The Idiosyncratic Financial Factor: An Explanation for the Role of Size Factors and the Weak Intertemporal Risk-Return Relation*

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

November 2024

Abstract

In addition to a priced, dominant market factor (DMF), the value-weighted stock market return contains an "idiosyncratic financial factor" (IFF) related to overweighting of large-cap stocks. The IFF carries no risk premium, is unrelated to macroeconomic factors and returns in other markets, and significantly impacts systematic risk estimates. Size factors separate exposures to the DMF from the IFF. Consistent with a model with nontraded assets, using the DMF as an alternative market factor resolves the size anomaly and obviates the need for size factors in multifactor models. Finally, the DMF generates a stronger intertemporal risk-return tradeoff.

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

* *Keywords*: Idiosyncratic risk; Market factor; Size effect; Size factors; Intertemporal risk-return relation

*This is a substantially revised version of an earlier manuscript that was circulated under the title "Idiosyncratic financial risk and a reevaluation of the market risk-return tradeoff." Results here also substantially extend and supersede those from earlier work entitled "Real risk or paper risk? Mis-measured factors, granular measurement errors, and empirical asset pricing tests." We are grateful to John Campbell, John Cochrane, Zhi Da, Xavier Gabaix, Stefano Giglio, Daniel Greenwald, Evan Jo, Leonid Kogan, Ralph Koijen, Arvind Krishnamurthy, Eben Lazarus, Gang Li, Lira Mota, Jonathan Parker, Seth Pruitt, Cesare Robotti, Andrea Tamoni, Jun Tu, Yaxuan Wen, Nancy Xu and seminar participants at the American Finance Association, the Northern Finance Association, the Financial Intermediation Research Society, the BI-SHoF Conference on Asset Pricing, the China International Conference in Finance, the Eastern Finance Association, the NBER Asset Pricing meeting, the City University of Hong Kong, McGill, MIT, Rutgers, Stanford, the University of Rochester, and the University of Southern California 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.

†Supervisory Policy, 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 value-weighted US equity market index is not a well-diversified portfolio. Its returns are particularly sensitive to the returns on a small number of large-cap stocks that dominate the index.[1] This lack of diversification is not a recent phenomenon. Rather, it persists throughout the entire history of the US stock market. Whether such a feature is desirable is unclear. Concentrated market values could theoretically reflect a concentration in overall economic activity. Indeed, under the assumptions that deliver the Capital Asset Pricing Model (CAPM), the value-weighted return on the market is proportional to the marginal value of wealth and it emerges as the single pricing factor. However, in practice, the value-weighted *US equity* market return is typically used as a proxy for this object despite the fact that publicly traded equities comprise a small fraction of most investors' total wealth. In such a context, stocks that move the value of the equity market but are disconnected from broader components of wealth do not command high risk premia in most models (see, e.g., Mayers, 1973; Roll, 1977).

In this paper, we identify a component of the standard CRSP value-weighted market factor (VMF) that is related to the relative overweighting of large-cap stocks in the index. We call this component the "idiosyncratic financial factor" (IFF). We show that the IFF carries no risk premium and is effectively unrelated to the broader economy or to returns in other asset classes. Moreover, we provide theoretical and empirical evidence on how the existence of the IFF can be used to explain two asset pricing "puzzles": 1) the size anomaly (along with the need to include size factors in standard multifactor models) and 2) the weak intertemporal relationship between the market's risk premium and its conditional variance.

Our paper makes four distinct contributions. First, we build a theoretical model featuring an IFF that is generated by a disconnect between the returns on financial and non-financial wealth. We use the model to develop predictions for asset pricing tests that use two distinct market proxies: 1) the VMF that is standard in the literature and 2) a factor that we call the

[1]For instance, in November 2018, *The Economist* magazine 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."

dominant market factor (DMF), which is constructed to explain as much common variation across groups of stocks as possible and prices the average stock. The model implies that the IFF is simply the difference between the VMF and the DMF. Second, we estimate empirical proxies for the DMF and IFF and show that these proxies are consistent with our model's main assumption that the IFF reflects a macro-financial disconnect. Third, we provide empirical evidence consistent with four predictions from our model related to attenuation in estimated market betas, the size anomaly and the role played by size factors, and IFF pricing. Fourth, we document a stronger intertemporal risk-return tradeoff when estimated using the DMF compared to when using the VMF. All of these results are consistent with the IFF adding unpriced risk into the VMF, which leads to mismeasurement in the amount of priced market risk and assets' exposures to that risk.

We begin by presenting two motivating facts related to the concentration of weights across firms in the VMF. First, a small number of large firms constitute a large share of the value-weighted market. Second, there are dramatic fluctuations in the industry composition across all firms in the equity market (particularly among the largest firms), and these fluctuations do not reflect the composition of the broader economy. This evidence suggests that some of the necessary conditions for the existence of an IFF are present in the equity market.

Motivated by this evidence, we build a static model featuring a macro-financial disconnect that arises when investors hold at least some non-equity market wealth. This disconnect generates an IFF, which is comprised of factor tilts specific to large firms and a "granular residual" (in the spirit of Gabaix, 2011), neither of which diversify away when value weights are used to construct the market factor. While the model nests the CAPM as a special case, in the empirically relevant case where investors finance most of their consumption from income sources other than equity market wealth, the model makes a sharp distinction between the DMF and the IFF. When the IFF exhibits little comovement with other sources of non-equity market wealth, as we show is the case in the data, its risk price is small and even zero in some limiting cases. Under plausible assumptions, our model generates the size anomaly when the VMF is used as the only pricing factor, which is resolved if the DMF is used or if a size factor is included.

We estimate the DMF as a projection of the VMF onto the first principal component (PC) from a principal component analysis (PCA) on a large set of standard equal-weighted stock portfolios. Using equal-weighted portfolios eliminates the outsized effect of large-cap stocks and ensures that any granular residuals diversify away. Our estimate of the IFF – the difference between the VMF and DMF – has an annualized standard deviation of approximately 6%, or about 40% of the VMF's standard deviation.[2] Our results are robust (and estimated IFF and DMF series are quite similar) to constructing the DMF using several different approaches that address different measurement concerns while still avoiding substantially overweighting a small number of large-cap stocks (see Section 4.2).

Our first empirical contribution is to validate a key assumption of our model, namely that there is a disconnect between the IFF and the broader economy. To this end, we estimate local projections of multiple variables that are linked to investor marginal utility in different classes of asset pricing models onto both the DMF and the IFF. These models include standard representative agent models, limited participation models, incomplete markets models, and intermediary asset pricing models. We also study local projections of direct proxies for other components of investor wealth such as bond and real estate returns. Across all variables, responses to the DMF are large, statistically significant, and have signs consistent with what we would expect from theory. By contrast, none of the associated variables respond to the IFF in a direction that would generate a positive risk premium based on theory, and coefficients are typically statistically insignificant (but are occasionally statistically significant with signs *opposite* those associated with the DMF). These results are consistent with our notion that the IFF is largely "idiosyncratic" with respect to the broader economy.

Our second empirical contribution is to test several of our model's predictions for the cross-section of expected returns. Our model implies that many stocks' exposures to priced market risk (i.e., market betas) are attenuated downwards when measured with respect to the VMF compared to the DMF. Accordingly, we find quantitatively large differences in

[2]In theory, the IFF could be driven by two components related to the overweighting of large-cap stocks in the VMF: 1) exposure to factors important for large stocks and 2) a granular residual, which is the weighted average of firm-specific shocks that do not diversify away. In practice, we find that the IFF is driven mostly by factor exposures.

estimated market betas with respect to the VMF and the DMF for standard test portfolios despite the high correlation between these market proxies. Consistent with our model, differences between VMF and DMF betas are usually negative, especially for smaller and high beta stocks. These differences in market betas are often large: The cross-sectional standard deviation of the *difference* between VMF and DMF betas is 0.13, which is roughly 2/3 the size of the cross-sectional standard deviation of VMF betas themselves (0.19).

Relatedly, our model offers a simple explanation for the size anomaly and the role of size factors (e.g., Banz, 1981; Fama and French, 1992, 1993). When the VMF contains unpriced risk that is mostly related to the returns on large-cap stocks (the IFF), it leads to an underestimation of most other stocks' exposures to the priced component of market risk (the DMF). As a result, small-cap stocks earn positive alphas with respect to the VMF. Size factors, both in the model and data, help to separate exposures to the priced (DMF) and unpriced (IFF) components of the VMF. This logic points to a new interpretation for the role of size factors: they correct for mismeasured market risk induced by the IFF. Empirically, we investigate model-implied connections between the IFF and size in two ways.

First, when we use the VMF in a number of standard factor models and exclude size factors from models that have them, alphas on all size-sorted portfolios below the top decile are positive and are increasing as we move across larger to smaller portfolios. This fact has been extensively documented and is part of the motivation behind including size factors in these models in the first place. However, when we replace the VMF with the DMF, alphas on size-sorted portfolios below the top decile are effectively eliminated as is their previously strong relationship with portfolio size.[3]

Second, we investigate our model's predictions for the role played by size factors themselves. When we drop size factors from factor models and use the DMF as the market factor, the pared-down models price size factors. By contrast, models that use the VMF do not price size factors. Thus, after replacing the VMF with the DMF, size factors no longer improve these models' ability to price assets (Barillas and Shanken, 2017). Relatedly, size factors do

[3]The DMF's ability to price size-sorted portfolios is not a mechanical artifact from giving micro-cap stocks a higher weight in the DMF, rather it is due to its lower weight on large-cap stocks. For example, we obtain similar results when we explicitly exclude micro- and large-cap stocks from our DMF construction.

not improve mean absolute pricing errors (MAPEs) across a large set of test assets when the DMF is used as the market proxy. By contrast, size factors substantially reduce these errors when the VMF is used instead. Across all the factor models we consider, we document a strong relationship between portfolios' loadings on size factors and the difference between loadings on the VMF and DMF, which illustrates how including these factors helps to correct for inconsistent measurement of priced market risk across portfolios.

When the stock market's share of wealth is relatively small compared to wealth held in other assets, our model implies that the IFF has a negative alpha with respect to the VMF and a zero alpha with respect to the DMF. We test this implication by running factor spanning tests on the IFF using a number of standard factor models, varying whether we use the VMF or the DMF as the market factor. In models that use the VMF, the IFF generates a negative alpha across all models and is typically statistically significant. By contrast, in models that use the DMF, we typically cannot reject the null that the IFF alpha is zero.

As our third empirical contribution, we conduct a test in the time series. When the IFF is not priced, it contributes to the VMF's conditional variance but not its risk premium, thus making the VMF risk-return tradeoff weaker and unstable over time. In line with this prediction, 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 but statistically insignificant (e.g., Nelson, 1991; Glosten et al., 1993; Koopman and Hol Uspensky, 2002). By contrast, when the analysis is repeated using the DMF, the risk-return tradeoff coefficient is two to four times larger and is statistically significant across a wide variety of specifications.

Contribution to the Literature: Although the IFF could, in principle, affect any value-weighted factor, we focus on the VMF due to its historical importance and ubiquity as a pricing factor. Thus, our paper contributes to the broad literature that addresses challenges associated with a mismeasured market risk factor that dates back to Roll's (1977) critique. We focus on insights beginning with Mayers (1973) that imply a particular type of mismeasurement in the VMF that can occur when investors have large components of wealth

outside of public equity markets. Our model explains why value-weighted portfolios may not be ideally suited to serve as pricing factors and we provide empirical evidence consistent with this logic.[4]

Historically, the VMF has been used as a proxy for investor wealth in the context of the CAPM with some notable studies implying a rejection of that model (e.g., Fama and MacBeth, 1973; Fama and French, 1992). Researchers have tried to address this failure by incorporating observable proxies for omitted components of wealth into the market portfolio including other asset classes (Stambaugh, 1982), other sources of wealth such as human capital (e.g., Mayers, 1973; Campbell, 1996; Jagannathan and Wang, 1996; Heaton and Lucas, 2000; Campbell et al., 2016), or aggregate financial and non-financial holdings of households (Cederburg and O'Doherty, 2019). Our approach is more in line with that in Daniel et al. (2020) in that we attempt to remove an unpriced component of the VMF.[5]

While there is considerable disagreement about which pricing factors matter, the vast majority of workhorse models include at least two: a value-weighted market factor and a size factor constructed by buying small stocks and selling large ones. While there is little controversy about including the market factor, there are still active debates regarding what economic mechanism generates the size effect and whether the size factors earn a positive risk premia at all.[6] However, the size anomaly may simply capture any component of discount rates that are not captured by a particular misspecified model (Berk, 1995). We argue that the standard practice of using the value-weighted equity market return as a pricing

[4]Malevergne et al. (2009) recognize similar issues with the VMF and define a second pricing factor (the "Zipf factor") to be the difference between the VMF and the equal-weighted market portfolio. Rather than arguing that this factor is an unpriced component of the VMF, they include the Zipf factor as a second pricing factor and provide evidence that it has a non-zero risk premium. This is in stark contrast to our finding that the IFF has no risk premium and is unrelated to proxies for investor marginal utility.

[5]As opposed to studying a general disconnect between characteristics-sorted portfolios and asset covariance structures as in Daniel et al. (2020), we focus on a very specific diversification issue related to the standard VMF. Our adjustments are distinct both conceptually and empirically: We find that our DMF (IFF) has a correlation with the Daniel et al. (2020)-corrected market factor of approximately 41% (48%) (i.e., their corrected market factor is far from perfectly correlated with the DMF and it loads on the IFF).

[6]See van Dijk (2011) and Alquist et al. (2018) for summaries of this debate. More recently, Asness et al. (2018) and Hou and van Dijk (2019) provide evidence that the size effect can be resurrected if one controls for firm quality or profitability shocks, respectively. Furthermore, even if the size anomaly is small, it can still have a large effect on allocational efficiency in the real economy (van Binsbergen and Opp, 2019). Size factors are also appealing on more pragmatic grounds, as they substantially improve models' pricing performance, as our result show.

factor provides a strong rationale for the inclusion of a size factor or, alternatively, for using a different market benchmark (i.e., the DMF). We show that size portfolios are no longer mispriced when we replace the VMF with the DMF, suggesting that the key source of misspecification addressed by size factors is the attenuation in market risk exposures due to the IFF. Further, as our results also illustrate, size factors play an important role in reducing cross-sectional pricing errors even during the post-1993 period in which the size factor premium is small. This result is consistent with our model, which implies that size factors help to isolate covariance with the value-weighted market factor related to our (priced) dominant market factor from that related to the (unpriced) idiosyncratic financial factor, even if the size premium is small or zero.

We also contribute to the literature that studies the intertemporal risk-return tradeoff. Many studies have focused on various approaches to better estimate conditional variance, and it has been found that results are sensitive to these modeling choices.[7] 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 VMF variance. Accordingly, we find that adjusting the VMF to purge it of the IFF yields a stronger risk-return tradeoff regardless of how we measure conditional variance.

2 Motivating evidence

There are at least two possible causes of a disconnect between the VMF and the broader economy. First, idiosyncratic shocks to a small number of large-cap firms may not diversify and thus may have a non-trivial influence on the VMF. Second, the composition of the VMF may differ from that of the broader economy. In this section, we present two related motivating facts about the US stock market that could generate an IFF.

[7] 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. Using alternative approaches to GARCH 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.

2.1 Fact 1: The VMF is dominated by a few large firms

The popular news media is rife with stories highlighting concerns that recent stock market performance is dominated by a few mega-cap companies, including the "Magnificent Seven,"[8] but this concern is not unique to the present. For instance, in the late 1960s and early 1970s, the "Nifty Fifty" dominated the market. When a small number of large firms dominate the market, it is possible that their idiosyncratic shocks do not diversify and contribute to the aggregate market return via a granular residual (e.g., Gabaix, 2011). In this subsection, we explore whether conditions exist such that a granular residual might affect the VMF.

