample. Additional details on the estimation procedure can be found in Online Appendix OA.4.

Figure OA.3: Impulse response functions to PCs (additional PCs, rolling estimation)



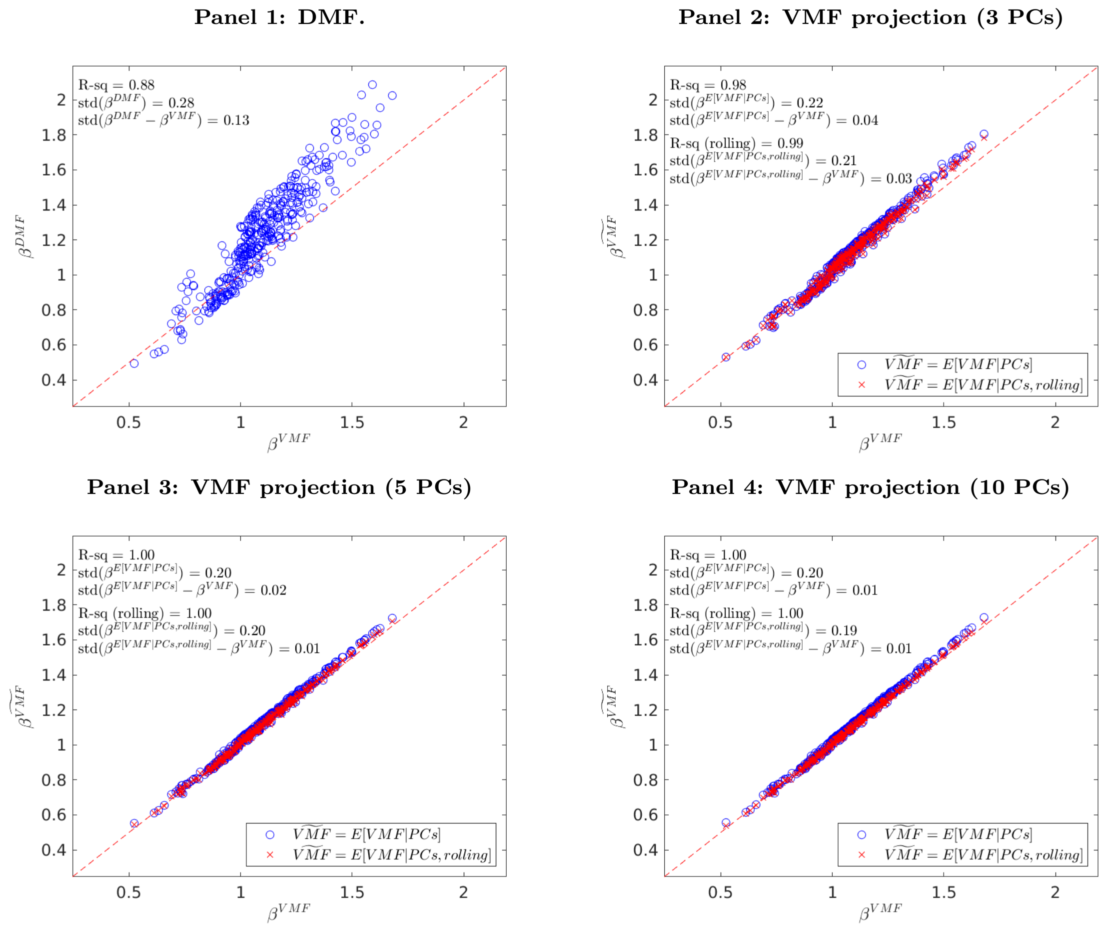
Note. This figure displays impulse response functions of selected macroeconomic aggregates to scaled versions of the DMF and to a linear projection of VMF onto PCs two through three, two through five, and two through ten where the weights are obtained from a rolling multivariate linear regression of the CRSP value-weighted market excess returns on PCs as described by equation (OA.5) using a centered 60-month rolling window. Monthly returns are compounded to the quarterly frequency to match that of the macroeconomic aggregates. The point estimates represent the impact of a standard deviation increase in the predictor variable on each macroeconomic aggregate. The initial point estimates (i.e., when the horizon is equal to one) correspond to the contemporaneous response of each macroeconomic aggregate to each predictor variable. The error bars indicate 95% confidence intervals computed using Newey-West standard errors with 10 lags. The descriptions of the macroeconomic variables can be found in the main text. All data are from the 1963-2019 sample. Additional details on the estimation procedure can be found in Online Appendix OA.4.

Figure OA.4: Impulse response functions to PCs (including GR, rolling estimation)



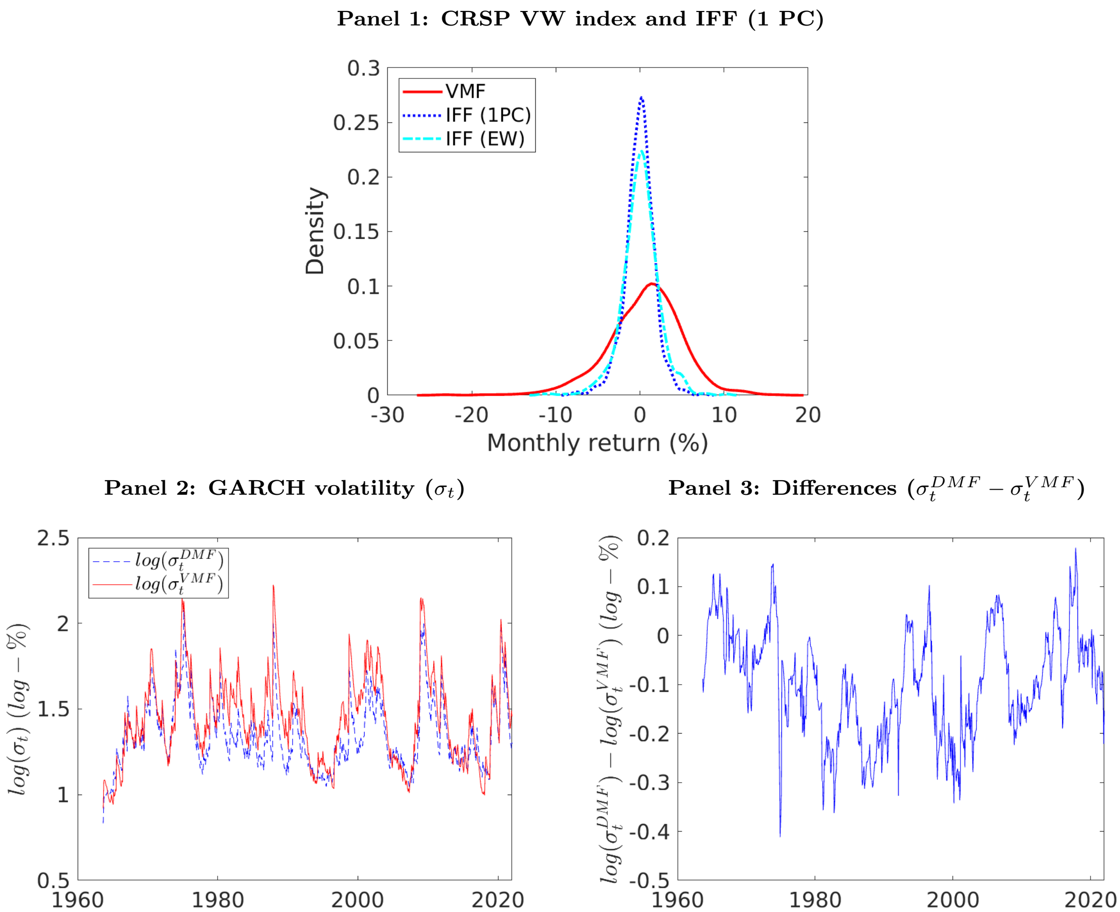
Note. This figure displays impulse response functions of selected macroeconomic aggregates to scaled versions of the DMF, to a linear projection of VMF onto PCs two through five where the weights are obtained from a rolling multivariate linear regression of the CRSP value-weighted market excess returns on PCs as described by equation (OA.5), and to the associated estimated granular residual (i.e., the residual from the regression) using a centered 60-month rolling window. Monthly returns are compounded to the quarterly frequency to match that of the macroeconomic aggregates. The point estimates represent the impact of a standard deviation increase in the predictor variable on each macroeconomic aggregate. The initial point estimates (i.e., when the horizon is equal to one) correspond to the contemporaneous response of each macroeconomic aggregate to each predictor variable. The error bars indicate 95% confidence intervals computed using Newey-West standard errors with 10 lags. The descriptions of the macroeconomic variables can be found in the main text. All data are from the 1963-2019 sample. Additional details on the estimation procedure can be found in Online Appendix OA.4.