Figure 1 plots the fraction of market value associated with the largest 5, 10, 25, and 50 stocks in the CRSP universe. Throughout the period from 1926 to 2021, the largest five stocks are associated with between 10 and 25 percent of overall market value. The combined weight of the top 25 stocks ranges from a low of approximately 20% during the 1980s to a high of almost 50% in the 1930s. As of 2021, this weight is approximately 35%.

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. Nevertheless, 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 possible that the VMF is not well diversified with respect to the idiosyncratic components of mega-cap stock returns. If this aggregated granular residual is large and carries a different price of risk than the DMF, it could cause issues for standard asset pricing inference when using the VMF. Below, we develop motivation for why one might expect the granular residual to be priced differently than the DMF in a theoretical model. Although such a granular residual could, in principle, affect asset pricing inference, we provide evidence that it is actually small in our setting and is unlikely to be the main driver of the disconnect between the DMF and VMF.

[8] See, for instance, the Wall Street Journal article "It's the Magnificent Seven's Market. The Other Stocks Are Just Living in It" from December 17, 2023.

2.2 Fact 2: VMF industry composition differs from that of the broader economy and varies dramatically over time

Even if the granular residual induced by mega-cap firms is small, large firms could generate a disconnect between the VMF and the broader economy if they are not fully representative of the broader economy. We explore one dimension over which such a disconnect could exist: industry composition. We show that the VMF's industry composition varies dramatically over time and is not representative of the real economy's industry composition.

Figure 2 plots the share of the US public equity value among all publicly traded stocks (Panel 1), public equity value among the top 2.5% of stocks by market capitalization using a NYSE breakpoint (Panel 2), and employment (Panel 3) for each of the 10 Fama-French industries over time. The composition of the total US equity market varies dramatically over time. For instance, in 1963, the manufacturing (high-tech) sector accounted for 25% (7%) of the equity market compared to 8% (36%) in 2021. Panel 2 shows that an important source of instability is the shifting industry composition of the largest publicly traded firms in the top 2.5% of stocks. In this subsample, shifting industry composition is even more dramatic, with the high-tech industry comprising approximately 73% of the value of these stocks in 2021. This dramatic fluctuation in the composition of the top 2.5% of stocks is important to recognize, since it can be seen that short term variation in Panel 1 among the broader equity market is often driven by these few large firms.

Are these compositional dynamics representative of the broader economy? To help gain insight into this question, we look to industry-level employment dynamics over the same time period. Panel 3 indicates that employment shares are much more stable across industries over time. For instance, in 1978, the manufacturing (high-tech) sector accounted for 17% (3%) of employment compared to 6% (5%) in 2021. So, the share of employment in high-tech industry composition was relatively stable over this time period, whereas that of the manufacturing industry underwent a percent decline similar to that indicated by the total equity market value.

The results in Panels 1 and 3 also highlight broad and persistent compositional differences between the industry shares among public equities and the broader economy based on the

industry shares in employment. For instance, the high-tech and utilities industries typically comprise a larger share of public equities than they do of employment because they demand more funding through the issuance of public equity. By contrast, healthcare comprises a relatively large share of employment, potentially because many workers in that industry work for privately owned companies such as hospitals and doctor's offices.

This evidence implies a disconnect between the industry composition of the equity market and that of the real economy. Although there are many dimensions other than industry composition that could also contribute to such a disconnect, we do not (nor do we need to) take a stand on the exact sources of this disconnect when constructing our theoretical model and testing its predictions.

3 Theoretical framework

3.1 A model in which the IFF is priced differently than the DMF

In this section, we outline a static one-period theoretical model that explicitly incorporates a macro-financial disconnect in the economy. We relegate additional details to Online Appendix OA.2. Our model features agents who hold both tradable financial wealth and an additional source of nontradable wealth. While the latter component captures many potential additional resources to fund consumption, we will refer to it as "labor income" to fix ideas. An important feature in this model is that there is a disconnect between returns on financial wealth (the VMF analog in the model) and those on labor income.

Motivated by the empirical evidence on the size distribution of stocks in Figure 1, investors can trade two types of stocks in the financial sector: K large stocks and $N - K$ small stocks, where $N \gg K$. We focus on the limit in which K is fixed and $N \to \infty$. Thus, idiosyncratic shocks to small stocks are perfectly diversified in the VMF, but idiosyncratic shocks to large stocks are not. Our interpretation of "small stocks" in the model is fairly broad. They are defined to have weights that are sufficiently small so that their idiosyncratic shocks are diversified in the VMF, and such that the average small-cap stock does not load

on factors that are particularly relevant for the largest firms in the economy.

We assume that realized returns in the financial sector are governed by two orthogonal factors, f and g, with the excess return on stock i given by

$$r_i = \beta_i f + \zeta_i g + \eta_i \tag{1}$$

where β_i and ζ_i are heterogeneous across stocks and η_i reflects a mean-zero idiosyncratic shock that is orthogonal to f and g. We define f to be a level factor normalized such that the value-weighted financial market return has a loading of one on f (i.e., $\sum_{i=1}^{N} w_i \beta_i = 1$ where w_i reflect value weights). g is an additional systematic factor rotated such that the average small stock has a zero loading on it (i.e., $\frac{1}{N-K} \sum_{i=K+1}^{N} \zeta_i = 0$), but the average large stock has a non-zero loading on it. f is meant to reflect the DMF, and g is meant to reflect tilts away from the DMF.

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

$$\mathbb{E}[r_i] = RRA \cdot [(1-\phi)Cov(r_i, r_m)) + \phi Cov(r_i, r_h)] \tag{2}$$

involving two priced factors, where RRA is relative risk aversion and ϕ captures the expected share of future consumption coming from labor income, which we will refer to as the human wealth share. The first priced factor is the VMF return, r_m, which is given by

$$r_m = f + \zeta_m g + \eta \equiv f + \epsilon, \tag{3}$$

where ζ_m captures the market's exposure to g, and η is the granular residual coming solely from the aggregation of η_i across large stocks (since the law of large numbers holds for the aggregation of η_i across small stocks). r_m reflects the model analog for the VMF, which is the sum of the DMF (f) and the IFF (ϵ). The second priced factor, r_h, captures the innovation in labor income and is given by

$$r_h = \psi_f (f - \mathbb{E}[f]) + \psi_\xi \xi \tag{4}$$

where the ψs are model constants (with $\psi_f \geq 0$), and ξ is an aggregate labor return shock that is assumed to be uncorrelated with the financial sector. We assume that labor does not load on ϵ, which is consistent with the empirical evidence we present in Section 5 that the IFF does not have a strong contemporaneous correlation with nor does it predict future labor income growth (or other macroeconomic variables, for that matter).

Equation (2) implies that our model nests the standard CAPM when there are no non-traded assets ($\phi = 0$) . In this case, the stochastic discount factor is linear in r_m, and g or η shocks are priced identically to f shocks. However, when nontraded 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 r_m becomes very weak when the nontraded assets do not load on ϵ. In the empirically-relevant case with $\phi \neq 0$, the model yields the following four testable predictions for which we find empirical support:

1. **Attenuation of small-cap betas:** Regardless of the value of ϕ, the true beta of a portfolio of small-cap stocks with respect to f, $\beta_S \equiv \frac{1}{N-K}\sum_{i=K+1}^{N}\beta_i$, is larger than its beta with respect to r_m as long as $\beta_S > 0$.

2. **Small-cap alphas:** Regardless of the value of ϕ, the alpha of a portfolio of small-cap stocks is positive with respect to r_m and zero with respect to f.

3. **Size factor and spanning:** Let $r_{smb} \equiv \frac{1}{N-K}\sum_{i=K+1}^{N} r_i - \frac{1}{K}\sum_{i=1}^{K} r_i$ be the return on a portfolio that is long small-cap stocks and short large-cap stocks, and suppose r_{smb} and r_m are linearly independent. For arbitrary ϕ, expected returns are linear in the bivariate betas on (r_m, r_{smb}) or (f, r_{smb}). In the limit as $\phi \to 1$, r_{smb} is spanned by f and unnecessary to include when f is used as the market factor; by contrast, r_{smb} is required as a pricing factor if r_m is used as the market factor. Furthermore, if $\beta_S > 1$ and $\phi \to 1$, $E[r_{smb}] \geq 0$.

4. **IFF alphas:** In the limit as $\phi \to 1$, the IFF has a risk premium of zero (i.e., an average excess return of zero); in addition, its alpha is negative with respect to a factor model that uses r_m and zero with respect to one that uses f.

The intuition behind Prediction 1 is that r_m contains noise relative to f (i.e., terms related to g and η in equation (3)), on which small-cap stocks do not load on average, leading to attenuation bias when r_m is used to estimate small-cap betas. Prediction 2 follows from this small-cap beta attenuation bias.

The traditional rationale for including size factors is that small stocks may be exposed to extra systematic risks (e.g., those related to financial constraints or common liquidity shocks) that are irrelevant for large stocks. Prediction 3 points to a very different interpretation: Large stocks – which also have substantial loadings on the DMF – are exposed to sources of common variation (ϵ) that are unlikely to correlate with investors' marginal utility. The result follows from the fact that f and ϵ can be recovered from a linear combination of r_m and r_{smb}. As such, including the size factor allows the model to separately estimate priced exposures to f and ϵ. However, when ϵ is not priced, size factors are unnecessary when f is used as the market proxy but remain necessary when r_m is used. Relatedly, Prediction 4 follows from the fact that ϵ's beta with respect to r_m is biased upwards relative to its beta with respect to f (because there are terms related to ϵ in r_m but not in f).

As we discuss further in Online Appendix OA.4, 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 it (i.e., it is likely empirically the case that ϕ is close to 1). Accordingly, we find empirical support for each of these predictions.

3.2 Linear asset pricing framework

In this section, we develop a simple linear asset pricing framework to operationalize and test predictions from our theoretical model in our empirical work. We extend the model framework to incorporate multiple additional priced risk factors and time-varying volatility, where the latter feature allows us to speak to the intertemporal risk-return relationship.

3.2.1 Maintained assumptions

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,

$$VMF_{t+1} = \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}, \tag{5}$$

where $w_{i,t}$ is the value-weight of stock i based on its market capitalization, and N_t is the number of stocks in the index at the end of period t.

We assume that each stock's excess return is described by a linear factor model such that

$$r_{i,t+1} = \beta_i \cdot f_{t+1} + \zeta_i' \cdot g_{t+1} + \eta_{i,t+1}, \tag{6}$$

where f_{t+1} is a scalar random variable that represents the DMF, β_i is stock i's exposure to f_{t+1}, g_{t+1} is a vector of additional common factors with $g_{t+1} \perp f_{t+1}$, ζ_i is stock i's exposure to g_{t+1}, and $\eta_{i,t+1}$ is a finite-variance idiosyncratic shock with $\mathbb{E}[\eta_{i,t+1}|f_{t+1}, g_{t+1}] = 0$.

3.2.2 Defining the DMF and IFF

We separate f_{t+1} from the other factors g_{t+1} via the following restrictions:

$$\mathbb{E}\left[\sum_{i=1}^{N_t} w_{i,t}\beta_i\right] = 1, \text{ and} \tag{7}$$

$$\mathbb{E}\left[\frac{1}{N_t}\sum_{i=1}^{N_t} \zeta_i\right] = 0. \tag{8}$$

In other words, the scale of f_{t+1} is defined such that the VMF has an unconditional average loading of one on it, and the remaining factors (g_{t+1}) are rotated so that the average stock has zero loading on them. Therefore, f_{t+1} captures the common factor exposure of the average stock, and g_{t+1} capture tilts away from f_{t+1}. This is similar in spirit to standard factor models that use the VMF as a level factor, with other factors constructed from long-short portfolios to capture tilts away from the VMF.

Combining equations (5) and (6), the VMF can be expressed as

$$\begin{aligned} VMF_{t+1} &= \beta_{m,t} \cdot f_{t+1} + \zeta'_{m,t}\ g_{t+1} + \sum_{i=1}^{N_t} w_{it}\eta_{i,t+1} \\ &\equiv f_{t+1} + \underbrace{(\beta_{m,t}-1)\cdot f_{t+1}}_{\substack{\text{time-varying exposure} \\ \text{to true market factor } f_{t+1}}} + \underbrace{\zeta'_{m,t}\ g_{t+1}}_{\substack{\text{(potentially) time-varying} \\ \text{exposure to other factors } g_{t+1}}} + \underbrace{\eta_{t+1}}_{\substack{\text{granular} \\ \text{residual}}} \\ &\equiv \underbrace{f_{t+1}}_{\substack{\text{dominant} \\ \text{market factor}}} + \underbrace{\epsilon_{t+1}}_{\substack{\text{idiosyncratic} \\ \text{financial factor}}} . \end{aligned} \tag{9}$$

where $\beta_{m,t} \equiv \sum_{i=1}^{N_t} w_{it}\beta_i$ and $\zeta_{m,t} \equiv \sum_{i=1}^{N_t} w_{it}\zeta_i$. Although the assumed stock-level loadings in equation (6) are time invariant, the VMF loadings are in general time varying due to the changing composition of stocks that comprise the equity market. We define $\epsilon_{t+1} \equiv VMF_{t+1} - f_{t+1}$ as the IFF, and it is is the sum of three distinct components. The first component captures the fact that the VMF may have conditional loadings on f_{t+1} that are above or below the time series average loading of one. The second captures the value-weighted average of conditional loadings on other orthogonal factors g_{t+1} (a deviation relative 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 (e.g., Figure 2), these factor exposures likely vary over time.

Following Gabaix (2011), we 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 the conditional variance of η_{t+1} collapses to zero depends on the extent to which the law of large numbers applies.

3.2.3 Risk premia on the DMF and VMF in the linear framework

Assuming there are no arbitrage opportunities, there exists 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. The linear factor structure in equation (6) implies that the SDF takes the form

$$M_{t+1} - \mathbb{E}_t[M_{t+1}] = -\gamma \cdot (f_{t+1} - \mathbb{E}_t[f_{t+1}])\mathbb{E}_t[M_{t+1}] - \Lambda(g_{t+1} - \mathbb{E}_t g_{t+1})\mathbb{E}_t[M_{t+1}]. \tag{10}$$

In addition to the standard assumption of constant risk prices (γ and Λ), the other key assumption implied by equation (10) is that the granular residual, η_{t+1}, is not priced, which holds in our theoretical model when $\phi \to 1$. This does not imply that the IFF is not priced but, rather, that just one component of the IFF is not priced. We make this assumption to simplify the exposition, but it does not affect the interpretation of our empirical results.