Figure OA.5: Beta comparison between various market proxies and the VMF



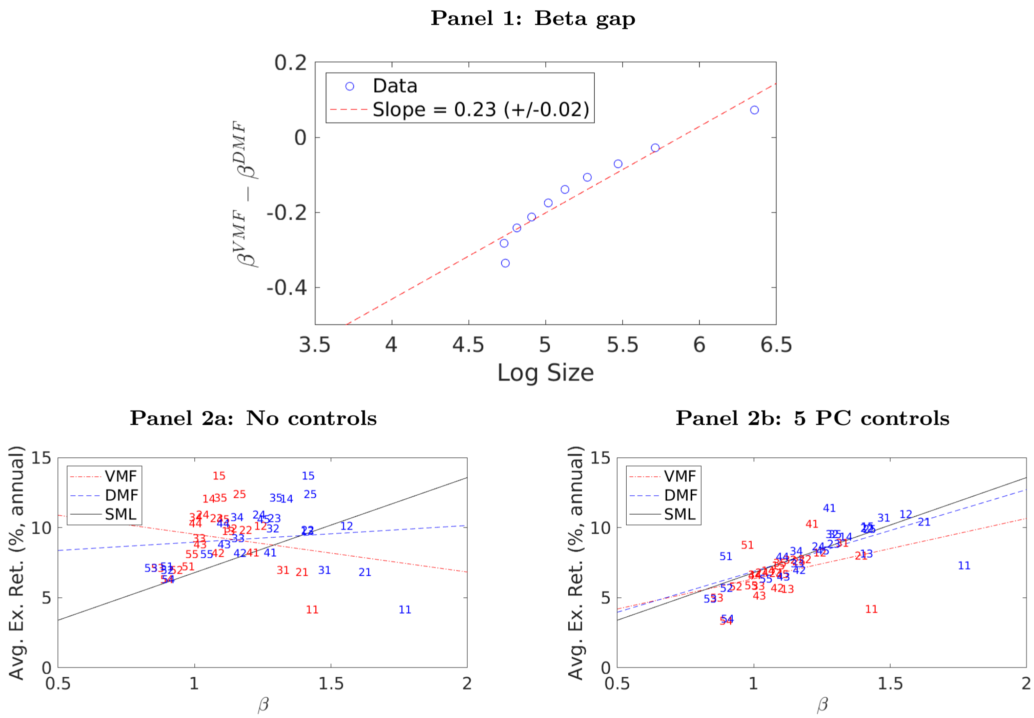
Note. This figure compares portfolio betas estimated using different proxies for the market portfolio to those estimated using the VMF. Panel 1 uses the DMF as a market proxy and Panels 2-4 use the VMF projected onto various numbers of PCs (3, 5, and 10, respectively) using the projection described in equation (OA.5). Projections of the VMF onto PCs are done using either the full sample of data or using a centered rolling window of 60 months. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Each data point represents an estimated beta from among our full (value-weighted) 372 portfolio set. The red dotted line in each plot is a 45 degree line. All data are from the 1963-2021 sample.

Figure OA.6: Time-series behavior of VMF, DMF, and IFFs estimated from alternative models



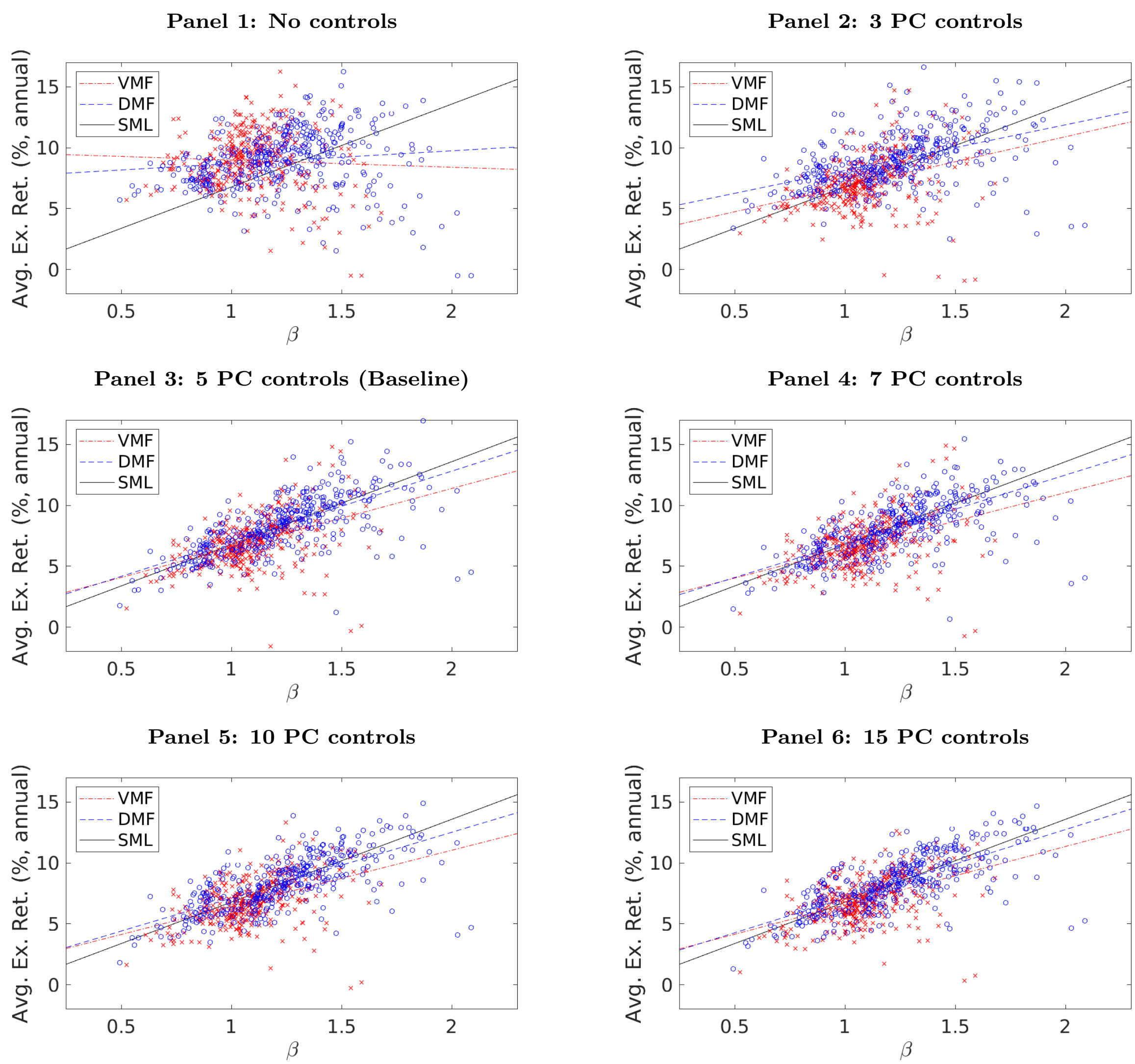
Note. This figure highlights the magnitudes of estimated IFFs and estimated volatilities of related indexes. Panel 1 plots kernel density estimates of monthly excess returns on the CRSP value-weighted index (VMF) together with densities for IFFs from DMFs estimated using either the first PC or the CRSP equal-weighted market index. Panel 2 plots in-sample forecasts for the conditional volatility of the VMF (on a log scale), obtained by estimating a GARCH(1,1) model. The blue dotted line represents the conditional volatility forecasts obtained from the DMF (using a model with one PC), whereas red solid line represents those from the VMF. Panel 3 plots log differences in those volatility forecasts. All data are from the 1963-2021 sample.

Figure OA.7: Beta gap and expected return-beta relationship for 25 size-by-book-to-market-sorted portfolios