Given the SDF in equation (10), asset i's risk premium is given by

$$\mathbb{E}_t\left[r_{i,t+1}\right] = \gamma \cdot \beta_i \cdot \sigma^2_{f,t} + \Lambda' \Sigma'_{g,t} \Sigma_{g,t} \zeta_i, \tag{11}$$

where $\sigma^2_{f,t}$ and $\Sigma'_{g,t}\Sigma_{g,t}$ capture the conditional variance of f and the covariance matrix of the g factors, respectively. The conditional risk premia on the DMF and VMF are thus given by

$$\mathbb{E}_t[DMF_{t+1}] = \mathbb{E}_t[f_{t+1}] = \gamma \cdot \sigma^2_{f,t} \text{ and} \tag{12}$$

$$\mathbb{E}_t\left[VMF_{t+1}\right] = \gamma \cdot \sigma^2_{f,t} + \gamma \cdot (\beta_{m,t} - 1) \cdot \sigma^2_{f,t} + \Lambda' \Sigma'_{g,t} \Sigma_{g,t} \zeta_{m,t}. \tag{13}$$

4 The empirical DMF and IFF

4.1 Estimating the DMF and IFF

We have defined the the theoretical DMF (f_{t+1}) in our linear asset pricing framework in equations (6)–(8) and distinguished it from the VMF in equation (9). Our goal here is to identify an empirical proxy for f_{t+1} that serves as a level pricing factor with as little exposure as possible to the IFF defined in equation (9) as possible. To this end, we perform a PCA on the excess returns of a broad set of equal-weighted characteristics-sorted portfolios so that each PC is a linear combination of excess returns on the portfolios.[9]

We use a standard set of 372 portfolios with data from 1963-2021. We obtain all portfolio

[9]Early studies that make use of PCA 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). We use PCA as a tool to identify a level factor (the first PC) that helps define our DMF and addresses our diversification concerns about the VMF, but we do not make use of higher-order PCs in our analysis.

data from Kenneth French's Data Library and provide a summary of the portfolios in Online Appendix Table OA.1.[10] We choose these portfolios for our PCA to balance a few tradeoffs that we discuss further in Online Appendix OA.5; we also discuss some additional details related to the PCA in that section. Most importantly, using equal-weighted portfolios helps alleviate the concern that dramatic time-variation in the composition of large-cap firms (e.g., such as in Figure 2) drives variation in our DMF. Using equal-weighted portfolios also ensures that any granular residuals are diversified out of the DMF.[11]

The first PC can be thought of as a scaled empirical proxy for the DMF. In order to ensure that our empirical DMF satisfies our scale condition in equation (7), we define the DMF to be a full-sample projection of the VMF onto the first PC. The IFF is then defined as the difference between the two (i.e., $IFF \equiv VMF - DMF$), consistent with its definition in equation (9). Higher PCs can be thought of as proxies for g_{t+1}, because they are orthogonal to the first PC and explain variation in asset returns that is unexplained by the first PC.

In addition to mechanically satisfying our scale restriction on the DMF, we find empirical support that the PCA framework satisfies other key assumptions about f_{t+1} and g_{t+1}. First, consistent with our interpretation of the DMF as a level factor, we find that the portfolio weights used to construct the first PC are all positive and have a relatively tight distribution. Consistent with our restriction in equation (8), we also find that the average loading of the 372 portfolios (both equal- and value-weighted versions) on each of the higher PCs is approximately zero. See Online Appendix OA.5.1 and Figures OA.1–OA.2 for further discussion and related results.

4.2 Alternative DMF specifications considered for robustness

The main DMF specification addresses our VMF diversification concerns, relies on readily accessible portfolio data, and is straightforward to construct. However, there are some potential shortcomings in relying on equal-weighted portfolios. In this section, we describe

[10] In some robustness checks, we use portfolios with weights that are not provided in Kenneth French's Data Library. In these cases, we construct the portfolios on our own using data from Compustat and CRSP.

[11] This is true as long as the equal-weighted portfolios are well diversified such that the law of large numbers approximately holds. In this case, idiosyncratic shocks, $\eta_{i,t+1}$, will diversify away such that they do not impact our DMF.

four alternative DMF constructions that address these potential concerns.

First, equal-weighted portfolios are costly to trade and are exposed to return reversals. To address these potential concerns, we construct an alternative DMF that uses capped value-weighted versions of the same 372 portfolios, where we weight stocks by their market equity winsorized at the eightieth NYSE percentile as in Jensen et al. (2023). By limiting the weights of large-cap stocks, this construction also mitigates our concerns about the VMF being unduly influenced by large-cap stocks.

Second, our size anomaly results might be "mechanically" related to overweighting small-cap stocks in our main DMF. To address this potential concern and also limit the effects of large-cap stocks, we construct an alternative DMF that uses equal-weighted versions of the same 372 portfolios where we drop stocks in the lowest and highest 20th percentiles of the size distribution (by NYSE breakpoints) before constructing the portfolios. The small-cap filter eliminates a significant fraction of stock-month observations (55%) from our sample; these are the stocks that populate the lowest deciles of size-sorted portfolios and are the ones whose average returns we aim to explain with the DMF in our size anomaly analyses. By contrast, the large-cap filter only eliminates 8% of stock-month observations form our sample due to the skewed size distribution.

Third, our PCA uses a full-sample analysis to identify the DMF. Thus, the DMF is not tradable. To address this potential concern, we also construct a tradable version of the DMF using a rolling PCA to identify ex ante DMF weights on each of the 372 portfolios.

Finally, assuming the market composition of stocks does not change, we show that the equal-weighted market is a scaled version of the DMF in Online Appendix OA.6. Thus, we consider using the VMF projected onto the CRSP equal-weighted market as a fourth alternative construction of the DMF. Despite the restrictive assumption needed to generate this implication, we provide robustness checks using this proxy for the DMF because its construction is straightforward and does not rely on any statistical factor identification as in our PCA analyses.

None of these modifications materially alter our results or conclusions. Thus, for brevity, all results in the main text are based on our main DMF specification unless otherwise noted.

We provide robustness checks using each alternative DMF specification in the Online Appendix and refer to these results in footnotes when introducing each of our main results.

4.3 DMF and IFF summary statistics

Table 1 reports summary statistics on the monthly excess return time series for the VMF, the DMF, and the IFF. We include results based on our main DMF (constructed using the PCA based on equal-weighted portfolios) and on the four alternative constructions discussed in the previous section. Panel A presents statistics related to the DMF, and Panel B presents statistics related to the IFF. Average returns on the VMF and various DMF constructions are similar at approximately 7% per year. The DMF has a lower volatility than the VMF, is less negatively skewed, and has higher kurtosis.

If the IFF captures an important source of priced risk, one would expect it to earn positive excess returns on average (given that it is a tradable portfolio). We find little empirical support for this in the data. The average return on the IFF is economically small at -0.53% per year, and, consistent with Prediction 4 of our model, we cannot reject the null hypothesis that it is equal to zero. Relatedly, we show that the VMF is priced by the DMF in Online Appendix OA.7, with the DMF explaining 84% of the variation in the VMF. Furthermore, we find that the VMF's loading on higher PCs either does not add significant explanatory power to its risk premium, and in some cases reduces explanatory power (e.g., in some cases the VMF has negative loadings on PCs with positive risk premia). These results are consistent with our hypothesis that the main source of risk in the VMF for which investors demand a positive risk premium is captured by the DMF.

The ratio of IFF-to-VMF variance is 15.7%, implying that the fraction of total variability in the VMF induced by the IFF is moderate compared to that of the DMF. Thus, it has sufficient scope to affect asset pricing tests when using the VMF as a market factor. We provide an analysis in Online Appendix OA.9 that indicates that the ratio of the granular residual's variance to that of the VMF is at most approximately 1-2%, implying that this component of the IFF is unlikely to be the primary driver of any differences in asset pricing inferences between the VMF and the DMF.

We provide correlation matrices for the DMF and IFF variants constructed via different methods in Panels A and B, respectively. All DMF variants are highly correlated with our main DMF specification. This, paired with their similar average returns, explains their similar performance in robustness checks on our main results and suggests that the DMF is an easily identifiable level factor regardless of whether the base portfolios in the PCA use equal weights, use capped value weights, or drop small- and large-cap stocks.

5 The IFF and the macro-financial disconnect

In this section, we test a key assumption in our theoretical model, namely that the DMF is strongly linked to broader economic conditions (i.e., proxies for changes in the value of wealth outside of investors' US tradable stock portfolio) whereas the IFF is not. To this end, we examine whether the DMF and IFF have differential predictive power for macroeconomic aggregates and performance among other asset classes by estimating local projection impulse responses (LPIRs) (Jordà, 2005).

We estimate the LPIRs for fifteen different empirical proxies for investor marginal utility in several different classes of asset pricing models, some in the spirit of omitted sources of investor wealth. Our choice of macroeconomic aggregates and other time series are log per capita consumption growth (FRED), log GDP growth (FRED), total US income growth (World Inequality Database, https://wid.world/data/), real uncertainty (Jurado et al., 2015), top 0.1% US income growth (World Inequality Database, https://wid.world/data/), non-farm proprietary income growth (FRED) (Heaton and Lucas, 2000), the income risk measure based on cross-sectional skewness in household income growth (Schmidt, 2024), the unemployment rate (FRED), the common idiosyncratic volatility (CIV) factor (Herskovic et al., 2016), the intermediary capital risk factor (He et al., 2017), the common mutual fund flow factor (Dou et al., 2024), the credit spread (Moody's seasoned Baa corporate bond yield relative to the yield on 10-year treasury), the Bloomberg Barclays US Aggregate Bond Index, the Hedge Fund Research fund-of-funds diversified hedge fund return index (HFRIFOFD), and the FTSE NAREIT US Real Estate Index.

Figure 3 plots LPIRs with respect to either the DMF or the IFF.[12] The blue lines and confidence bands correspond to using the DMF as a forecasting variable, whereas the red lines and confidence bands correspond to using the IFF. Variables are signed such that increases in variable values correspond with decreases in marginal utility (i.e., we would expect these variables to have positive LPIR coefficients with respect to the DMF). Additional details on the LPIR estimation can be found in Online Appendix OA.10.

Representative agent models: Panels 1–4 show LPIRs for some common measures of aggregate risk that appear in common representative agent-based models with complete markets. For instance, Panel 1 shows that consumption growth (see, e.g., Lucas, 1978, and the many papers that followed) responds strongly and positively to the DMF but has no response to the IFF. This is also true of related aggregate consumption risk measures including GDP growth, total income growth, and (the negative of) real uncertainty (see, e.g., Bansal et al., 2014, for a discussion of the pricing of uncertainty). Note that the mean-reverting response of real uncertainty is expected since this is a stationary time series. This is also true of LPIRs for all other stationary time series we investigate.

Limited participation models: Given that a high fraction of aggregate stock market wealth is owned by a small subset of high-income/high-net worth investors, we also consider measures that are more relevant in models that emphasize the concentration of aggregate risks onto market participants (see, e.g., Basak and Cuoco, 1998). Panels 5–6 show LPIRs for some common measures of aggregate risk that appear in limited participation models such as the income growth of top earners (see, e.g., Malloy et al., 2009; Parker and Vissing-Jørgensen, 2009; Campbell et al., 2016) and proprietor's income as in Heaton and Lucas (2000). In both cases, the DMF generates a strong and positive response in the variables (albeit with some reversion in the latter case), whereas the IFF generates no response.

Incomplete markets models: In the presence of imperfect risk sharing and state-dependent distributions of idiosyncratic shocks to agents' consumption and wealth, higher moments of shocks that redistribute consumption across different agents enter the stochastic

[12]We report robustness checks on the results in Figure 3 in Online Appendix Figures OA.4, OA.5, OA.6, and OA.7 using the alternative DMF/IFF constructions described in Section 4.2.

discount factor (see, e.g., Panageas et al., 2020, for a survey). It is straightforward to extend our model to incorporate such features (see Online Appendix OA.2), which creates an additional rationale for the DMF carrying a higher risk premium than the IFF when the DMF is informative about idiosyncratic risk but the IFF is not. Panels 7–9 show LPIRs for some common measures related to undiversifiable risk in incomplete markets models including the skewness measure from Schmidt (2024), unemployment, and the common idiosyncratic volatility (CIV) measure from Herskovic et al. (2016). In all cases, the DMF generates a strong and positive response, whereas there is no response (or a negative response in the case of CIV) to the IFF. The positive coefficients on the DMF are expected in a model in which good news about the DMF signals lower idiosyncratic risk (higher skewness and lower volatility). Although these models are silent on IFF sign predictions, the zero or negative coefficients we estimate show the IFF has a relationship with these idiosyncratic risk measures distinct from that of the DMF.

Intermediary risk models: Panels 10–11 show LPIRs when using tradable factors associated with intermediary risk, including the intermediary risk factor from He et al. (2017) and the common fund flows factor from Dou et al. (2024). In both cases, the DMF generates a strong and positive response in the variables (consistent with negative DMF realizations being bad states for intermediaries). In contrast, the IFF fails to generate a response.

Other asset classes: According to Roll's (1977) critique, one reason the VMF may lead to a rejection of the CAPM is that it does not contain other sources of investors' wealth held outside the US equity market. These omitted sources of wealth can drive a wedge between the VMF and the theoretical ideal market index, even in a world in which CAPM assumptions hold. It is therefore informative to understand the comovement of both the DMF and the IFF with other sources of financial wealth, several key components of which are included in Panels 12–15. First, Panel 12 considers an important component of the return on risky debt, showing that (the negative of) credit spreads respond strongly and positively to the DMF at business frequencies but do not respond to the IFF. Panel 13 shows that the (negative of) bond returns are also positively related to the DMF, but not related to the IFF. Bond prices can be viewed as a rough proxy for expected marginal utility, such

that declines in bond prices reflect good news, consistent with the signs of the responses we observe.[13] Finally, the returns among other asset classes that reflect other components of wealth (such as hedge funds and REITs in Panels 14 and 15, respectively) respond strongly and positively to the DMF, whereas hedge funds show no response to the IFF, and REITs show a long-term negative response.

The evidence in this section therefore illustrates that, regardless of one's preferred micro-foundation for risk premia, a stark contrast emerges between the DMF and the IFF. While the DMF is tightly linked with variables that capture investors' marginal value of wealth, such a link is more tenuous (and sometimes incorrectly signed) for the IFF.

6 Testing model predictions

6.1 The IFF and market betas

In this section, we provide evidence consistent with Prediction 1 from our theoretical model that small-cap VMF betas are biased downward compared to DMF betas. We begin by characterizing beta bias in the presence of the IFF in a more empirically relevant setting than that described in our theoretical model, which yields some additional insights. Given our assumed factor structure in equation (6), the excess returns on a portfolio are

$$r_{p,t+1} = \alpha_p + \beta_p \cdot f_{t+1} + \zeta_p' \cdot g_{t+1} + \eta_{p,t+1}, \tag{14}$$

where $r_{p,t+1}$ is the excess portfolio return, and $\eta_{p,t+1}$ represents the portfolio's (potential) granular residual.[14] If one estimates the market risk exposure $\widehat{\beta}_p$ as the slope coefficient from a standard CAPM regression of $r_{p,t+1}$ on VMF_{t+1} and an intercept, the OLS estimate

[13] Consistent with this interpretation, Lustig et al. (2013) show that future bond returns are positively correlated with their measure of future returns on wealth (i.e., lower bond prices today forecast higher bond returns and wealth returns).

[14] We assume that the portfolio's composition is relatively stable so that the factor loadings are time invariant. This assumption simplifies the intuition for the beta bias we develop here and is consistent with standard empirical practice when estimating portfolio betas.

of β_p in equation (14) converges to

$$\underset{T\to\infty}{\text{plim}}\widehat{\beta}_p^{VMF} = \beta_p \cdot \underbrace{\left\{\frac{Var\left[f_{t+1}\right]}{Var\left[f_{t+1}+\epsilon_{t+1}\right]}\right\}}_{\text{attenuation bias}} + \underbrace{\frac{Cov\left[\epsilon_{t+1}, r_{p,t+1}\right]}{Var\left[f_{t+1}+\epsilon_{t+1}\right]}}_{\text{additive bias}}. \tag{15}$$

The IFF generates two sources of bias in $\widehat{\beta}_p^{VMF}$: 1) An attenuation bias represented by the term in brackets and 2) An additive bias represented by the direct correlation between the IFF (ϵ_{t+1}) and the the portfolio return $r_{p,t+1}$, which is mostly driven by $\zeta_p' g_{t+1} + \eta_{p,t+1}$. In practice, we expect the additive bias to be larger for large-cap portfolios, because the covariance between $r_{p,t+1}$ and ϵ_{t+1} is likely higher in this case (recall that this correlation is zero for the average small stock). Regarding the attenuation bias, the term in brackets is less than 1, implying that $\widehat{\beta}_p^{VMF}$ is likely attenuated downwards for most individual stocks (since the second term is likely small for all but very large stocks). However, when using value-weighted portfolios as test assets, the direction of the bias is ambiguous. We expect downward bias for small-cap portfolios, but we expect this downward attenuation bias to be partially nullified or even made positive by the additive bias for large-cap portfolios. Finally, equation (15) also implies that, conditional on the additive bias, the betas of high-beta portfolios are likely more severely biased than those of low-beta portfolios due to the attenuation bias. In practice, we use the DMF as a proxy for f_{t+1} and simply refer to the difference $\beta_p^{VMF} - \beta_p^{DMF}$ as the "beta gap" rather than calling it a "bias."