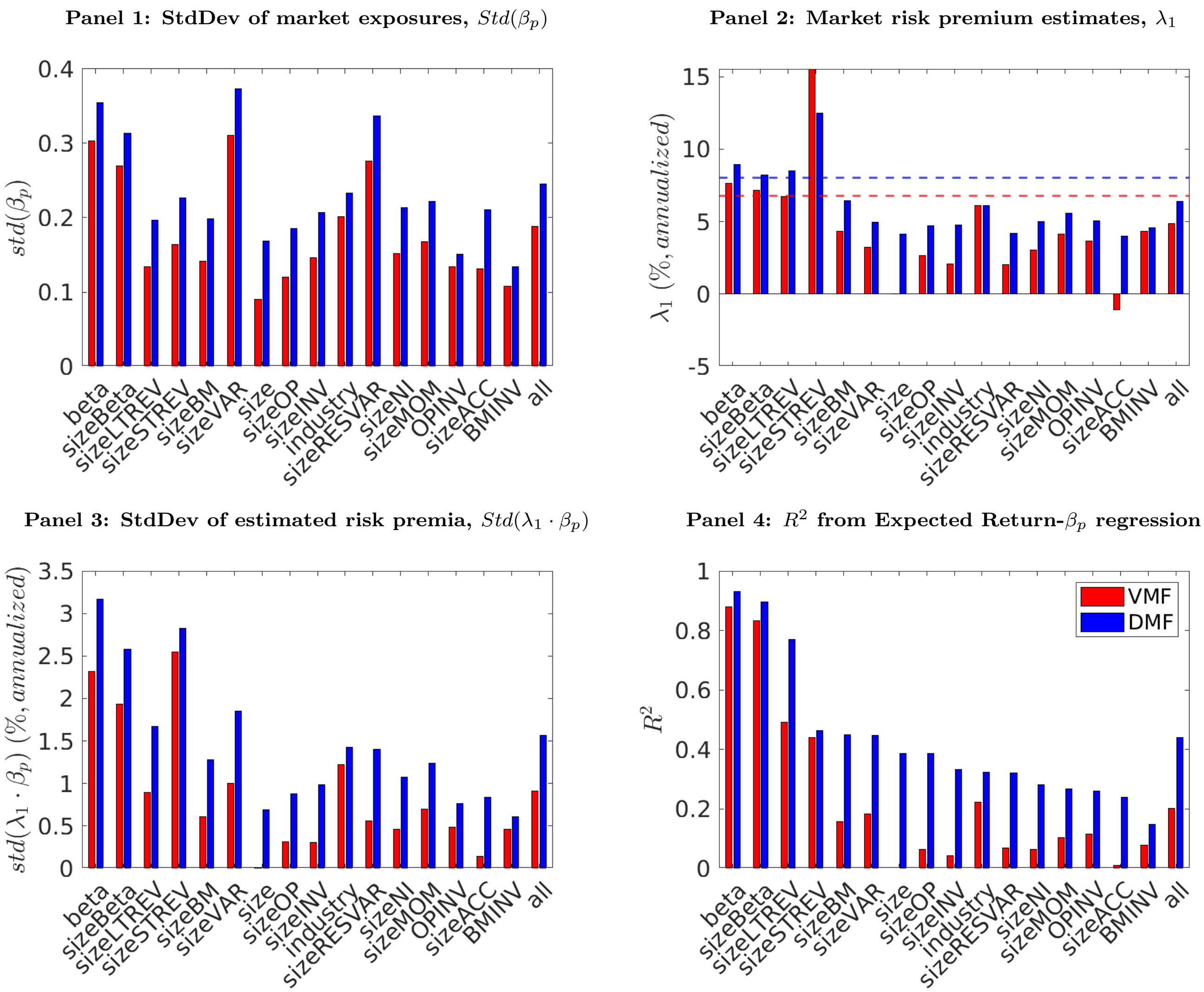
Note. This figure plots the beta gap ($\beta_p^{VMF} - \beta_p^{DMF}$) as a function of each portfolio's average log market equity (Panel 1), and SML estimates with and without the hedging procedure described by equation (25) (Panel 2) for size-by-book-to-market-sorted portfolios. Panel 1: The red line in each plot is the OLS best fit line. Log-size is computed as the average of each portfolio's monthly log market capitalization. Panel 2: Panel 2a uses unhedged portfolio returns, while Panel 2b uses returns that are hedged of exposure to the first five PCs. The red (blue) line in is constructed from full-sample market betas estimated using the VMF (DMF). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. The black line in is the SML implied by the average of VMF. Each number refers to the portfolio number in the sort. The first number refers to the first sorting variable (size), and the second number refers to the second sorting variable (book-to-market). For example, portfolio "15" is "small-value" and portfolio "51" is "large-growth." All data are from the 1963-2021 sample.

Figure OA.8: Expected return-beta relationship for all portfolios (with varying number of PCs used for hedging)



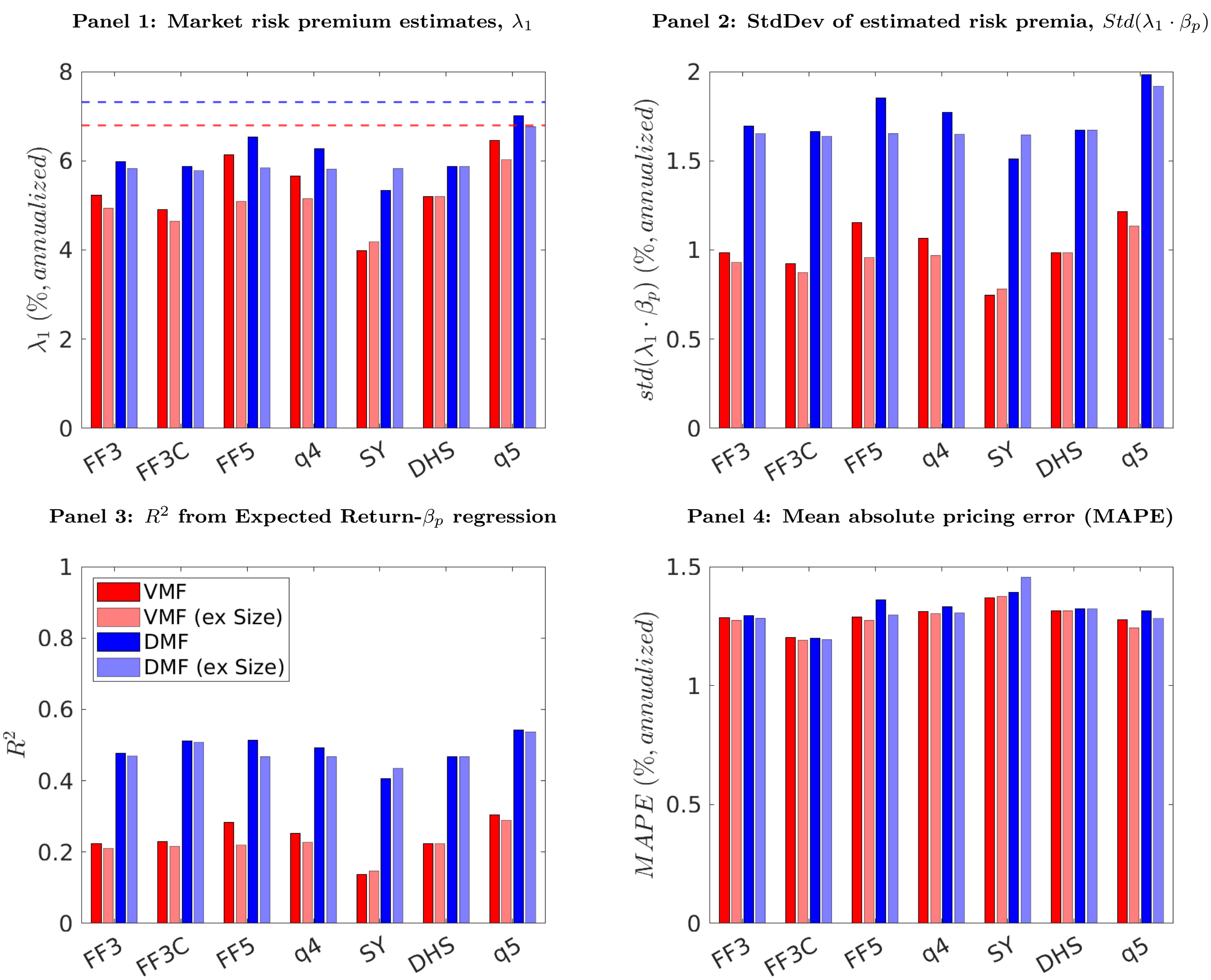
Note. This figure displays SML estimates for various portfolio sorts, with and without the hedging procedure described by equation (25) using up to 15 PCs as controls. The red (blue) line is constructed from full-sample market betas estimated using the VMF (DMF). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. The black line is the SML implied by the average of VMF. We use all 372 portfolios from the PCA as test assets in this case, which are described in Online Appendix Table OA.1. All data are from the 1963-2021 sample.

Figure OA.9: Asset pricing statistics by portfolio set (imposing normalization restrictions)



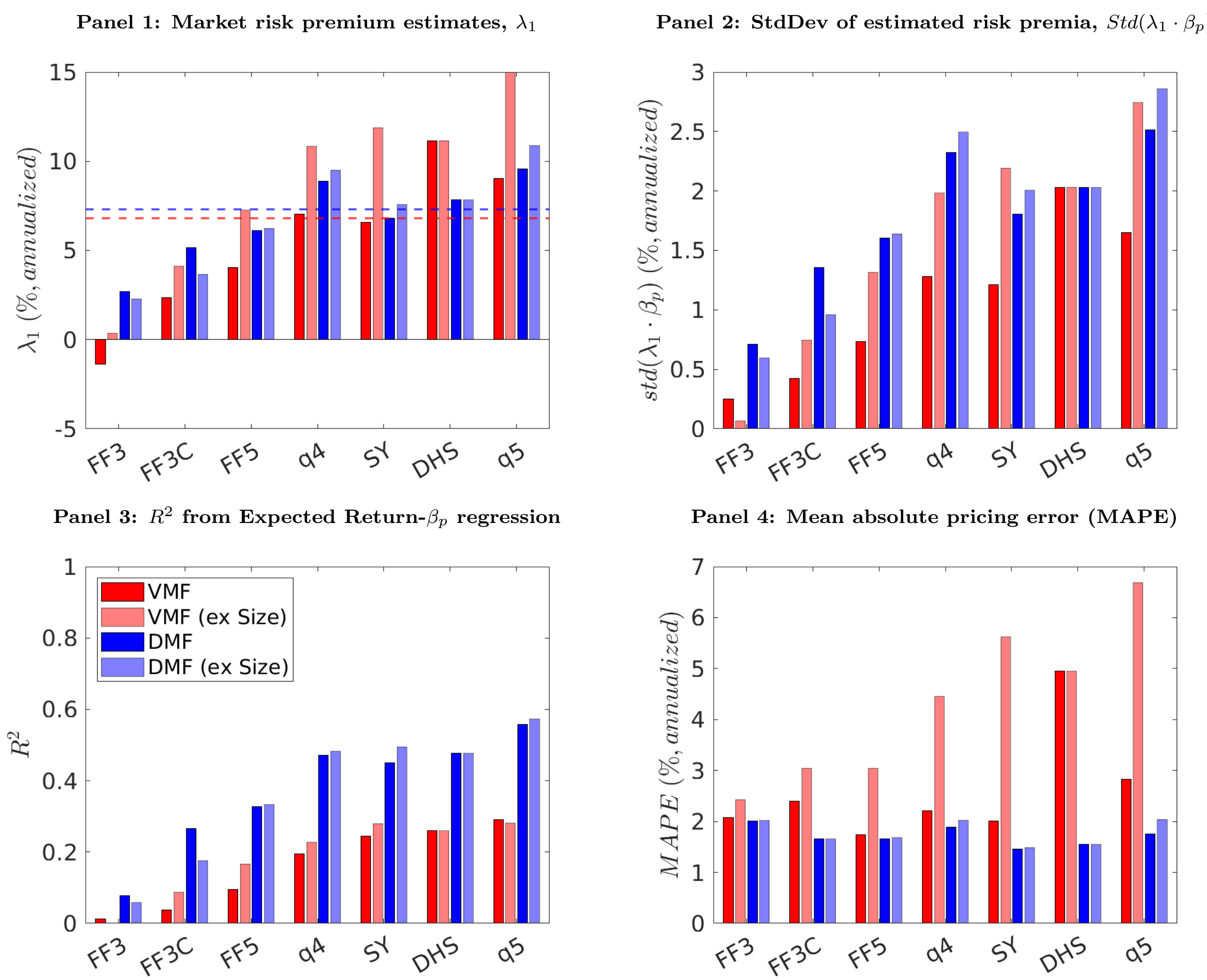
Note. This figure plots comparisons of statistics related to SMLs implied by the VMF and the DMF using specific sets of test assets (described on the x axes) to construct various respective SMLs. In this case, we rotate the first five PCs so that they satisfy the normalization restrictions in equations (3) and (4). All portfolio returns are hedged with respect to the rotated PCs using the procedure described by equation (25). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. The test asset groups are based on portfolio sets are based portfolios obtained from Kenneth French's webpage and described in Online Appendix Table OA.1. The "all" group uses all portfolios as test assets. Panel 1 plots the standard deviation of market betas of each portfolio group using each respective market proxy. Panel 2 plots the cross-sectional-regression-implied market risk premium from each set of test assets. Panel 3 plots the standard deviation of the risk premia for portfolios in each test asset set implied by each model. Panel 4 plots R-squared values implied by each cross-sectional regression. All data are from the 1963-2021 sample.

Figure OA.10: Asset pricing statistics by factor model



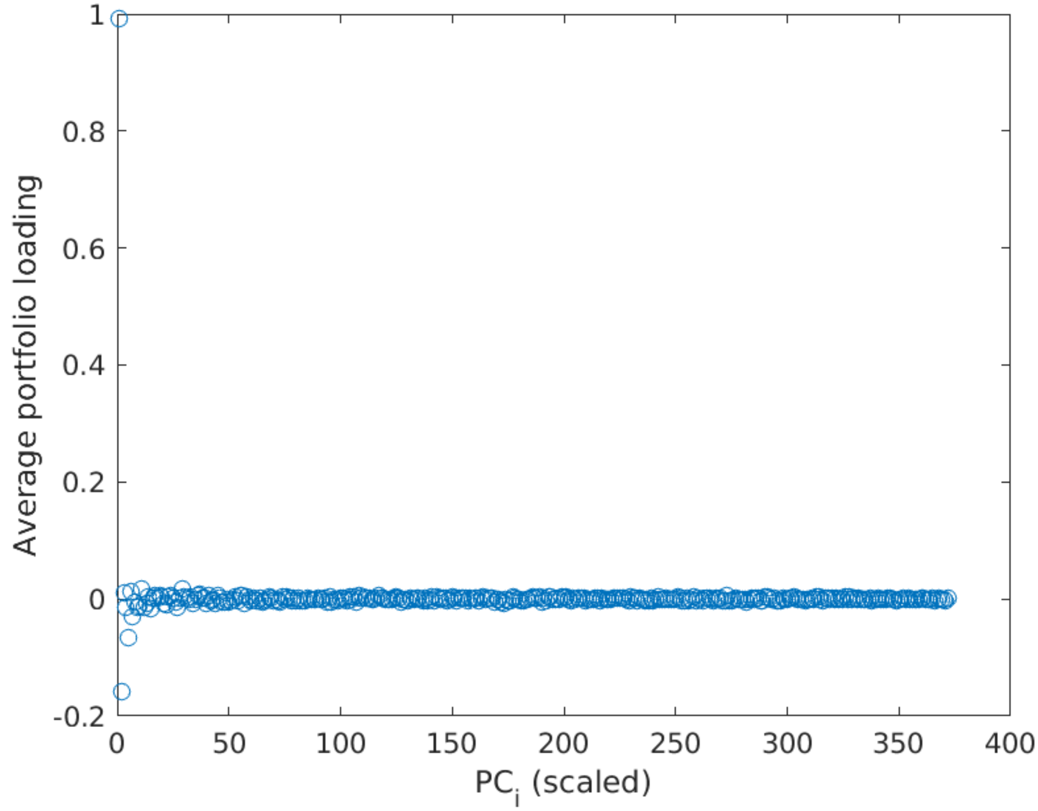
Note. This figure plots comparisons of asset pricing statistics implied by the VMF and the DMF. In this case we use portfolio returns that are hedged with respect to factors from popular factor models along with the usual five PCs using the procedure described by equation (25). In all cases, we remove the market factor from each model and replace it with the VMF or the DMF. In results "ex Size" we remove the size factor from each model (except for the DHS model, which does not include a size factor). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. We use all 372 value-weighted portfolios as test assets, which were obtained from Kenneth French's webpage and described in Online Appendix Table OA.1. Panel 1 plots the cross-sectional-regression-implied market risk premium from each factor model. Panel 2 plots the standard deviation of the risk premia for portfolios in each test asset set implied by each factor model. Panel 3 plots r-squared values implied by each cross-sectional regression. Panel 4 plots the mean absolute pricing error (α) from time series regressions of portfolio returns on all factors using each respective market proxy for each factor model. All data are from the 1963-2021 sample.

Figure OA.11: Asset pricing statistics by factor model (equal-weighted test assets)



Note. This figure plots comparisons of asset pricing statistics implied by the VMF and the DMF. In this case we use portfolio returns that are hedged with respect to factors from popular factor models rather than PCs as in Figure 8, but still according to equation (25). In all cases, we remove the market factor from each model and replace it with the VMF or the DMF. In results "ex Size" we remove the size factor from each model (except for the DHS model, which does not include a size factor). The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. We use all 372 equal-weighted portfolios as test assets, which were obtained from Kenneth French's webpage and described in Online Appendix Table OA.1. Panel 1 plots the cross-sectional-regression-implied market risk premium from each factor model. Panel 2 plots the standard deviation of the risk premia for portfolios in each test asset set implied by each factor model. Panel 3 plots r-squared values implied by each cross-sectional regression. Panel 4 plots the mean absolute pricing error (α) from time series regressions of portfolio returns on all factors using each respective market proxy for each factor model. All data are from the 1963-2021 sample.