Equation (15) implies that VMF betas are, on the whole, likely to have downward bias relative to DMF betas. We begin by confirming this implication in Figure 4, which plots kernel density estimates of β_p^{VMF} and β_p^{DMF} estimated using univariate regressions for each of the value-weighted versions of the 372 portfolios used in our PCA.[15] The red (blue) curve represents the β_p^{VMF} (β_p^{DMF}) density. The table beneath the figure shows percentiles of the beta distributions, as well as some summary statistics for the beta gap distribution. The β_p^{DMF} density has fatter tails, especially on the right, and is shifted rightwards relative to the β_p^{VMF} density. The β_p^{DMF} percentiles are mostly larger than the β_p^{VMF} percentiles, with

[15]We report robustness checks on the results in Figure 4 in Online Appendix Figure OA.8 using the alternative DMF constructions described in Section 4.2.

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 (0.13) and the standard deviation of β_p^{DMF} (0.28) are large relative to the standard deviation of β_p^{VMF}.

Consistent with the intuition from our theoretical model, equation (15) also implies that the beta gap is more severe for small-cap portfolios. Figure 5 Panel 1 plots the beta gap for each of the 372 portfolios against each portfolio's log-size. The OLS best-fit line is also plotted in red for ease of visualization. Consistent with the intuition from equation (15), 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.

Figure 5 Panel 2 plots β_p^{VMF} versus β_p^{DMF}, showing that the beta bias is typically more severe for high-beta stocks. Intriguingly, this result is similar to the finding in Decaire and Graham (2024) that VMF betas are typically smaller than the CAPM market betas used by equity analysts in their valuation models. They attribute this wedge to analysts' efforts to try to mitigate noise in model estimates, similar to our argument as to why one might want to remove the IFF from the VMF. If, for instance, analysts apply CAPM logic but use alternative market indices that are less susceptible to the IFF (e.g., such as the DMF), then such a result obtains quite naturally.

Finally, Table 2 reports estimates of the beta gap from size-by-beta-sorted portfolios and offers perhaps the most direct test of our predictions about beta gap behavior. Recall from equation (15) that true betas are attenuated downward by a multiplicative factor implying that, holding the covariance between ϵ_{t+1} and $r_{p,t+1}$ (which depends mostly on size) constant, the attenuation bias is larger for higher-beta portfolios. We closely approximate this comparative static by moving down columns of Table 2 (i.e., increasing beta, holding size fixed). Indeed, beta gaps become more negative as we move down *every* column in the table. Likewise, when holding β_p approximately constant by moving across columns within the same row, the beta gap becomes more negative for smaller-cap portfolios (i.e., moving

right to left), which is also consistent with the intuition from equation (15).

Overall, this evidence is consistent with Prediction 1 from our theoretical model that small-cap betas are biased downward and the additional prediction from equation (15) that high-beta portfolio betas are biased downward when estimated using the VMF compared to when estimated using the DMF.

6.2 The IFF and the size anomaly

In this section, we test Prediction 2 from our theoretical model that implies smaller stocks will earn positive abnormal returns when the VMF is used as the market factor and when size factors are omitted. While our model does not include any additional priced risk factors, we demonstrate empirically that a similar result holds even when controlling for other risk factors unrelated to size. We test this prediction in Figure 6, where we report alphas for 10 value-weighted size-sorted portfolios (a subset of our 372 portfolios) estimated using nine standard factor models including 1) the CAPM, 2) the Fama and French (1993) three-factor model ("FF3"), 3) the Carhart (1997) four-factor model ("FF3C"), 4) the Fama and French (2015) five-factor model ("FF5"), 5) the Hou et al. (2015) four-factor model ("q4"), 6) the Stambaugh and Yuan (2017) four-factor model ("SY"), 7) the Daniel et al. (2020) three-factor model ("DHS"), 8) the Hou et al. (2021) five-factor model ("q5"), and 9) the Chabi-Yo et al. (2024) three factor ICAPM-inspired model ("ICAPM").[16] We drop size factors from the six models with size factors (FF3, FF3C, FF5, q4, SY, and q5) in this analysis because, consistent with our theoretical model and as we argue below, the primary role of these factors is help correct for attenuation in the market beta estimates that we study here.

Figure 6 Panel 1 reports results when the VMF is used as the market factor in each model. In this case, the small-cap portfolio (decile 1) yields positive and statistically significant alphas across all models (excluding FF3 and FF3C). Within each model, the alphas are approximately monotonically decreasing in portfolio deciles. Panel 2 reports results when the VMF is replaced by the DMF in each model. In this case, the small-cap portfolio alphas

[16]We report robustness checks on the results in Figure 6 in Online Appendix Figure OA.9 using the alternative DMF constructions described in Section 4.2.

are statistically indistinguishable from zero in three cases (q5, q4, and FF5) and slightly negative in other cases. Within each model, there do not appear to be strong patterns in the alphas across portfolio deciles. Economically, the magnitudes of the small-cap alphas are much smaller when using the DMF than they are when using the VMF. Overall, these patterns are consistent with the prediction from our theoretical model that small-cap alphas are positive under the VMF and (approximately) zero under the DMF.

Two other features are worth noting about Figure 6. First, just as *most portfolios* have DMF betas larger than VMF betas (see Figure 4), *most stocks* do not load on the IFF and therefore have alphas that are positive when the VMF is used as the market proxy and size factors are omitted. In Figure 6 Panel 1, alphas are positive and statistically significant for most models in almost all size deciles below the top one. Through the lens of the model, the reason why alphas are larger for the smallest stocks is because they tend to have higher DMF betas, which increases the attenuation in their estimated exposures to priced market risk. Second, the model predicts that alphas are positive even if the factor premium on the size factor is zero, as would happen in the model if the average DMF beta of small stocks was equal to that of large stocks. Consistent with this point, we find similar results to those in Figure 6 when we restrict to a post-1993 sample (see Online Appendix Figure OA.12). During this sample, the SMB factor earned average annualized returns of 1.4%, considerably lower than its 3.3% return over the 1963-1992 period. Thus, independent of the debate about whether the size factor premium has "gone away" in recent years, we see that most workhorse models continue to struggle to price the average stock outside of the top size decile when size factors are omitted.

6.3 The IFF and the role of size factors

In this section, we test Prediction 3 from our theoretical model that, when the IFF is not priced (as is the case in the model as $\phi \to 1$), size factors become irrelevant if we use the DMF in place of the VMF.

6.3.1 DMF prices and obviates the need for size factors

Our model provides a new interpretation for the role of size factors. Rather than capturing risks that are particularly important for small stocks, they help to address the fact that the VMF includes the IFF, which carries a low price of risk and is mostly relevant for large stocks. Accordingly, Prediction 3 from our theoretical model implies size factors are spanned by the DMF and therefore do not improve pricing errors when the DMF is used as a market proxy. By contrast, size factors are not spanned when the VMF is used and are therefore required to correctly price assets. To capture forces outside of the model, we test these spanning predictions after again controlling for additional factors other than size.

Table 3 reports size factor alphas from regressions of size factors in each model on the other model factors when using either the VMF (Panel A) or the DMF (Panel B) as the market factor (for the six models that include size factors).[17] 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. This is one piece of motivation for including size factors in the first place – a model's pricing ability is improved when unspanned factors are added to the model. However, when we use the DMF, the size factors are spanned by the other factors in all models (i.e., the size factor alphas are both economically and statistically insignificant) implying these factors do not improve each model's pricing ability (Barillas and Shanken, 2017). We also report betas of the size factors with respect to both the VMF and the DMF. In line with our theory, size factors have higher loadings on the DMF than on the VMF, which allows models using the DMF to span size factors. Thus, using the DMF in place of the VMF effectively allows us to remove the size factors from all of these models without impacting their pricing implications.

Another way to highlight how replacing the VMF with the DMF obviates the need to include size factors is by studying each model's ability to price a broad set of test assets. Thus, we estimate alphas for each of the 372 portfolios we study using four different versions of each model: 1) Model using the VMF with the size factor included, 2) Model using the

[17] We report robustness checks on the results in Table 3 in Online Appendix Table OA.4 using the alternative DMF constructions described in Section 4.2.

VMF with the size factor excluded, 3) Model using the DMF with the size factor included, and 4) Model using the DMF with the size factor excluded. We report the resulting mean absolute pricing errors (MAPEs) across all 372 portfolios in Figure 7. Panel 1 plots results for value-weighted portfolios, and Panel 2 plots results for equal-weighted portfolios.[18]

When the VMF is used (red bars), the MAPE is always higher when the size factor is excluded as a control for both value- and equal-weighted portfolios. By contrast, when the DMF is used (blue bars), the MAPE is comparable between the models that include and those that exclude the size factors. The numbers above bars in the graph represent p-values from a test of the null hypothesis that the MAPE from a given model without the size factor is lower than the MAPE from the same model with the size factor for a given VMF- or DMF-based model. For instance, the p-value of 25 associated with the DMF-based FF3 model in Panel 1 indicates that we cannot reject the hypothesis that the MAPE from the DMF model without the size factor is lower than the MAPE from the model that includes the size factor. For the VMF-based models, we can typically reject the null at standard levels (except for the FF3C model with value-weighted portfolios and the FF3 and FF3C models with equal-weighted portfolios). On the other hand, we can never reject the null for the DMF-based models. In short, removing size factors from VMF-based models significantly decreases their ability to price these test assets, which is not the case for DMF-based models.

Among models that exclude size factors, DMF-based MAPEs are always lower than VMF-based MAPEs. These results are in line with the size-sorted portfolio pricing tests in Figure 6 that imply many assets other than the largest-cap stocks are exposed to the IFF. In this case, using the DMF in place of the VMF helps to more correctly capture the market risk of these assets even when a given model does not include size factors.

6.3.2 Inspecting the mechanism: Size factors address beta gaps

The results we demonstrate above are consistent with Prediction 3 from our theoretical model. In the model, including a size factor in addition to the VMF allows the model to separately estimate the exposures of a given asset's return to the DMF and the IFF and

[18]We report robustness checks on the results in Figure 7 in Online Appendix Figure OA.10 using the alternative DMF constructions described in Section 4.2.

therefore price assets correctly. Adding these factors therefore addresses the attenuation in estimated exposures to the priced DMF factor, analogous to the control function approach for addressing endogeneity problems in linear regressions. This "cleansing role" of the size factor is especially transparent to see in the special case of our model in which $\beta_S = 1$. When this is the case $r_{smb} = -\frac{\epsilon}{\bar{v}}$. From the Frisch-Waugh-Lovell theorem, the loading on the market from a bivariate regression of asset i's return, r_i, on r_m and r_{smb} is equal the loading of r_i on the DMF f. In this special case, even though the SMB factor would not earn a risk premium as $\phi \to 1$, controlling for r_{smb} addresses the attenuation in market betas due to the IFF. In the empirically relevant case with $\beta_S > 1$, the SMB factor combines a positive loading on the DMF, $\beta_{smb} > 0$, with a negative loading on the IFF. Nonetheless, both SMB and the VMF contain independent information about the DMF and the IFF, so a two factor model with r_{smb} along with either the VMF or the DMF correctly prices assets.

Our model also predicts that there is a tight connection between the beta gaps documented above and loadings on size factors once they are included along with the VMF. Specifically, if $r_i = \beta_i f + \gamma_i \epsilon + u_i$, then can also be expressed as (see Online Appendix OA.2)

$$r_i = \left(\beta_i - \frac{\beta_{smb}(\beta_i - \gamma_i)}{\beta_{smb} + \frac{1}{\bar{v}}}\right) r_m + \left(\frac{\beta_i - \gamma_i}{\beta_{smb} + \frac{1}{\bar{v}}}\right) r_{smb} + u_i. \tag{16}$$

The exposure to the SMB factor is increasing in the DMF loading β_i and decreasing in the IFF loading γ_i. Conversely, the difference between the DMF and VMF beta in equation (15) is decreasing in β_i and increasing in γ_i, so in the model SMB exposures are inversely related to the beta gaps discussed in Section 6.1.

Figure 8 illustrates that the data tightly conform with this prediction of the theory.[19] Specifically, for each of the factor models we consider, we compare the difference between the estimated market risk exposures when the VMF and DMF are used as risk factors, respectively. These estimates are analogous to the beta gaps discussed above except that they additionally hedge exposure to other factors. Across all of the models we consider, there is a very strong negative relationship between estimated loadings on size factors and the beta

[19] We report robustness checks on the results in Figure 8 in Online Appendix Figure OA.11 using the alternative DMF constructions described in Section 4.2.

gaps. Hence, size factors quite directly "undo" the attenuation in market risk exposures due to the IFF, capturing an important dimension of omitted risk (namely, assets' heterogeneous exposures to the DMF versus the IFF).

What changes when we instead use the DMF as the market factor in the model? The component of the SMB portfolio orthogonal to the DMF equals the IFF in the model. However, since these additional risks are unpriced as $\phi \to 1$, omitting them from the model does not reduce pricing errors. Of course, in reality, the orthogonal component of the SMB factor may also include additional variation from common factors that are relevant for small stocks, but the data suggest that these factors are not priced as is evident from the similar MAPEs in the DMF-based models that either include or exclude size factors in Figure 7.

6.4 IFF pricing

We showed that we cannot reject the null hypothesis that the IFF has no risk premium in Section 4.3. In this section, to test the remainder of Prediction 4 from our theoretical model that the IFF has a negative alpha with respect to the VMF and a zero alpha with respect to the DMF, we run IFF spanning tests using the nine factor models described in Section 6.2. We again consider versions of the models that use either the VMF or the DMF as a market factor and either include or exclude size factors. As implied by the previous subsection, we expect the IFF to be closely related to the 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 the size factors because they short large-cap stocks.

Table 4 presents results from 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.[20] When we use the VMF as the market proxy, the IFF alphas are always negative and typically statistically significant. However, including size factors helps reduce these alphas because the IFF has a negative loading on size. On the other hand, when we use the DMF as the market proxy, we typically cannot reject the null that the IFF alphas are equal to zero regardless of whether

[20]We report robustness checks on the results in Table 4 in Online Appendix Table OA.5 using the alternative DMF constructions described in Section 4.2.

we include size factors. Overall, these results are consistent with Prediction 4 from our model. Furthermore, the fact that the IFF generates a negative alpha with respect to the VMF implies that it is priced quite differently by investors than what we would expect for a market proxy, providing further evidence for the distinction between the DMF and IFF.

7 The IFF and the intertemporal risk-return tradeoff

In this section, we show how the IFF induces attenuation bias in estimates of the market's intertemporal risk-return tradeoff. Next, we provide evidence that this bias generates a weak intertemporal risk-return tradeoff for the VMF compared to that for the DMF.