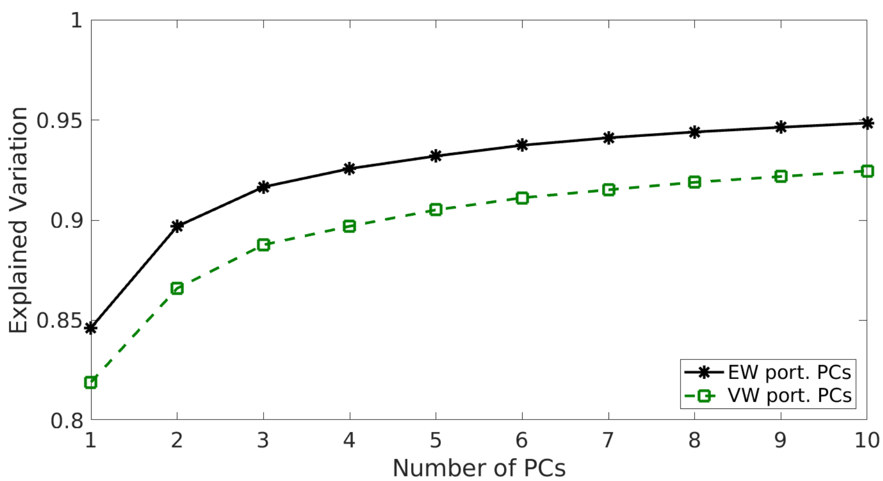
Figure OA.12: Value-weighted portfolio loadings on PCs



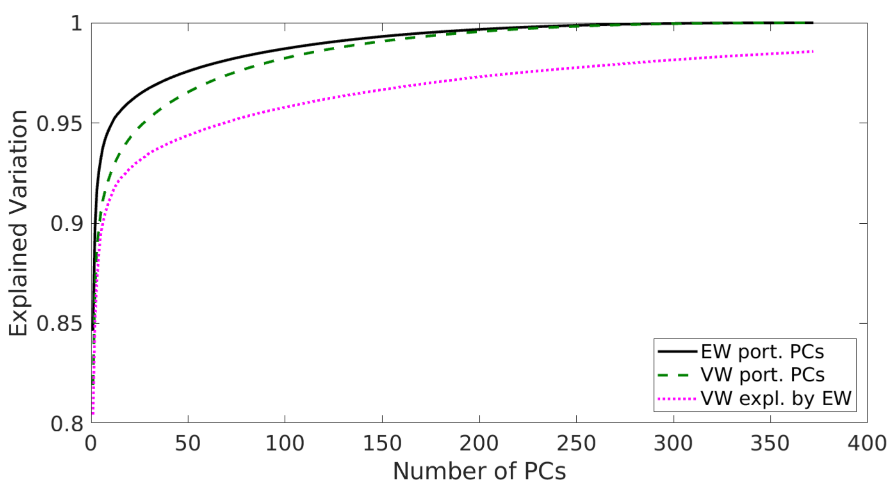
Note. This figure plots the average loadings of all 372 value-weighted portfolios we use as test assets (described in Online Appendix Table OA.1 on scaled PCs. In particular, the first PC is simply the DMF (i.e., a projection of VMF onto the first PC). Higher PCs are scaled to have the same volatilities as the VMF. The average loading of portfolios on the first (scaled) PC is computed according to equation (3) where the weights, w_i are computed as using average market capitalizations of each portfolio over the entire sample, and takes on a value of 0.993. Average portfolio loadings on higher (scaled) PCs are computed as equal-weighted averages according to equation (4). All data are from the 1963-2021 sample.

Figure OA.13: PCA explained variation

Panel 1: EW and VW PCA explained variation

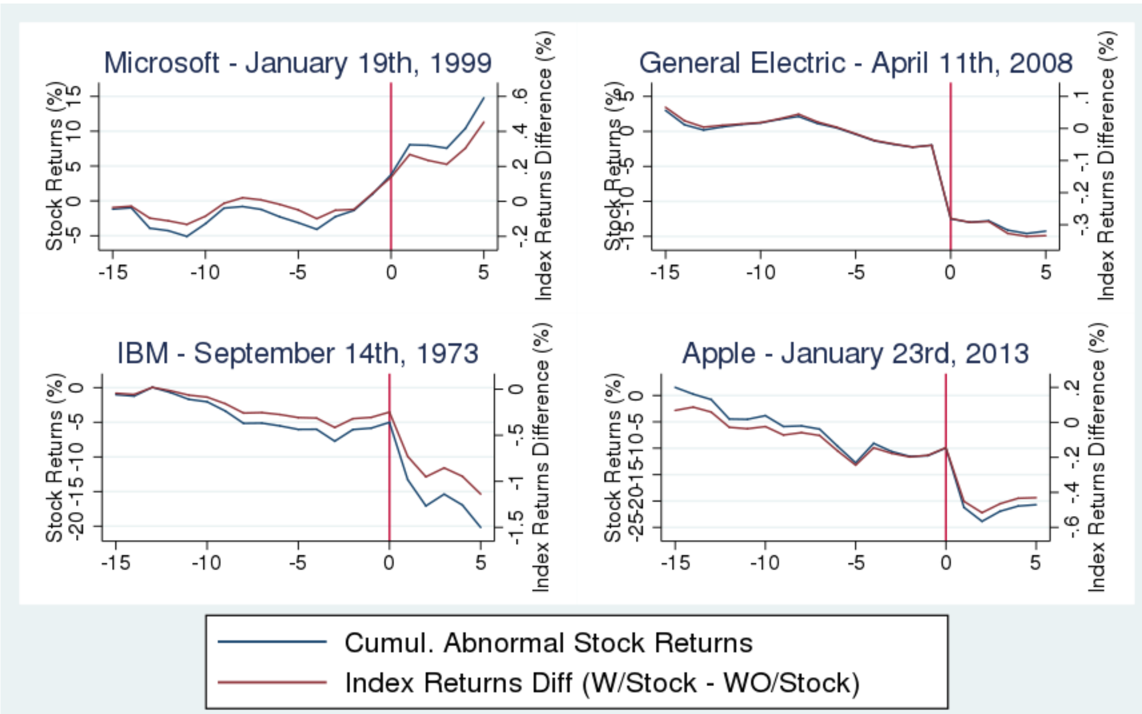


Panel 2: VW explained by EW



Note. Panel 1: This figure plots the cumulative share of variance explained by the first k principal components as a function of k, for two large sets of characteristic-sorted portfolios over the post-1963 sample. The black line plots this share of variance explained from a PCA using equal-weighted portfolios, while the green line plots the analogous object for value-weighted portfolios. Panel 2: This figure plots the same cumulative shares of variance from Panel 1 for all PCs, and plots the share of VW PC variance explained by k EW PCs. These PCAs use 372 portfolios described in Online Appendix Table OA.1. All data are from the 1963-2021 sample.

Figure OA.14: Examples of aggregate effects of announcements for megacap companies



Notes: All stock price data comes from CRSP whereas the market factors and risk-free rates to calculate abnormal returns comes from Fama-French. The market index used is value-weighted where the difference represents the value-weighted index with the individual stock contribution minus the value-weighted index without the stock contribution. Individual company betas are calculated using the three months prior to the event.

Table OA.1: Portfolios Used in Cross-Sectional Tests

Sorting variables	**Abbreviation**	**Start year**	**No. ports.**
Size/Book-to-market	sizeBM	1926	25
Size/Operating profitability	sizeOP	1963	25
Size/Investment	sizeINV	1963	25
Book-to-market/Investment	BMINV	1963	25
Operating profitability/Investment	OPINV	1963	25
Size/Momentum	sizeMOM	1927	25
Size/Short-term reversal	sizeSTREV	1926	25
Size/Long-term reversal	sizeLTREV	1931	25
Size/Accruals	sizeACC	1963	25
Size/Market beta	sizeBeta	1963	25
Size/Net share issues	sizeNI	1963	35
Size/Variance	sizeVAR	1963	25
Size/Residual variance	sizeRESVAR	1963	25
Industry	industry	1926	17
Size	size	1926	10
Market beta	beta	1963	10

Note: This table lists the portfolio sorts from Kenneth French's website used for both the principal components analysis (equal-weighted versions of the portfolios) as well as the cross-sectional asset pricing tests (value-weighted versions of the portfolios). This provides a total of 372 portfolios over our main sample period, which spans 1963-2021. The listed abbreviations are those used for plotting.

Table OA.2: How Large is the Granular Residual?

Variable	Stdev. (%)	Skew.	Kurt.	IQR (%)	$\sigma_i^2/\sigma_{VMF}^2$	$\sigma_i^2/\sigma_{IFF}^2$	σ_i/σ_{VMF}	σ_i/σ_{IFF}
Panel A: Full-sample regressions								
VMF	15.36	-0.56	5.07	18.70	100.00	636.42	100.00	252.27
IFF	6.09	-0.37	5.22	6.85	15.71	100.00	39.64	100.00
GR (1 PC)	6.09	-0.37	5.22	6.85	15.71	100.00	39.64	100.00
GR (2 PC)	4.12	-0.42	6.89	3.81	7.21	45.86	26.85	67.72
GR (3 PC)	3.59	-0.48	5.33	3.74	5.47	34.79	23.38	58.98
GR (5 PC)	2.65	-0.32	6.05	2.80	2.97	18.88	17.23	43.45
GR (7 PC)	2.62	-0.30	5.78	2.83	2.90	18.47	17.03	42.97
GR (10 PC)	2.48	-0.38	6.19	2.80	2.60	16.56	16.13	40.69
GR (15 PC)	2.21	-0.45	5.50	2.56	2.08	13.23	14.42	36.38
Panel B: Rolling regressions								
VMF	15.36	-0.56	5.07	18.70	100.00	636.42	100.00	252.27
IFF	6.09	-0.37	5.22	6.85	15.71	100.00	39.64	100.00
GR (1 PC)	5.76	-0.31	6.03	6.42	14.08	89.63	37.53	94.67
GR (2 PC)	3.76	-0.23	5.92	3.56	6.01	38.22	24.51	61.82
GR (3 PC)	3.29	-0.53	5.19	3.64	4.58	29.14	21.40	53.98
GR (5 PC)	2.21	-0.21	5.14	2.58	2.06	13.14	14.37	36.24
GR (7 PC)	2.11	-0.09	4.65	2.45	1.89	12.00	13.73	34.65
GR (10 PC)	1.92	-0.10	4.80	2.22	1.55	9.89	12.47	31.45
GR (15 PC)	1.61	-0.17	4.73	1.92	1.10	7.02	10.51	26.50

Note: This table presents summary statistic for the granular residual (GR) estimated as the residual from regressing IFF on progressively more PCs. It also provides summary statistics for the VMF and the IFF for comparison. The VMF is the CRSP value-weighted market index and the IFF is that implied when the DMF is constructed using one PC. The IFF is the sum of two components: 1) time-varying market exposure to the first PC, 2) time-varying market exposures to higher PCs, and 2) the GR. We estimate the GR by regressing the IFF on various numbers of PCs listed in columns of the table, and computing the implied GR as $GR \approx IFF - IFF_{proj}$ where IFF_proj is the projection of the IFF onto the specified number of PCs. Panel A presents results that run full-sample regressions, effectively imposing that IFF loadings on PCs are constant. Panel B presents results using centered rolling regressions of the IFF on the specified number of PCs with a 60-month window (i.e., in any 60-month window, if month t is the month of interest for estimating the GR, the regression is estimated using data from months $t-30$ to $t+29$). The last four columns present variance or standard deviation estimates of the variance and standard deviation of each variable in column one relative to those from two benchmarks: the VMF and the IFF. σ_i represents the resulting time series standard deviation with $i \in \{GR, VMF, IFF\}$ and where GR is estimated as described above. All statistics are annualized and in percent. We annualize the Standard Deviation and IQR statistics by multiplying monthly frequency values by $\sqrt{12}$. All data are from the 1963-2021 sample.

Table OA.3: Risk-return tradeoff coefficient (γ) across GARCH models (1927-2021)

Risk defn.	Parameter	VMF	DMF (1PC)	DMF (EW)
Panel A: GARCH				
Variance	γ	1.47*	2.56***	2.51***
(σ_t^2)		(0.76)	(0.73)	(0.77)
	$\gamma^{VMF} - \gamma^{DMF}$		-1.09	-1.04
Volatility	γ	0.16	0.35***	0.32***
(σ_t)		(0.11)	(0.09)	(0.09)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.19	-0.16
Panel B: GJR-GARCH				
Variance	γ	1.25*	3.21***	3.06***
(σ_t^2)		(0.67)	(0.50)	(0.57)
	$\gamma^{VMF} - \gamma^{DMF}$		-1.96	-1.81
Volatility	γ	0.12	0.33***	0.31***
(σ_t)		(0.10)	(0.08)	(0.08)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.21	-0.19
Panel C: E-GARCH				
Variance	γ	1.78**	3.60***	3.60***
(σ_t^2)		(0.78)	(0.72)	(0.76)
	$\gamma^{VMF} - \gamma^{DMF}$		-1.82	-1.82
Volatility	γ	0.14	0.23***	0.26***
(σ_t)		(0.10)	(0.09)	(0.09)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.09	-0.11

Note: This table compares γ estimates and standard errors (in parentheses) from the VMF with those from DMF variants. The first column corresponds with results using the VMF, which is the CRSP value-weighted index. Columns two and three correspond to results from DMF variants that are constructed using either the first PC from the PCA on equal-weighted portfolios described in Online Appendix Table OA.1 or the CRSP equal-weighted market index. All returns are in excess of the risk-free rate. We consider three popular GARCH volatility forecasting models such as a GARCH (1,1) (Panel A), a Threshold GARCH model of Glosten et al. (1993) (Panel B) and an exponential GARCH model of Nelson (1991) (Panel C). Within each panel, the first (second) set of results is based on the conditional variance (volatility) as a proxy for risk. Rows labeled $\gamma^{VMF} - \gamma^{DMF}$ report the difference between γ's estimated using the VMF and the respective DMF variants. We report outer product standard errors (Hamilton, Time Series Analysis, 1994, pp. 133-148) in parentheses. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1927-2021 sample.

Table OA.4: Difference between unadjusted and adjusted index betas for size × beta-sorted portfolios (1993-2021)

Market beta quintile	Size quintile					Small - Large
	Small	2	3	4	Large	
Low	-0.24	-0.17	-0.08	-0.02	0.08	-0.32
	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)	(0.07)
2	-0.25	-0.19	-0.13	-0.06	0.05	-0.30
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
3	-0.28	-0.21	-0.15	-0.10	0.03	-0.31
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
4	-0.30	-0.25	-0.19	-0.14	-0.01	-0.30
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
High	-0.40	-0.32	-0.23	-0.18	-0.04	-0.36
	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.06)
High-Low	-0.15	-0.15	-0.15	-0.16	-0.12	
	(0.06)	(0.03)	(0.03)	(0.04)	(0.03)	

Note: This table reports estimates of the bias induced in market betas associated with the presence of the IFF in cross-sectional asset pricing tests, which tests the predictions of equation (27). In particular, this table reports $\beta^{VMF} - \beta^{DMF}$. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Our model predicts this difference is increasing in β and decreasing in size, consistent with the results in the table. Estimates are from monthly returns on 25 size-by-beta-sorted portfolios from Kenneth French's data library. Standard errors are in parentheses and are computed using the bootstrap procedure described in Online Appendix OA.11. All data are from the 1993-2021 sample.

Table OA.5: Comparing VMF and DMF asset pricing statistics (1993-2021 sample)

	Number of orthogonalization PCs					
	0	3	5	7	10	15
Panel A: Market risk premium estimates (λ_1)						
DMF	0.29	3.39	5.04*	6.46**	6.41**	6.59**
$[\mathbb{E}\left[r_{t+1}^{DMF}\right]$ = 8.26, se = 2.51]	(3.18)	(2.77)	(2.66)	(2.62)	(2.62)	(2.60)
VMF	-0.69	4.82	5.04*	6.84**	6.57**	6.99**
$[\mathbb{E}\left[r_{t+1}^{VMF}\right]$ = 8.80, se = 2.79]	(4.22)	(3.08)	(3.06)	(2.97)	(2.94)	(2.92)
DMF - VMF	0.98	-1.43	0.00	-0.38	-0.16	-0.41
$[\mathbb{E}\left[r_{t+1}^{DMF} - r_{t+1}^{VMF}\right]$ = -0.54, se = 1.11]	(1.34)	(1.78)	(1.43)	(1.30)	(1.29)	(1.27)
$p\,(DMF < VMF)$	0.23	0.77	0.50	0.60	0.54	0.60
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)						
DMF	0.09	1.08	1.61*	2.06**	2.04**	2.10**
$[std\left(\beta_p^{DMF}\right)$ = 0.32, se = 0.02]	(0.64)	(0.79)	(0.85)	(0.87)	(0.87)	(0.86)
VMF	0.16	1.09*	1.14*	1.54**	1.48**	1.58**
$[std\left(\beta_p^{VMF}\right)$ = 0.23, se = 0.02]	(0.59)	(0.66)	(0.67)	(0.69)	(0.68)	(0.68)
DMF - VMF	-0.06	-0.01	0.47	0.52	0.56	0.52
$[std\left(\beta_p^{DMF}\right) - std\left(\beta_p^{VMF}\right)$ = 0.09, se = 0.01]	(0.34)	(0.49)	(0.42)	(0.39)	(0.40)	(0.39)
$p\,(DMF < VMF)$	0.46	0.42	0.12	0.08	0.06	0.07
Panel C: SML R-squared						
DMF	0.00	0.19	0.41**	0.55***	0.55***	0.55***
	(0.13)	(0.17)	(0.20)	(0.19)	(0.19)	(0.19)
VMF	0.01	0.21	0.25	0.38**	0.36**	0.37**
	(0.12)	(0.14)	(0.16)	(0.17)	(0.17)	(0.17)
DMF - VMF	0.00	-0.02	0.17	0.17	0.19	0.19*
	(0.07)	(0.11)	(0.10)	(0.09)	(0.09)	(0.09)
$p\,(DMF < VMF)$	0.46	0.47	0.12	0.08	0.06	0.05
Panel D: Mean absolute pricing error (α)						
DMF	2.55***	2.09***	1.64***	1.51***	1.52***	1.54***
	(0.49)	(0.26)	(0.17)	(0.13)	(0.13)	(0.12)
VMF	2.59***	1.66***	1.64***	1.53***	1.53***	1.61***
	(0.58)	(0.18)	(0.15)	(0.12)	(0.12)	(0.12)
DMF - VMF	-0.04	0.43**	0.00	-0.02	-0.01	-0.06
	(0.47)	(0.21)	(0.07)	(0.05)	(0.05)	(0.04)
$p\,(DMF < VMF)$	0.54	0.01	0.43	0.58	0.60	0.91

Note: This table presents asset pricing statistics when using all value-weighted portfolios as test assets and hedging their returns with respect to different numbers of PCs using the hedging procedure described by equation (25). We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports time series averages of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$), the VMF ($\mathbb{E}[r_{t+1}^{VMF}]$), and their difference. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. The first column in Panel B also reports the standard deviation of portfolio exposures to the DMF ($std(\beta_p^{DMF})$), the VMF ($std(\beta_p^{VMF})$), and their difference. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that use either the DMF or the VMF, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported DMF statistics are less than the VMF statistics based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1993-2021 sample.