7.1 Econometric intuition

Suppose a researcher is trying to test the risk-return relationship in equation (12) but only has access to the VMF (and not the DMF). Equation (9) implies that the difference between the VMF and DMF variances is given by

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

where we see the role of the three separate components: time-varying exposures to f variance, time-varying exposures to g covariance, and the granular residual variance. Other than the ambiguous role of the time-varying $\beta_{m,t}$ (which we argue in Online Appendix OA.8 is likely to be small), the last two terms unambiguously increase the variance of the VMF relative to that of the DMF. If these terms are time varying, they also add noise to $Var_t[VMF_{t+1}]$ relative to $\sigma^2_{f,t+1}$. Therefore, we expect the risk-return tradeoff coefficient estimated using the VMF to be biased downward compared to that estimated using the DMF.

To formalize this intuition, we next develop an expression for the risk-return tradeoff coefficient under the VMF. We make three additional assumptions here for expositional simplicity. We can relax each assumption (these are not imposed in our estimation procedure), but doing so increases the complexity of the expressions without adding much intuition. First, we assume that $\beta_{m,t} = 1$, which is approximately satisfied in the data (see Figure

OA.3). Second, consistent with the small economic magnitude and statistical insignificance of the IFF average (Table 1), we assume the IFF risk premium is zero – i.e., $E_t[\epsilon_{t+1}] = 0$. Third, even though conventional estimates of the conditional VMF variance are likely to be misspecified, we abstract away from this issue and assume that the researcher has access to a "perfect" conditional variance estimate for the VMF, $\sigma^2_{VMF,t}$.

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 $\sigma^2_{VMF,t}$. However, when the IFF is present, the OLS regression coefficient will be biased, most likely towards zero due to the additional volatility in the VMF relative to the DMF. As a simple way to see this, the ratio of the true market 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 the IFF. In this case, the OLS estimate of the slope coefficient on $\sigma^2_{m,t+1}$ will not be consistent for γ, instead converging to

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

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 multiplicative term on γ in the above equation will induce a downward bias provided that $Var\left[\sigma^2_{\epsilon,t}\right] \geq -Cov\left[\sigma^2_{f,t}, \sigma^2_{\epsilon,t}\right]$.[21] 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.

7.2 The DMF yields a stronger intertemporal risk-return tradeoff

In this subsection, we investigate how accounting for the IFF affects estimates of the market's intertemporal risk-return tradeoff using GARCH variance forecasting models. To demonstrate the effects of the IFF, we estimate conditional variances of the VMF and DMF using

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

a GARCH(1,1) specification as

$$
\begin{aligned}
r_{t+1} &= \phi_0 + \phi_1 \cdot r_t + \gamma \cdot \sigma_t^2 + u_{t+1}, && (19)\\
u_{t+1} &= \sigma_t v_{t+1}, && (20)\\
\sigma_{t+1}^2 &= \alpha_0 + \alpha_1 u_t^2 + \alpha_2 \sigma_t^2, && (21)
\end{aligned}
$$

where r_{t+1} is either VMF_{t+1} or DMF_{t+1}, u_t is a forecasting error from the previous period, and v_{t+1} has unit variance.

While we adopt the GARCH(1,1) as our baseline variance forecasting model, there are many GARCH model extensions aimed at better capturing pervasive features in the data such as volatility persistence, asymmetry, and the leptokurtic distribution of financial asset returns. Thus, we complement our baseline GARCH(1,1) with estimates from two additional nonlinear GARCH models: the 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 are widely used in the empirical literature to capture "leverage effects."[22] Additional details related to the models and estimation are provided in Online Appendix OA.11.

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

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

We provide estimates of the conditional volatility from the baseline GARCH specification in Figure 9. As expected from the unconditional analysis in Table 1 and our commentary above, the conditional DMF volatility is typically lower than the conditional VMF volatility. 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

[22]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).

to the lagged forecasting errors used to generate the conditional variance forecasts within the GARCH forecasting models based on the VMF.

Table 5 reports γ values estimated according to the above specifications. 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 return proxy. The first column uses the VMF, the second column uses our main DMF specification, and the remainder use the alternative DMF specifications described in Section 4.2. The reported γ differences are differences between γs estimated using the VMF and γs estimated using different versions of the DMF. The first set of results in each panel defines risk according to the variance specification in equation (19), and the second set of results defines risk according to the volatility specification in equation (22). The DMF generates positive and statistically significant γ estimates across all specifications. However, the VMF generates positive estimates that are always statistically insignificant. Furthermore, our VMF-based γ estimates are always lower than our DMF-based estimates, consistent with the IFF generating attenuation bias in the VMF-based estimates.

While we focus on the relationship between the conditional mean and variance of the stock market for brevity, our argument about the IFF generating instability in predictive regressions applies more broadly. Consistent with this idea and motivated by our findings, Schmidt (2024) provides evidence that the rate of involuntary job loss and a number of other standard predictors from the literature (including the realized variance) have considerably stronger forecasting power for the DMF than for the VMF.

8 Conclusion

We provide evidence that the standard value-weighted market factor contains an "idiosyncratic" component (the IFF) that is unrelated to proxies for investor marginal utility and carries no risk premium. We show, both theoretically and empirically, how the existence of the IFF can be used to explain the size effect. The crux of our argument is that the IFF induces unpriced risk in the VMF that leads to a mismeasurement of market risk that is particularly severe for small-cap firms. We show how replacing the standard value-weighted

market factor with our alternative dominant market factor explains average returns across size-sorted portfolios (i.e., the size premium) and obviates the need to include size factors in all standard factor models we investigate. Furthermore, consistent with the notion that the IFF introduces unpriced risk into the VMF, the DMF implies a significant intertemporal market risk-return tradeoff whereas the VMF does not.

While value weighting has many advantages (e.g., lower rebalancing costs, theoretical justification from the CAPM), we uncover a potential disadvantage related to overweighting large-cap stocks. Although we focus on implications with respect to the VMF, we do not explore implications for other pricing factors from standard factor models that traditionally use value weights, which we leave for future research. Furthermore, our mechanism could help explain some related "puzzling" results, including the observations that the CAPM works better on announcement days (Savor and Wilson, 2014) and that the CAPM works well for jumps in market returns but not well for diffusive shocks (Bollerslev et al., 2016).[23]

Other aggregate measures commonly used in empirical asset pricing may be affected by instability coming from placing high weights on a small number of large-cap firms. For instance, a similar effect could contaminate aggregate valuation ratios with terms unrelated to the key source of predictability of interest. As such, making similar corrections to the aggregate price-dividend ratio could generate stronger return/cash flow predictability. The overarching point we highlight is that value-weighted indices may reflect many sources of risk that are not important for the average firm or investor marginal utility and that stripping out these risks can lead to appreciably different conclusions when testing related theories.

References

Alquist, R., R. Israel, and T. Moskowitz (2018): "Fact, Fiction, and the Size Effect," *The Journal of Portfolio Management*, 45, 3–30.

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.

[23]Our framework could explain these results if the conditional variance of the DMF relative to that of the IFF is larger during these isolated time periods since it is more related to macroeconomic aggregates, causing standard measures to more accurately reflect the quantity of risk during these periods.

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

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): "Volatility, the macroeconomy, and asset prices," *The Journal of Finance*, 69, 2471–2511.

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.

Basak, S. and D. Cuoco (1998): "An equilibrium model with restricted stock market participation," *The Review of Financial Studies*, 11, 309–341.

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

Black, F. (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.

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

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. 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 (2024): "An Intertemporal Risk Factor Model," *Management Science (forthcoming)*.

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. (2009): *Asset pricing: Revised edition*, 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.

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.

Decaire, P. H. and J. R. Graham (2024): "Valuation Fundamentals," Working paper.

Dou, W. W., L. Kogan, and W. Wu (2024): "Common Fund Flows: Flow Hedging and Factor Pricing," *Journal of Finance, Forthcoming*.

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.

Fama, E. F. and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607–636.

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.

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.

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.

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.

He, Z., B. Kelly, and A. Manela (2017): "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, 126, 1–35.

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

Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh (2016): "The common factor in idiosyncratic volatility: Quantitative asset pricing implications," *Journal of Financial Economics*, 119, 249–283.

Hou, K. and M. A. van Dijk (2019): "Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns," *Review of Financial Studies*, 32, 2850–2889.

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.

Jensen, T. I., B. Kelly, and L. H. Pedersen (2023): "Is There a Replication Crisis in Finance?" *Journal of Finance*, 78, 2465–2518.

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

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105, 1177–1216.

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.

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.

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

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

Lucas, Robert E., J. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 631–653.

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

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.

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

Newey, W. K. and K. D. West (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 703–708.

——— (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

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.

Panageas, S. et al. (2020): "The implications of heterogeneity and inequality for asset pricing," *Foundations and Trends® in Finance*, 12, 199–275.

Parker, J. A. and A. Vissing-Jørgensen (2009): "Who bears aggregate fluctuations and how?" *American Economic Review*, 99, 399–405.

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. (2024): "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.

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.

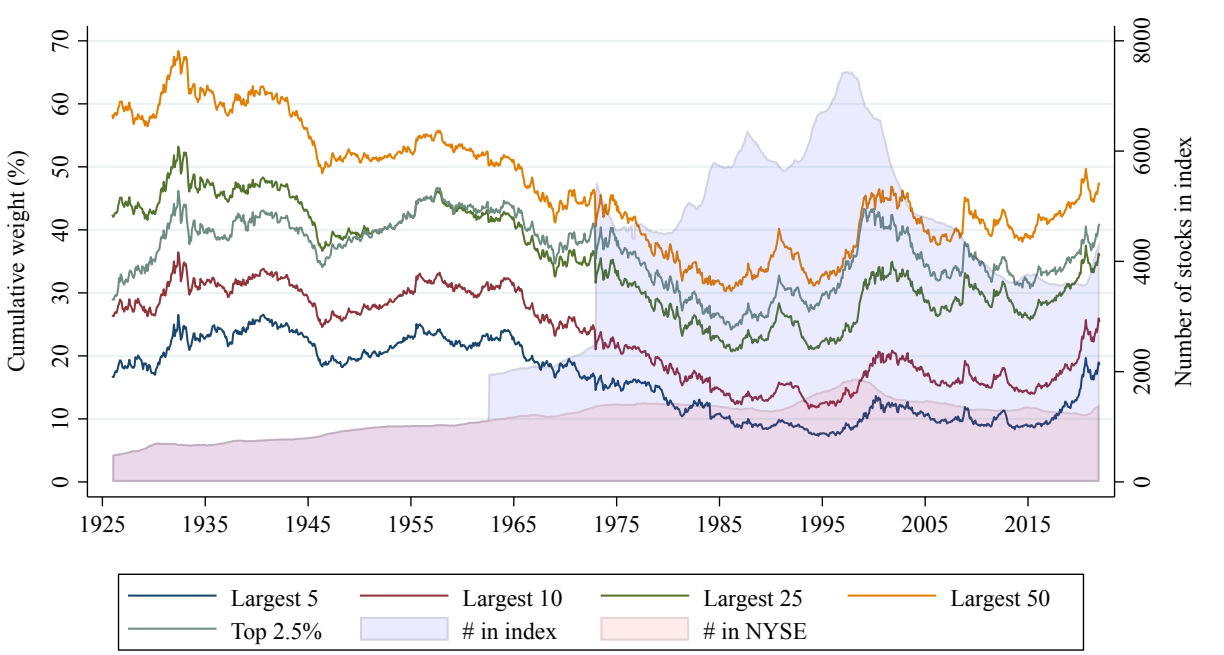
Stock, J. H. and M. W. Watson (2002): "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

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

van Dijk, M. A. (2011): "Is size dead? A review of the size effect in equity returns," *Journal of Banking and Finance*, 35, 3263–3274.

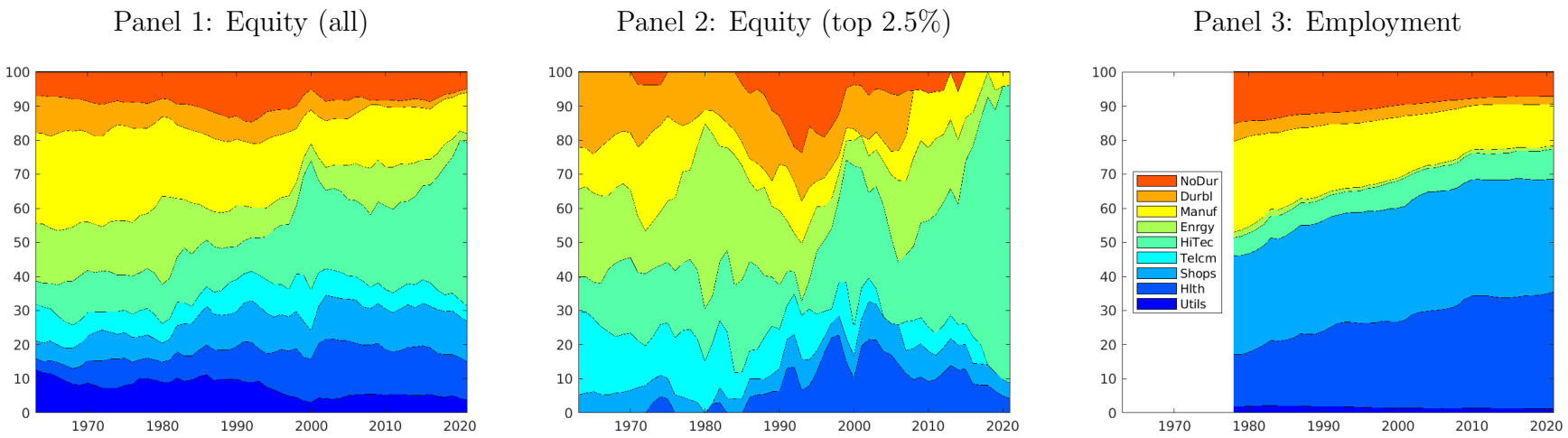
Whitelaw, R. F. (1994): "Time variations and covariations in the expectation and volatility of stock market returns," *Journal of Finance*, 49, 515–541.