Table OA.6: Comparing VMF and DMF asset pricing statistics (equal-weighted test assets)

	Number of orthogonalization PCs					
	0	3	5	7	10	15
Panel A: Market risk premium estimates (λ_1)						
DMF	1.06	5.28***	6.94***	7.05***	6.75***	6.95***
$[\mathbb{E}\left[r_{t+1}^{DMF}\right] = 7.31$, se $= 1.83]$	(2.50)	(1.88)	(1.86)	(1.84)	(1.84)	(1.84)
VMF	-2.21	4.55**	6.05***	6.22***	5.80***	6.12***
$[\mathbb{E}\left[r_{t+1}^{VMF}\right] = 6.79$, se $= 2.00]$	(3.10)	(2.08)	(2.07)	(2.04)	(2.01)	(2.01)
DMF - VMF	3.27***	0.72	0.89	0.83	0.94	0.83
$[\mathbb{E}\left[r_{t+1}^{DMF} - r_{t+1}^{VMF}\right] = 0.53$, se $= 0.79]$	(1.21)	(1.07)	(0.84)	(0.82)	(0.80)	(0.79)
$p\,(DMF < VMF)$	0.00	0.25	0.16	0.17	0.14	0.16
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)						
DMF	0.28	1.39***	1.82***	1.85***	1.77***	1.83***
$[std\left(\beta_p^{DMF}\right) = 0.28$, se $= 0.01]$	(0.43)	(0.52)	(0.52)	(0.51)	(0.52)	(0.51)
VMF	0.40	0.83**	1.10***	1.13***	1.05***	1.11***
$[std\left(\beta_p^{VMF}\right) = 0.19$, se $= 0.01]$	(0.41)	(0.38)	(0.39)	(0.38)	(0.38)	(0.38)
DMF - VMF	-0.12	0.56**	0.73***	0.72***	0.72***	0.71***
$[std\left(\beta_p^{DMF}\right) - std\left(\beta_p^{VMF}\right) = 0.10$, se $= 0.01]$	(0.58)	(0.28)	(0.24)	(0.24)	(0.24)	(0.23)
$p\,(DMF < VMF)$	0.55	0.02	0.00	0.00	0.00	0.00
Panel C: SML R-squared						
DMF	0.01	0.29**	0.48***	0.50***	0.54***	0.58***
	(0.06)	(0.13)	(0.14)	(0.14)	(0.14)	(0.14)
VMF	0.02	0.14	0.25**	0.27**	0.29**	0.33**
	(0.07)	(0.09)	(0.11)	(0.12)	(0.13)	(0.13)
DMF - VMF	-0.01	0.15*	0.23***	0.23***	0.25***	0.24***
	(0.08)	(0.08)	(0.06)	(0.06)	(0.07)	(0.06)
$p\,(DMF < VMF)$	0.55	0.04	0.00	0.00	0.00	0.00
Panel D: Mean absolute pricing error (α)						
DMF	2.41***	1.58***	1.35***	1.28***	1.20***	1.11***
	(0.24)	(0.11)	(0.08)	(0.07)	(0.06)	(0.05)
VMF	3.31***	1.47***	1.36***	1.29***	1.20***	1.11***
	(0.59)	(0.09)	(0.08)	(0.07)	(0.06)	(0.05)
DMF - VMF	-0.90*	0.11*	-0.01	-0.01	-0.01	0.00
	(0.52)	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)
$p\,(DMF < VMF)$	0.97	0.03	0.75	0.58	0.65	0.68

Note: This table presents asset pricing statistics when using all equal-weighted portfolios as test assets and hedging their returns with respect to different numbers of PCs using the hedging procedure described by equation (25). We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports time series averages of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$), the VMF ($\mathbb{E}[r_{t+1}^{VMF}]$), and their difference. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. The first column in Panel B also reports the standard deviation of portfolio exposures to the DMF ($std(\beta_p^{DMF})$), the VMF ($std(\beta_p^{VMF})$), and their difference. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that use either the DMF or the VMF, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported DMF statistics are less than the VMF statistics based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table OA.7: Comparing VMF and DMF asset pricing statistics (only factor model controls)

	Factor model control							
	None	FF3	FF3C	FF5	q4	SY	DHS	q5
Panel A: Market risk premium estimates (λ_1)								
DMF	1.05	2.28	3.90*	5.42***	7.04***	5.33**	6.09**	8.23***
$[\mathbb{E}\left[r_{t+1}^{DMF}\right] = 7.31$, se $= 1.83]$	(2.34)	(1.96)	(2.04)	(1.86)	(2.05)	(2.20)	(2.60)	(2.09)
VMF	-0.59	-0.94	1.99	4.58**	7.07***	5.70**	12.86***	9.17***
$[\mathbb{E}\left[r_{t+1}^{VMF}\right] = 6.79$, se $= 2.00]$	(3.20)	(2.38)	(2.44)	(2.22)	(2.46)	(2.64)	(3.39)	(2.54)
DMF - VMF	1.64	3.22***	1.91**	0.84	-0.02	-0.37	-6.77***	-0.94
$[\mathbb{E}\left[r_{t+1}^{DMF} - r_{t+1}^{VMF}\right] = 0.53$, se $= 0.79]$	(1.12)	(0.99)	(0.95)	(0.91)	(1.01)	(1.03)	(1.45)	(1.03)
$p\,(DMF < VMF)$	0.06	0.00	0.02	0.17	0.49	0.64	1.00	0.80
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)								
DMF	0.30	0.65	1.10*	1.54***	1.99***	1.53**	1.70**	2.33***
$[std\left(\beta_p^{DMF}\right) = 0.28$, se $= 0.01]$	(0.43)	(0.47)	(0.56)	(0.55)	(0.61)	(0.65)	(0.74)	(0.62)
VMF	0.11	0.18	0.37	0.86**	1.34***	1.09**	2.38***	1.73***
$[std\left(\beta_p^{VMF}\right) = 0.19$, se $= 0.01]$	(0.38)	(0.29)	(0.35)	(0.41)	(0.48)	(0.51)	(0.67)	(0.50)
DMF - VMF	0.19	0.47	0.73	0.68**	0.66***	0.44*	-0.68**	0.59**
$[std\left(\beta_p^{DMF}\right) - std\left(\beta_p^{VMF}\right) = 0.10$, se $= 0.01]$	(0.39)	(0.59)	(0.38)	(0.25)	(0.26)	(0.26)	(0.32)	(0.26)
$p\,(DMF < VMF)$	0.45	0.32	0.08	0.01	0.00	0.05	0.98	0.01
Panel C: SML R-squared								
DMF	0.01	0.08	0.26	0.37**	0.57***	0.42**	0.50***	0.65***
	(0.08)	(0.10)	(0.16)	(0.15)	(0.15)	(0.18)	(0.19)	(0.14)
VMF	0.00	0.01	0.04	0.17	0.39***	0.28*	0.48***	0.51***
	(0.07)	(0.05)	(0.09)	(0.11)	(0.14)	(0.15)	(0.12)	(0.14)
DMF - VMF	0.01	0.07	0.22	0.20**	0.18**	0.13	0.02	0.15**
	(0.07)	(0.12)	(0.10)	(0.07)	(0.07)	(0.07)	(0.15)	(0.07)
$p\,(DMF < VMF)$	0.45	0.33	0.08	0.01	0.01	0.06	0.56	0.02
Panel D: Mean absolute pricing error (α)								
DMF	2.06***	1.76***	1.56***	1.38***	1.66***	1.63***	1.50***	1.69***
	(0.25)	(0.18)	(0.16)	(0.11)	(0.26)	(0.24)	(0.28)	(0.31)
VMF	2.61***	1.77***	1.58***	1.31***	1.27***	1.17***	2.92***	1.49***
	(0.41)	(0.17)	(0.19)	(0.10)	(0.12)	(0.12)	(0.74)	(0.18)
DMF - VMF	-0.55	0.00	-0.02	0.07*	0.39*	0.45**	-1.42*	0.20
	(0.36)	(0.09)	(0.19)	(0.07)	(0.22)	(0.23)	(0.71)	(0.27)
$p\,(DMF < VMF)$	0.90	0.45	0.53	0.05	0.04	0.02	0.97	0.22

Note: This table presents asset pricing statistics when using all value-weighted portfolios as test assets and hedging their returns with respect to factors from prominent models using the hedging procedure described by equation (25) (and not controlling for any PCs). We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Each column corresponds to controlling for factors in different models where the acronyms are defined in the text. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports time series averages of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$), the VMF ($\mathbb{E}[r_{t+1}^{VMF}]$), and their difference. The VMF is the CRSP value-weighted market index and the DMF is that constructed using one PC. Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. The first column in Panel B also reports the standard deviation of portfolio exposures to the DMF ($std(\beta_p^{DMF})$), the VMF ($std(\beta_p^{VMF})$), and their difference. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that use either the DMF or the VMF, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported DMF statistics are less than the VMF statistics based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table OA.8: Comparing VMF asset pricing statistics with and without size factors

	Factor model control							
	None	FF3	FF3C	FF5	q4	SY	DHS	q5
Panel A: Market risk premium estimates (λ_1)								
VMF (w/ size factor)	-0.59	-0.94	1.99	4.58**	7.07***	5.70**	12.86***	9.17***
$[\mathbb{E}\left[r_{t+1}^{VMF}\right] = 6.79$, se $= 2.00]$	(3.20)	(2.38)	(2.44)	(2.22)	(2.46)	(2.64)	(3.39)	(2.54)
VMF (w/o size factor)		1.39	4.37	9.16***	12.85***	14.03***		18.34***
		(3.17)	(3.32)	(2.86)	(3.16)	(3.33)		(3.29)
w/ Size. - w/o Size		-2.33	-2.38	-4.58**	-5.79***	-8.33***		-9.16***
		(2.11)	(2.11)	(1.84)	(2.04)	(1.98)		(2.21)
$p\,(w/\,Size. < w/o\,Size)$		0.84	0.84	0.99	1.00	1.00		1.00
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)								
VMF (w/ size factor)	0.11	0.18	0.37	0.86**	1.34***	1.09**	2.38***	1.73***
	(0.38)	(0.29)	(0.35)	(0.41)	(0.48)	(0.51)	(0.67)	(0.50)
VMF (w/o size factor)		0.26	0.82	1.72***	2.43***	2.67***		3.47***
		(0.39)	(0.54)	(0.55)	(0.62)	(0.67)		(0.65)
w/ Size. - w/o Size		-0.08	-0.45	-0.86**	-1.09***	-1.59***		-1.73***
		(0.43)	(0.40)	(0.35)	(0.39)	(0.39)		(0.42)
$p\,(w/\,Size. < w/o\,Size)$		0.59	0.80	0.99	1.00	1.00		1.00
Panel C: SML R-squared								
VMF (w/ size factor)	0.00	0.01	0.04	0.17	0.39***	0.28*	0.48***	0.51***
	(0.07)	(0.05)	(0.09)	(0.11)	(0.14)	(0.15)	(0.12)	(0.14)
VMF (w/o size factor)		0.01	0.15	0.35***	0.47***	0.49***		0.50***
		(0.07)	(0.13)	(0.12)	(0.11)	(0.11)		(0.09)
w/ Size. - w/o Size		-0.01	-0.11	-0.18**	-0.08*	-0.21**		0.01
		(0.07)	(0.08)	(0.06)	(0.06)	(0.09)		(0.08)
$p\,(w/\,Size. < w/o\,Size)$		0.58	0.80	0.99	0.96	0.99		0.75
Panel D: Mean absolute pricing error (α)								
VMF (w/ size factor)	2.61***	1.77***	1.58***	1.31***	1.27***	1.17***	2.92***	1.49***
	(0.41)	(0.17)	(0.19)	(0.10)	(0.12)	(0.12)	(0.74)	(0.18)
VMF (w/o size factor)		1.96***	1.88***	2.20***	2.79***	3.44***		4.48***
		(0.29)	(0.37)	(0.48)	(0.64)	(0.73)		(0.77)
w/ Size. - w/o Size		-0.19	-0.30	-0.89***	-1.52***	-2.27***		-2.99***
		(0.23)	(0.32)	(0.47)	(0.62)	(0.75)		(0.74)
$p\,(w/\,Size. < w/o\,Size)$		0.90	0.86	1.00	1.00	1.00		1.00

Note: This table presents asset pricing statistics when using all value-weighted portfolios as test assets and hedging their returns with respect to factors from prominent models using the hedging procedure described by equation (25) (and not controlling for any PCs). In this case, we only consider models that use the VMF (the CRSP value-weighted excess market return) and compare results when controlling for models including or excluding their size factors. We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Each column corresponds to controlling for factors in different models where the acronyms are defined in the text. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports the time series average of the VMF ($\mathbb{E}[r_{t+1}^{VMF}]$). Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that either include or exclude each model's size factor, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported statistics from including size factors are less than statistics from excluding size factors based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. Note that the DHS model does not include a size factor. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.

Table OA.9: Comparing DMF asset pricing statistics with and without size factors

	Factor model control							
	None	FF3	FF3C	FF5	q4	SY	DHS	q5
Panel A: Market risk premium estimates (λ_1)								
DMF (w/ size factor)	1.05	2.28	3.90*	5.42***	7.04***	5.33**	6.09**	8.23***
$[\mathbb{E}\left[r_{t+1}^{DMF}\right] = 7.31$, se $= 1.83]$	(2.34)	(1.96)	(2.04)	(1.86)	(2.05)	(2.20)	(2.60)	(2.09)
DMF (w/o size factor)		1.79	2.22	5.57**	7.84***	6.25**		9.87***
		(2.33)	(2.40)	(2.24)	(2.54)	(2.52)		(2.63)
w/ Size. - w/o Size		0.49	1.68	-0.15	-0.80	-0.92		-1.64
		(1.30)	(1.29)	(1.26)	(1.50)	(1.34)		(1.58)
$p\,(w/\,Size. < w/o\,Size)$		0.31	0.09	0.51	0.67	0.73		0.84
Panel B: Cross-sectional dispersion in estimated risk premia ($std\,(\lambda_1 \cdot \beta_p)$)								
DMF (w/ size factor)	0.30	0.65	1.10*	1.54***	1.99***	1.53**	1.70**	2.33***
	(0.43)	(0.47)	(0.56)	(0.55)	(0.61)	(0.65)	(0.74)	(0.62)
DMF (w/o size factor)		0.51	0.63	1.58**	2.22***	1.79**		2.79***
		(0.48)	(0.52)	(0.65)	(0.75)	(0.74)		(0.78)
w/ Size. - w/o Size		0.14	0.48	-0.04	-0.23	-0.26		-0.46
		(0.35)	(0.41)	(0.36)	(0.43)	(0.38)		(0.45)
$p\,(w/\,Size. < w/o\,Size)$		0.40	0.17	0.51	0.67	0.73		0.84
Panel C: SML R-squared								
DMF (w/ size factor)	0.01	0.08	0.26	0.37**	0.57***	0.42**	0.50***	0.65***
	(0.08)	(0.10)	(0.16)	(0.15)	(0.15)	(0.18)	(0.19)	(0.14)
DMF (w/o size factor)		0.05	0.11	0.38**	0.60***	0.50***		0.69***
		(0.10)	(0.14)	(0.16)	(0.16)	(0.19)		(0.13)
w/ Size. - w/o Size		0.03	0.15	-0.01	-0.03	-0.09		-0.04
		(0.07)	(0.10)	(0.08)	(0.07)	(0.10)		(0.06)
$p\,(w/\,Size. < w/o\,Size)$		0.41	0.18	0.51	0.66	0.73		0.83
Panel D: Mean absolute pricing error (α)								
DMF (w/ size factor)	2.06***	1.76***	1.56***	1.38***	1.66***	1.63***	1.50***	1.69***
	(0.25)	(0.18)	(0.16)	(0.11)	(0.26)	(0.24)	(0.28)	(0.31)
DMF (w/o size factor)		1.80***	1.56***	1.39***	1.81***	1.70***		2.05***
		(0.22)	(0.18)	(0.15)	(0.34)	(0.30)		(0.43)
w/ Size. - w/o Size		-0.04	-0.01	-0.01	-0.14	-0.07		-0.36
		(0.14)	(0.17)	(0.11)	(0.25)	(0.15)		(0.33)
$p\,(w/\,Size. < w/o\,Size)$		0.75	0.66	0.72	0.67	0.76		0.84

Note: This table presents asset pricing statistics when using all value-weighted portfolios as test assets and hedging their returns with respect to factors from prominent models using the hedging procedure described by equation (25) (and not controlling for any PCs). In this case, we only consider models that use the DMF (the CRSP value-weighted excess market return projected onto the first PC) and compare results when controlling for models including or excluding their size factors. We use all 372 value-weighted portfolios (obtained from Kenneth French's data library and described in more detail in Online Appendix Table OA.1) as test assets. Each column corresponds to controlling for factors in different models where the acronyms are defined in the text. Panel A reports the cross-sectional-regression-implied market risk premium. The first column in Panel A also reports the time series average of the DMF ($\mathbb{E}[r_{t+1}^{DMF}]$). Panel B reports the standard deviation of the risk premia on all portfolios implied by each model. Panel C reports R-squared values implied by each full-sample cross-sectional regression. Panel D reports the mean absolute pricing error (α_p) across all portfolios from time series regressions that either include or exclude each model's size factor, as well as their difference. The last row in each panel reports the p-value associated with the test that the reported statistics from including size factors are less than statistics from excluding size factors based on the bootstrap procedure described in Online Appendix OA.11. Results presented in Panels A, B, and D are annualized and in percent. Note that the DHS model does not include a size factor. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from the 1963-2021 sample.