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



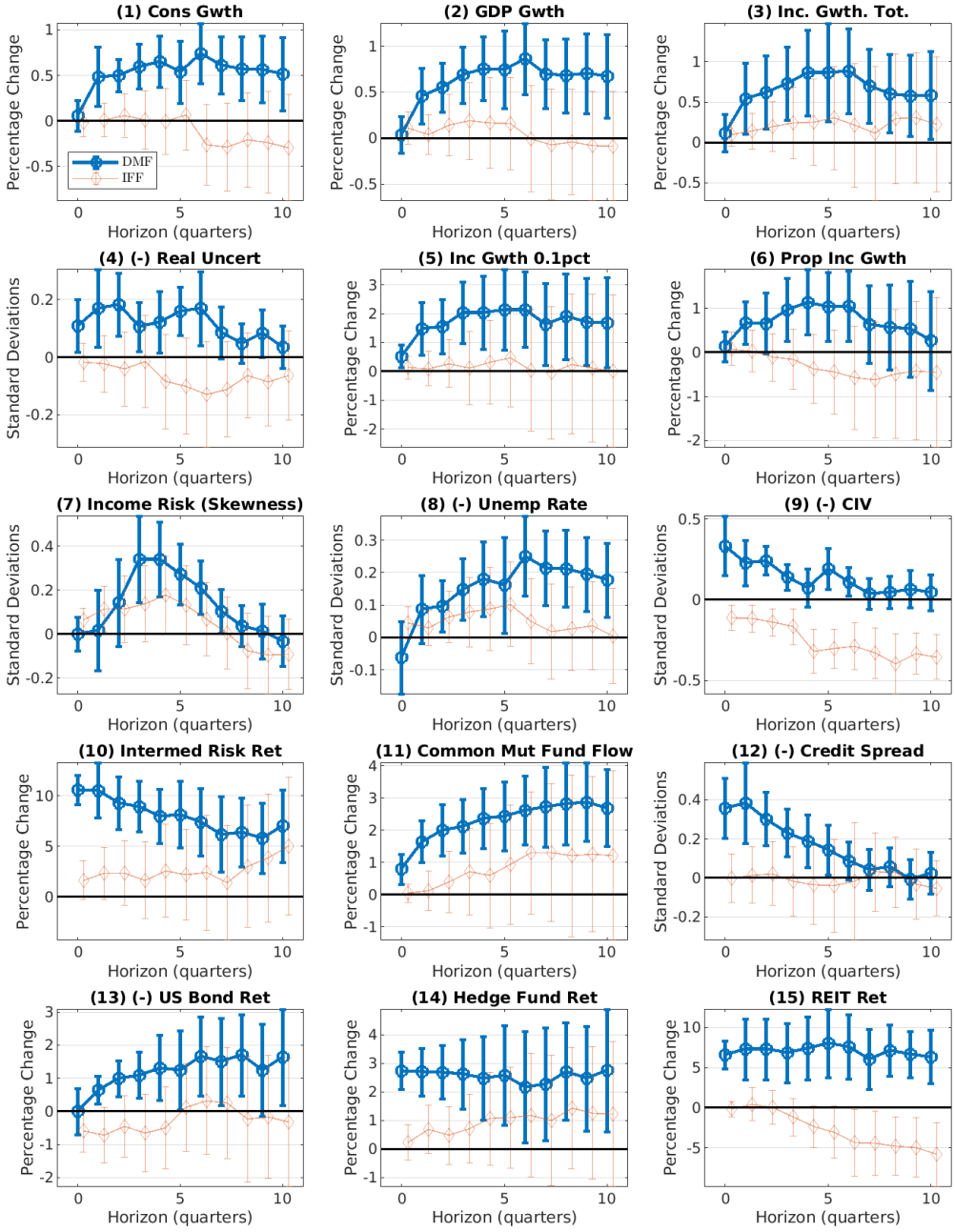
Note. This figure plots the combined weights of individual stocks for the 5, 10, 25, and 50 largest stocks in the CRSP value-weighted index (left axis). The top 2.5% line, also on the left axis, plots the weight associated with stocks whose market capitalization puts them among the largest stocks with the rank cutoff based on 2.5% of the total firms in the NYSE. We use the NYSE breakpoint to avoid mechanical decreases in the series as many (mostly small) AMEX and NASDAQ firms are added to the database over time. The right axis plots the total number of stocks included in the CRSP index as well as in 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 other firms in the CRSP value-weighted index.

Figure 2: Industry compositions within public equities and employment



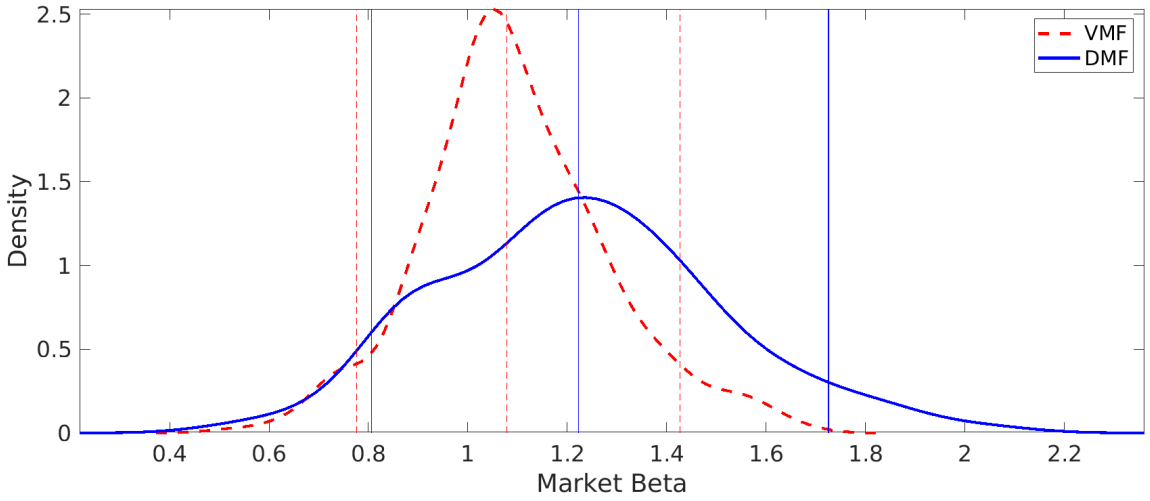
Note: This figure plots the share of public equity value among all publicly traded stocks (Panel 1), public equity value among the top 2.5% of stocks by NYSE market capitalization (Panel 2), and total employment (Panel 3) by each of the 10 Fama-French industry categories by year. See Kenneth French's website for detailed industry definitions by SIC code and description. We exclude the Other category since it is comprised of a number of industries and makes up a large (but stable) share of employment (∼45%), making visual comparisons across other industries less clear. However, the percentages cited in Section 2.2 include the Other category in their calculation. Panel 1 is constructed by aggregating the value of each stock in each industry using SIC codes from Compustat when available (otherwise, we use SIC codes from CRSP). We use an NYSE breakpoint to define the top 2.5% of stocks by size each period to construct Panel 2 to avoid generating mechanical increases in the composition of stocks in the top category, as many small AMEX and NASDAQ stocks are added to the database at various points in its history. Panel 3 is constructed by aggregating employment count data from the US Census by NAICS code. NAICS codes are matched to the SIC codes used to define the 10 Fama-French industries using the linking table provided by David Dorn (https://www.ddorn.net/data). The time periods represented in each panel are based on data availability. Data for the US equity market span 1963–2021 and that for employment span 1978–2021. The x-axes in each plot have the same scale for ease of comparison.

Figure 3: Local projection impulse responses to the DMF and IFF



Note. This figure plots local projection impulse responses of selected macroeconomic aggregates and financial return time series to the DMF and IFF. The indexes are defined in Section 5 and are signed so that increases in the variable value correspond to decreases in marginal utility in standard models. A "(-)" in the panel title indicates that the negative value of a time series is used to conform to this logic. Monthly returns are compounded to 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 response variable. Response variables are in units of percent in the case of return or growth time series (e.g., consumption growth, US bond returns, etc.) or units of standard deviation in the case of stationary time series (e.g., unemployment rate, real uncertainty, etc.). The initial point estimates (i.e., when the horizon is equal to zero) 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. All data are from 1963–2021 (conditional on variable availability). Additional details on the estimation procedure can be found in Online Appendix OA.10.

Figure 4: Distribution of β_p^{VMF} and β_p^{DMF} for a large number of test assets

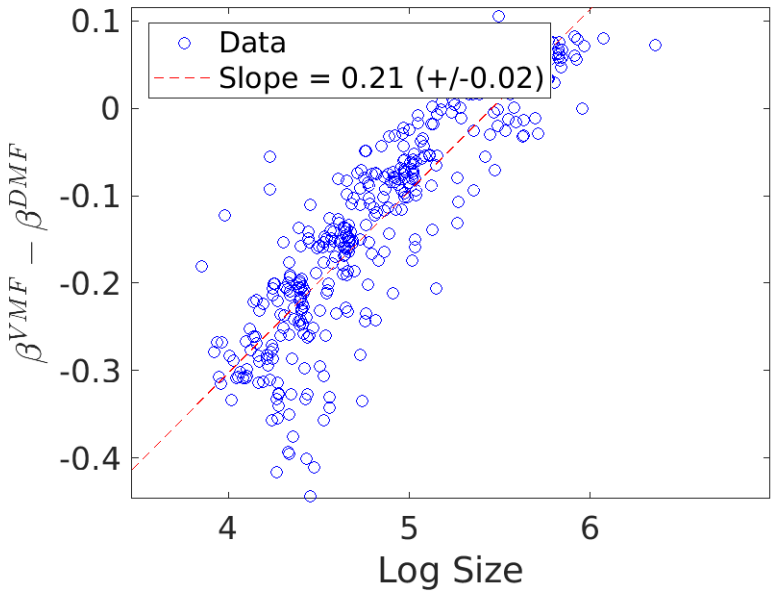


			Market Beta Percentile								
	Mean	Stdev.	1	5	10	25	50	75	90	95	99
β_p^{VMF}	1.09	0.19	0.67	0.78	0.88	0.99	1.08	1.21	1.34	1.43	1.60
β_p^{DMF}	1.22	0.28	0.59	0.81	0.87	1.02	1.22	1.40	1.60	1.73	1.94
$\beta_p^{VMF} - \beta_p^{DMF}$	-0.13	0.13	0.08	0.06	0.05	-0.03	-0.13	-0.22	-0.30	-0.33	-0.42

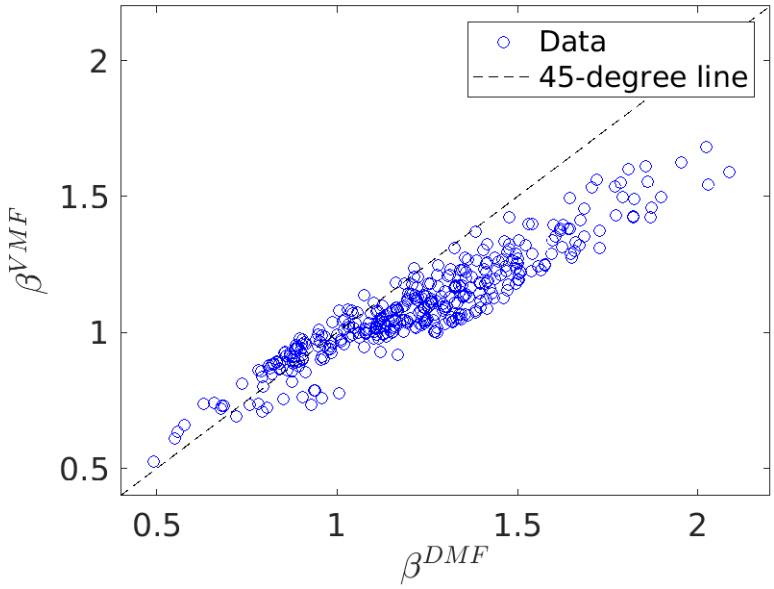
Note. This figure plots kernel density estimates of market betas for our set of 372 value-weighted portfolios estimated using the VMF and DMF. See Online Appendix Table OA.1 for additional details about these portfolios. Color-coded vertical lines indicate the 5th, 50th, and 95th percentiles of the two distributions. We also tabulate a number of univariate summary statistics for each distribution. All data are from 1963–2021.

Figure 5: VMF and DMF beta comparisons using a large number of test assets

Panel 1: Beta gap vs. portfolio size



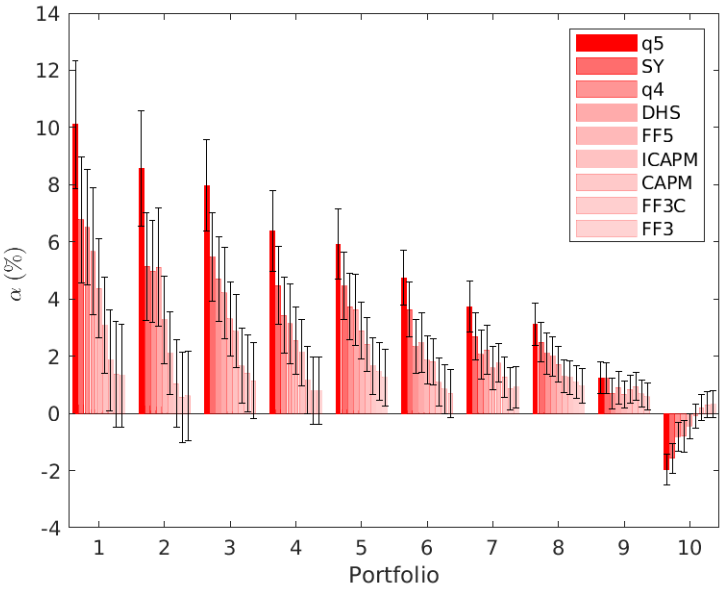
Panel 2: VMF vs. DMF betas



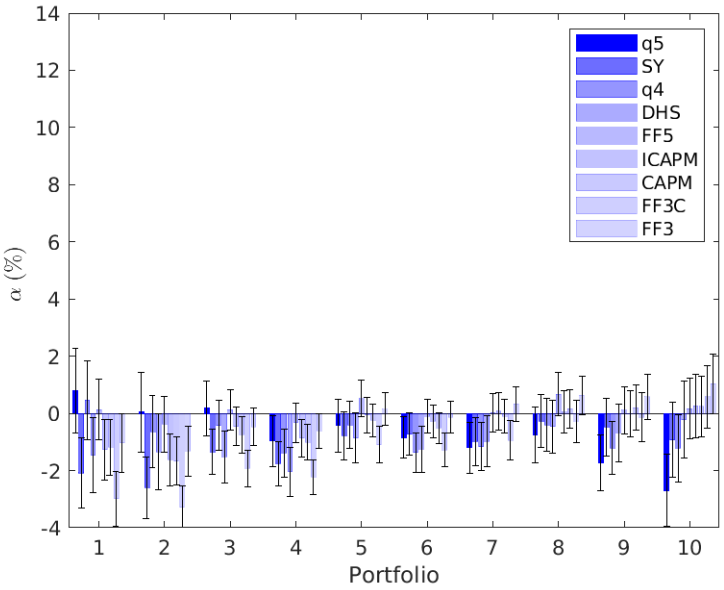
Note. This figure compares VMF and DMF betas using our set of 372 standard characteristics-sorted portfolios. Panel 1 compares the beta gap ($\beta_p^{VMF} - \beta_p^{DMF}$) as a function of each portfolio's average log market equity. The dotted red line is the OLS best fit line. Log-size is computed as the average of each portfolio's monthly log market capitalization. Panel 2 provides a scatterplot of β_p^{VMF} versus β_p^{DMF}. All data are from 1963–2021.

Figure 6: Alphas for size-sorted decile portfolios using the VMF or the DMF

Panel 1: VMF-based models (w/o size factors)

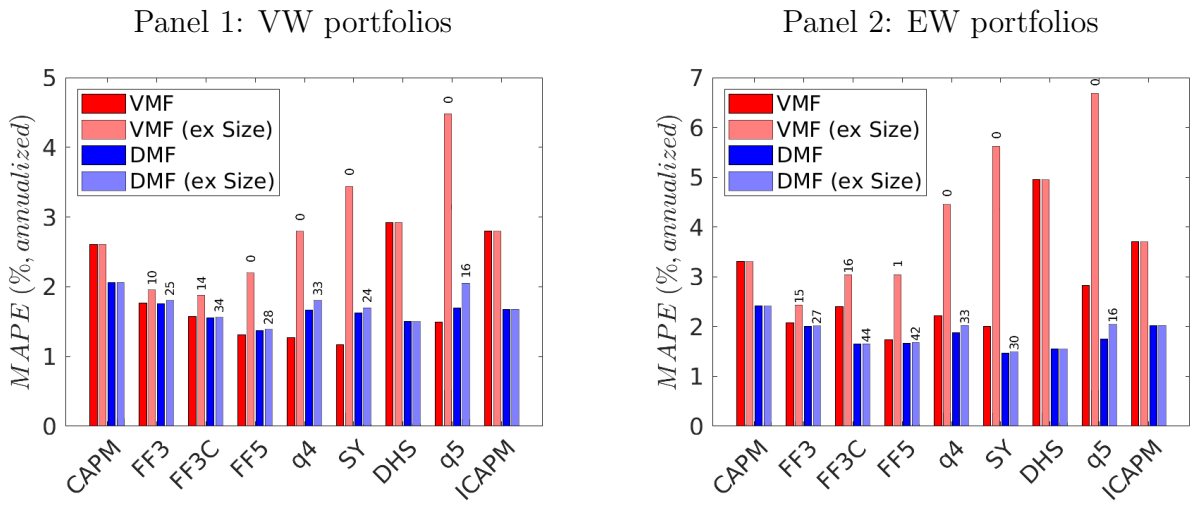


Panel 2: DMF-based models (w/o size factors)



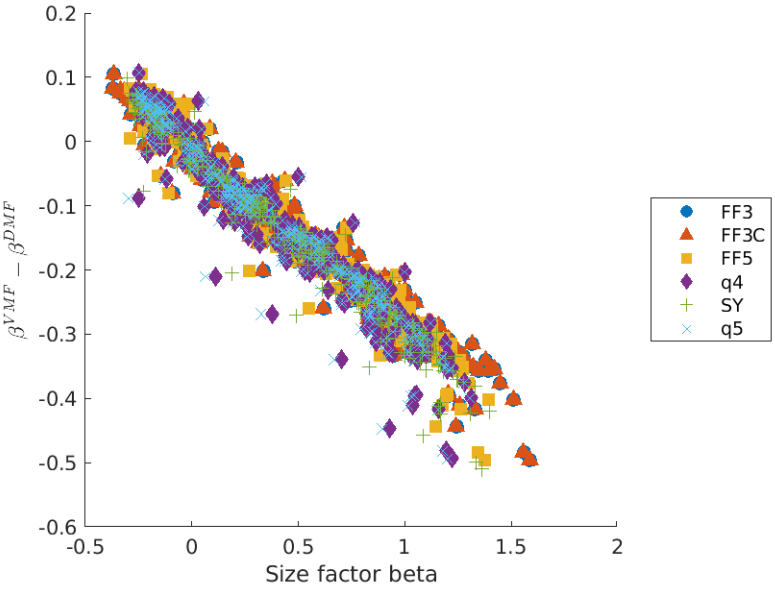
Note. This figure reports alphas for value-weighted size-sorted decile portfolios (obtained from our set of 372 characteristics-sorted portfolios) for each of the nine factor models described in Section 6.2. We drop size factors from the six models with size factors (FF3, FF3C, FF5, q4, SY, and q5) in this analysis because, consistent with our theoretical model and our empirical results, the primary role of these factors is help correct for attenuation in the market beta estimates that we study here. All portfolio returns were obtained from Kenneth French's webpage and are described in Online Appendix Table OA.1. Panel 1 uses the VMF as the market factor in each model. Panel 2 uses the DMF as the market factor in each model. The error bars represent 95% confidence intervals on each alpha estimate based on the bootstrap procedure described in Online Appendix OA.12. All data are from 1963–2021 (conditional on factor model availability).

Figure 7: Mean absolute pricing errors for a large number of test assets



Note. This figure plots the mean absolute pricing errors (α) for our set of 372 characteristics-sorted portfolios as test assets across the nine standard factor models described in Section 6.2 and listed in chronological order based on publication date. Panel 1 uses value-weighted portfolios. Panel 2 uses equal-weighted portfolios. All portfolio returns were obtained from Kenneth French's webpage and are described in Online Appendix Table OA.1. We study four versions of each model: 1) "VMF": Model that uses the VMF as the market factor and includes size factors when present, 2) "VMF (ex Size)": Model that uses the VMF as the market factor and excludes size factors when present, 3) "DMF": Model that uses the DMF as the market factor and includes size factors when present, and 4) "DMF (ex Size)": Model that uses the DMF as the market factor and excludes size factors when present. For models that include size factors, the numbers above the bars represent p-values (in percent) from testing the null hypothesis that models that exclude the size factors yield a lower mean absolute pricing error that the alternative model that excludes the size factor. P-values are based on the bootstrap procedure described in Online Appendix OA.12. All data are from 1963–2021 (conditional on factor model availability).

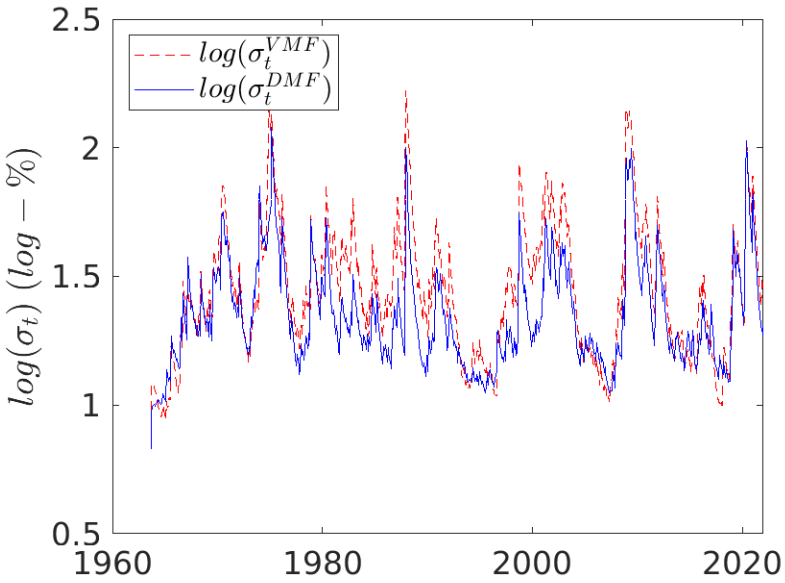
Figure 8: Beta gap vs. size factor betas



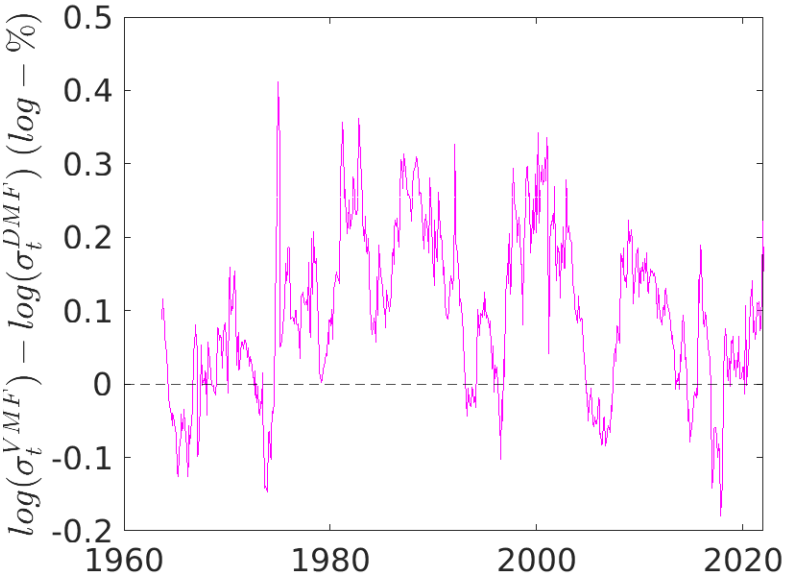
Note. This figure reports the beta gaps ($\beta_p^{VMF} - \beta_p^{DMF}$) versus size factor betas among the six factor models we study that include size factors (FF3, FF3C, FF5, q4, SY, and q5; see Section 6.2 for a description of these models) using value-weighted versions of our set of 372 characteristics-sorted portfolios as test assets. All portfolio returns were obtained from Kenneth French's webpage and are described in Online Appendix Table OA.1. β_p^{VMF} and β_p^{DMF} are estimated as multivariate betas within each model that use either the VMF or the DMF as the market factor. The size factor betas are estimated as multivariate betas within each model that uses the VMF as the market factor. All data are from 1963–2021 (conditional on factor model availability).

Figure 9: VMF and DMF conditional volatility comparison

Panel 1: Volatilities



Panel 2: Difference in volatilities



Note. This figure compares VMF and DMF conditional volatilities estimated using a GARCH(1,1) model. Panel 1 plots in-sample forecasts for the conditional volatility of each index (on a log scale). The dotted red line represents the conditional volatility forecast estimated for the VMF, whereas the solid blue line represents the forecast obtained for the DMF. Panel 2 plots the difference between the forecasts in Panel 1. All data are from 1963–2021.

Table 1: Summary statistics for the VMF, DMF, and IFF

	Statistic							Correlations					
	Mean (%)	t(Mean)	Stdev. (%)	Skew.	Kurt.	Med. (%)	IQR (%)	VMF	EW PCA	Capped VW PCA	No Sm/Lg EW PCA	Rolling EW PCA	EW MKT
Panel A: Market portfolio proxies													
VMF	6.79***	[3.23]	15.36	-0.56	5.07	11.12	18.70	1.00	0.92	0.94	0.93	0.92	0.87
DMF (EW PCA)	7.31***	[3.68]	14.10	-0.35	5.83	9.69	15.98	0.92	1.00	0.99	0.99	0.99	0.98
DMF (Capped VW PCA)	7.12***	[3.61]	14.51	-0.50	5.56	10.60	17.20	0.94	0.99	1.00	1.00	0.99	0.95
DMF (No Sm/Lg EW PCA)	7.08***	[3.72]	14.32	-0.45	5.70	10.39	16.41	0.93	0.99	1.00	1.00	0.98	0.94
DMF (Rolling EW PCA)	7.08***	[3.59]	14.19	-0.51	6.33	9.55	16.31	0.92	0.99	0.99	0.98	1.00	0.97
DMF (EW MKT)	6.43***	[3.25]	13.29	-0.23	6.10	8.78	14.71	0.87	0.98	0.95	0.94	0.97	1.00
Panel B: IFF proxies													
VMF								1.00	0.40	0.33	0.36	0.38	0.50
IFF (EW PCA)	-0.53	[-0.55]	6.09	-0.37	5.22	0.52	6.85	0.40	1.00	0.96	0.93	0.97	0.94
IFF (Capped VW PCA)	-0.33	[-0.41]	5.05	-0.24	4.50	0.15	5.66	0.33	0.96	1.00	0.98	0.92	0.83
IFF (No Sm/Lg EW PCA)	-0.29	[-0.32]	5.56	-0.20	4.67	0.62	6.08	0.36	0.93	0.98	1.00	0.90	0.78
IFF (Rolling EW PCA)	-0.29	[-0.33]	5.90	-0.45	5.78	1.03	6.87	0.38	0.97	0.92	0.90	1.00	0.91
IFF (EW MKT)	0.36	[0.32]	7.70	-0.33	6.00	1.42	8.23	0.50	0.94	0.83	0.78	0.91	1.00

Note: This table provides summary statistics for monthly excess returns on the VMF, DMF variants, and respective IFFs. Panel A reports results for the VMF and DMF variants. Panel B reports results for the IFF variants. We consider five different variants of the DMF: 1) "DMF (EW PCA)" is constructed as a projection of the VMF onto the first PC from a PCA using our set of 372 equal-weighted standard characteristics-sorted portfolios described in Online Appendix Table OA.1, 2) "DMF (Capped VW PCA)" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio, 3) "DMF (No Sm/Lg EW PCA)" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints, 4) "DMF (Rolling EW PCA)" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify PCA portfolio loadings each month (using the standard 372 equal-weighted portfolios), and 5) "DMF (EW MKT)" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index. The IFFs are obtained by subtracting the respective DMF variants from the VMF. We report the mean, standard deviation, skewness, and kurtosis for each variable. Given the high skewness and excess kurtosis of these distributions, we also report the median and the interquartile range (IQR) as robust measures of location and scale. We also provide correlation matrices among the VMF/DMF and the VMF/IFF. 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 ($\sqrt{12}$). T-statistics are in square brackets and are estimated according to Newey and West (1987, 1994). */**/*** represent statistical significance of the mean values at 90%, 95%, and 99%, respectively. All data are from 1963–2021.

Table 2: Difference between the 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 to test the predictions of equation (15). In particular, this table reports $\beta^{VMF} - \beta^{DMF}$. Estimates are from monthly returns on 25 value-weighted size-by-beta-sorted portfolios from Kenneth French's data library. Standard errors are in parentheses and are computed according to the bootstrap procedure described in Online Appendix OA.12. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from 1963–2021.

Table 3: Size factor spanning tests

	Factor model					
	FF3	FF3C	FF5	q4	SY	q5
Panel A: VMF as the market factor						
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: DMF as the market factor						
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]

Note: This table presents factor spanning tests for the size factors in each factor model when the VMF (Panel A) or the DMF (Panel B) is used as the market factor in each model. We only include models with size factors and test each model using its respective size factor. α^{VMF} (α^{DMF}) represents the alpha when the VMF (DMF) is used; it is annualized and in percent. β^{VMF}_{Size} (β^{DMF}_{Size}) represents the size factor beta with respect to the VMF (DMF). T-statistics are reported in brackets and are based on the bootstrap procedure described in Online Appendix OA.12. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from 1963–2021 (conditional on factor model availability).

Table 4: IFF spanning tests

	Factor model								
	CAPM	FF3	FF3C	FF5	q4	SY	DHS	q5	ICAPM
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)	-1.59**	-1.00	-2.04***	-2.09***	-3.37***	-4.33***	-3.80***	-5.18***	-2.23***
	[-2.24]	[-1.44]	[-2.74]	[-2.95]	[-4.29]	[-4.97]	[-3.94]	[-6.23]	[-3.33]
β_{IFF}^{VMF} (w/o size)	0.16***	0.14***	0.16***	0.17***	0.19***	0.24***	0.21***	0.22***	0.31***
	[8.15]	[7.74]	[9.33]	[9.46]	[11.30]	[10.44]	[9.38]	[12.54]	[13.18]
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)	-0.53	0.17	-0.57	-0.24	-1.48*	-1.09	-0.90	-2.32**	-0.47
	[-0.68]	[0.23]	[-0.71]	[-0.32]	[-1.71]	[-1.09]	[-0.90]	[-2.51]	[-0.59]
β_{IFF}^{VMF} (w/o size)	0.00	-0.01	0.01	0.00	0.04*	0.01	0.01	0.06**	0.04
	[0.00]	[-0.60]	[0.46]	[0.05]	[1.87]	[0.25]	[0.46]	[2.56]	[1.22]

Note: This table presents factor spanning tests for the IFF using each factor model when the VMF or the DMF is used as the market factor in each model. Panel A (Panel B) presents results when the VMF (DMF) is used as the market factor. Within each panel, we run tests using models that include their size factors (designated as "w/ size") and that exclude their size factors (designated as "w/o size"). Note that the CAPM, DHS, and ICAPM models do not include size factors, and therefore we do not present any results "w/ size" in these cases. α is annualized and in percent. β_{IFF}^{VMF} (β_{IFF}^{DMF}) represents the IFF market βs when the VMF (DMF) is used as the market proxy. β_{IFF}^{Size} represents the 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.12. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from 1963–2021 (conditional on factor model availability).

Table 5: Risk-return tradeoff coefficient (γ) across GARCH models

Risk defn.	Parameter	VMF	DMF (EW PCA)	DMF (Capped VW PCA)	DMF (No Sm/Lg EW PCA)	DMF (Rolling EW PCA)	DMF (EW MKT)
Panel A: GARCH							
Variance	γ	2.47	7.98***	7.37***	7.99***	5.67***	7.01***
(σ_t^2)		(2.01)	(2.62)	(2.72)	(2.70)	(2.17)	(2.53)
	$\gamma^{VMF} - \gamma^{DMF}$		-5.51	-4.89	-5.52	-3.20	-4.53
Volatility	γ	0.19	0.68***	0.65***	0.71***	0.47**	0.56***
(σ_t)		(0.18)	(0.22)	(0.25)	(0.24)	(0.18)	(0.20)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.49	-0.46	-0.52	-0.28	-0.38
Panel B: GJR-GARCH							
Variance	γ	1.01	4.17**	3.30**	3.23**	2.85*	5.29***
(σ_t^2)		(1.88)	(1.67)	(1.67)	(1.53)	(1.60)	(1.94)
	$\gamma^{VMF} - \gamma^{DMF}$		-3.16	-2.29	-2.22	-1.84	-4.28
Volatility	γ	0.15	0.44***	0.41**	0.40**	0.35**	0.49***
(σ_t)		(0.19)	(0.16)	(0.17)	(0.16)	(0.16)	(0.17)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.29	-0.26	-0.25	-0.20	-0.34
Panel C: E-GARCH							
Variance	γ	1.61	6.17***	6.08***	6.84***	5.08***	5.78**
(σ_t^2)		(1.88)	(2.12)	(2.20)	(2.15)	(1.96)	(2.26)
	$\gamma^{VMF} - \gamma^{DMF}$		-4.56	-4.48	-5.23	-3.47	-4.17
Volatility	γ	0.18	0.46***	0.49***	0.54***	0.41***	0.39**
(σ_t)		(0.17)	(0.17)	(0.18)	(0.18)	(0.15)	(0.17)
	$\gamma^{VMF} - \gamma^{DMF}$		-0.28	-0.31	-0.36	-0.23	-0.22

Note: This table compares γ estimates and standard errors (in parentheses) based on the VMF with those based on different DMF variants. We consider four different variants of the DMF: 1) "DMF (EW PCA)" is constructed as a projection of the VMF onto the first PC from a PCA using our set of 372 equal-weighted standard characteristics-sorted portfolios described in Online Appendix Table OA.1, 2) "DMF (Capped VW PCA)" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio, 3) "DMF (No Sm/Lg EW PCA)" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints, 4) "DMF (Rolling EW PCA)" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify the PCA portfolio loadings each month (using the standard 372 equal-weighted portfolios), and 5) "DMF (EW MKT)" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index. 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 γs 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 1963–2021.

Online Appendix

"The Idiosyncratic Financial Factor: An Explanation for the Role of Size Factors and the Weak Intertemporal Risk-Return Relation"

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

OA.1 Summary of additional tables & figures

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

OA.1.1 Figures

- Figure OA.1 provides evidence that the portfolio weights of each of our 372 equal-weighted portfolios in the PCA used to define our main DMF measure are positive and relatively similar to one another in magnitude. For more commentary, see Section 4.1 and Online Appendix OA.5.1.

- Figure OA.2 provides evidence that the average portfolio has a loading of zero on higher factors, g, represented here by higher PCs. For more commentary, see Section 4.1 and Online Appendix OA.5.1.

- Figure OA.3 provides estimates of time-varying VMF loadings on PCs 1-5. See Online Appendix OA.8 for additional details.

- Figure OA.4 reports robustness checks on the local projection impulse response (LPIR) results in Figure 3 using the capped value-weighted PCA-based DMF construction described in Section 4.2.

- Figure OA.5 reports robustness checks on the LPIR results in Figure 3 using a PCA-based DMF, where the PCA uses equal-weighted portfolios that drop the lowest and

highest 20th percentiles of stocks (by NYSE breakpoints) before portfolio construction as described in Section 4.2.

- Figure OA.6 reports robustness checks on the LPIR results in Figure 3 using the rolling equal-weighted PCA-based DMF construction described in Section 4.2.
- Figure OA.7 reports robustness checks on the LPIR results in Figure 3 using a DMF constructed by projecting the VMF onto excess returns on the CRSP equal-weighted market index, which is described in Section 4.2.
- Figure OA.8 reports robustness checks on the beta distribution results in Figure 4 using the alternative DMF constructions described in Section 4.2.
- Figure OA.9 reports robustness checks on the size-sorted portfolio alpha results in Figure 6 using the alternative DMF constructions described in Section 4.2.
- Figure OA.10 reports robustness checks on the mean absolute pricing error results in Figure 7 using the alternative DMF constructions described in Section 4.2.
- Figure OA.12 reports robustness checks on the size-sorted portfolio alpha results in Figure 6 using the 1993–2021 sample.
- Figure OA.13 reports robustness checks on the mean absolute pricing error results in Figure 7 using the 1993–2021 sample.

OA.1.2 Tables

- Table OA.1 summarizes the set of 372 portfolios we use in our PCA analysis and as test assets. These portfolios were obtained from Kenneth French's website. In some cases (as in our main PCA), we use equal-weighted versions of the portfolios. In others (as in our asset pricing tests), we use value-weighted versions of the portfolios. In yet others (as in our alternative capped value-weighted PCA exercise described in Section 4.2), we construct our own versions of these portfolios since they are not publicly available.

- Table OA.2 provides results from a regression of the VMF on the DMF and higher PCs. See Online Appendix Section OA.7 for additional details.

- Table OA.3 provides evidence that the granular residual component of the IFF is small. See Online Appendix Section OA.9 for additional details.

- Table OA.4 reports robustness checks on the size factor spanning test results in Table 3 using the alternative DMF constructions described in Section 4.2.

- Table OA.5 reports robustness checks on the IFF spanning test results in Table 4 using the alternative DMF constructions described in Section 4.2.

OA.2 Theoretical model: Construction

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 $i = 1, \ldots, K$ 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 $i = K+1, \ldots, N$ 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 $P_{K+1} = \cdots = P_N = (1-\overline{v})W/(N-K)$.[24]

[24]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 and not within the set of large stocks. Expected dividends are assumed to be sufficiently large or small to justify all stock valuations.

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 $i = 1, \ldots, N$ are assumed to follow a two-factor model:[25]

$$r_i - \mathbb{E}[r_i] = \beta_i \left(f - \mathbb{E}[f]\right) + \zeta_i \left(g - \mathbb{E}[g]\right) + \eta_i \tag{OA.1}$$

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,[26] 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

$$\frac{\overline{v}}{K} \sum_{i=1}^{K} \beta_i + \frac{1 - \overline{v}}{N - K} \sum_{i=K+1}^{N} \beta_i = 1, \tag{OA.2}$$

which implies that the loading of the market portfolio's return on f is equal to 1. We also assume that the average small stock has a loading of zero on g so that[27]

$$\frac{1}{N - K} \sum_{i=K+1}^{N} \zeta_i = 0. \tag{OA.3}$$

The market portfolio is then a function of both the common factors and a weighted average

[25] 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 two factors for parsimony.

[26] 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 individual assets over which the agents are choosing optimal allocations.

[27] 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 the number of 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.

of idiosyncratic components, η_i, of individual stocks that may or may not be diversified away as the number of stocks increases:

$$\begin{aligned} r_m &= \frac{\overline{v}}{K}\sum_{i=1}^{K} r_i + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N} r_i \\ &= \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\mathbb{E}[r_i] + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N}\mathbb{E}[r_i]\right)}_{=\mathbb{E}[r_m]} + \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\beta_i + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N}\beta_i\right)}_{=1}\cdot(f-\mathbb{E}[f]) \\ &\quad + \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\zeta_i+\right)}_{\equiv\zeta_m}\cdot(g-\mathbb{E}[g]) + \underbrace{\left(\frac{\overline{v}}{K}\sum_{i=1}^{K}\eta_i + \frac{1-\overline{v}}{N-K}\sum_{i=K+1}^{N}\eta_i\right)}_{\equiv\eta} \\ &= \mathbb{E}[r_m] + f - \mathbb{E}[f] + \zeta_m\cdot(g-\mathbb{E}[g]) + \eta. \end{aligned} \tag{OA.4}$$

Note that we have used the restrictions in equations OA.2 and OA.3 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

$$Y = \overline{Y} + \phi_f(f-\mathbb{E}[f]) + \phi_g(g-\mathbb{E}[g]) + \phi_\eta\frac{\overline{v}}{K}\sum_{i=1}^{K}\eta_i + \xi \tag{OA.5}$$

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 accommodate 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):

$$M_\varepsilon(t|f,g,\eta_1,\ldots,\eta_N,\xi) = \exp\left(\gamma_0(t) + \gamma_f(t)f + \gamma_g(t)g + \sum_{i=1}^{K}\gamma_i(t)\eta_i + \gamma_\xi(t)\xi\right) \tag{OA.6}$$

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:

$$\begin{aligned}
\tilde{Y} &\equiv Y - \frac{1}{A}\log\mathbb{E}[\exp(-A\varepsilon_j)|f,g,\eta_1,\ldots,\eta_N,\xi] \\
&= \left(\overline{Y} + \phi_f(f-\mathbb{E}[f]) + \phi_g(g-\mathbb{E}[g]) + \phi_\eta\frac{1-\overline{v}}{K}\sum_{i=1}^{K}\eta_i + \xi\right) \\
&\quad - \frac{1}{A}\left(\gamma_0(-A) + \gamma_f(-A)f + \gamma_g(-A)g + \sum_{i=1}^{K}\gamma_i(-A)\eta_i + \gamma_\xi(-A)\xi\right) \\
&\equiv \mathbb{E}\left[\tilde{Y}\right]\cdot\left(1 + \psi_f(f-\mathbb{E}[f]) + \psi_g(g-\mathbb{E}[g]) + \psi_\eta\sum_{i=1}^{K}\eta_i + \psi_\xi\xi\right)
\end{aligned} \tag{OA.7}$$

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:

$$C_j = Y_j + W(1 + r_p(\theta^j)) = Y_j + W\left(\sum_{i=1}^{N}\theta_i^j(1+r_i) + \left[1 - \sum_{i=1}^{N}\theta_i^j\right](1+r_f)\right) \tag{OA.8}$$

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.3, we show that investor j's portfolio choice problem

can be rewritten in the following form:

$$\underset{\theta^j \in \mathbb{R}^N}{\operatorname{argmax}} \{-\mathbb{E}[\exp(-AC_j)]\} = \underset{\theta^j \in \mathbb{R}^N}{\operatorname{argmax}} \left\{ \mathbb{E}[r_p(\theta^j)] - A\mathcal{W} \left(\frac{1-\phi}{2} Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j)) \right) \right\} \tag{OA.9}$$

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.3 derives the following first-order condition for θ_i^j, the portfolio share of any stock $i = 1, \ldots, N$:

$$\mathbb{E}[r_i] - r_f = A\mathcal{W} \left[(1-\phi) Cov(r_i, r_p(\theta^j)) + \phi Cov(r_i, r_h) \right] \tag{OA.10}$$

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 the equilibrium expected excess returns for each stock:

$$\mathbb{E}[r_i] - r_f = \begin{cases} A\mathcal{W} \left[(1-\phi) \left(\beta_i \sigma_f^2 + \zeta_i \zeta_m \sigma_g^2 + \frac{\overline{v}}{K} \sigma_i^2 \right) + \phi \left(\beta_i \psi_f \sigma_f^2 + \zeta_i \psi_g \sigma_g^2 + \psi_\eta \frac{\overline{v}}{K} \sigma_i^2 \right) \right] & \text{for } i = 1, \ldots, K \\ A\mathcal{W} \left[(1-\phi) \left(\beta_i \sigma_f^2 + \zeta_i \zeta_m \sigma_g^2 + \frac{1-\overline{v}}{N-K} \sigma_i^2 \right) + \phi \left(\beta_i \psi_f \sigma_f^2 + \zeta_i \psi_g \sigma_g^2 \right) \right] & \text{for } i = K+1, \ldots, N \end{cases} \tag{OA.11}$$

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.10). This equation pins down average returns on f, g, and η:

$$\mathbb{E}[f] = A\mathcal{W} \left(1 - \phi + \phi \psi_f \right) \sigma_f^2 \tag{OA.12}$$

$$\mathbb{E}[g] = A\mathcal{W} \left((1-\phi) \zeta_m + \phi \psi_g \right) \sigma_g^2 \tag{OA.13}$$

$$\mathbb{E}[r_\eta] = A\mathcal{W}\left((1-\phi)\,\sigma_\eta^2 + \phi\psi_\eta\sigma_{\eta,K}^2\right), \tag{OA.14}$$

where r_η is an η-mimicking portfolio, σ_η^2 is the variance of the market-level granular residual, η, in equation (OA.4), and $\sigma_{\eta,K}^2$ is the variance of the value-weighted large stock granular residuals. We can also use equation (OA.10) to pin down the expected excess return on the market:

$$\mathbb{E}[r_m] - r_f = E[f] + \zeta_m \cdot \mathbb{E}[g] + \mathbb{E}[r_\eta], \tag{OA.15}$$

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

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

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

$$\mathbb{E}[r_i] - r_f = \begin{cases} \beta_i\mathbb{E}[f] + \zeta_i\mathbb{E}[g] + A\mathcal{W}\left(1-\phi+\phi\psi_\eta\right)\frac{\overline{v}}{K}\sigma_i^2 & \textit{for } i = 1,\ldots,K \\ \beta_i\mathbb{E}[f] + \zeta_i\mathbb{E}[g] & \textit{for } i = K+1, K+2,\ldots \end{cases} \tag{OA.17}$$

Furthermore, the market risk premium is given by

$$\mathbb{E}[r_m] - r_f = E[f] + \zeta_m \cdot \mathbb{E}[g] + \mathbb{E}[r_\eta]. \tag{OA.18}$$

If we specialize equation (OA.5) such that labor income does not load on g or the granular residuals, we have $\phi_g = 0$, $\phi_\eta = 0$, $\psi_g = 0$, and $\psi_\eta = 0$. In this case, the risk premium expressions become in the limit as $N \to +\infty$):

$$\mathbb{E}[f] = A\mathcal{W}\left(1-\phi+\phi\psi_f\right)\sigma_f^2, \tag{OA.19}$$

$$\mathbb{E}[g] = A\mathcal{W}\left(\left(1-\phi\right)\zeta_m\right)\sigma_g^2, \tag{OA.20}$$

$$\mathbb{E}[r_\eta] = A\mathcal{W}\left(1-\phi\right)\sigma_\eta^2, \tag{OA.21}$$

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

and the market risk premium is still as given in equation (OA.18).

The following two propositions provide testable implications implied by the model, and we provide the related derivations in Online Appendix OA.3.

Proposition 2 (Attenuation of small-cap betas)
Let β_p^m and β_p^frepresent 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

$$\beta_S^m - \beta_S^f \;=\; -\left(\frac{\sigma_\epsilon^2}{\sigma_f^2+\sigma_\epsilon^2}\right)\beta_S, \tag{OA.23}$$

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

$$\beta_L^m - \beta_L^f \;=\; \frac{1}{\overline{\nu}}\left[\frac{\sigma_\epsilon^2}{\sigma_f^2+\sigma_\epsilon^2} - \beta_L\frac{\zeta_m^2\sigma_g^2 - \overline{\nu}\sigma_i^2}{\sigma_f^2+\sigma_\epsilon^2}\right]. \tag{OA.24}$$

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

Proposition 3 (Small-cap 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 that 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

$$\begin{aligned} \alpha_S^m &= \beta_S \mathbb{E}[f] \left[\frac{\phi\psi_f \sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi+\phi\psi_f)\sigma_\epsilon^2} \right] \\ &\overset{\phi\to 1}{\to} \frac{\sigma_\epsilon^2}{\sigma_f^2+\sigma_\epsilon^2} \beta_S \mathbb{E}[f] \text{ , and} \end{aligned} \tag{OA.25}$$

$$\alpha_S^f = 0. \tag{OA.26}$$

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

$$\begin{aligned} \alpha_\epsilon^m &= A\mathcal{W}\left(\frac{\sigma_f^2}{\sigma_f^2+\sigma_\epsilon^2}\right)\left((1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi)\zeta_m^2\sigma_g^2 + (1-\phi)\sigma_\eta^2\right) \\ &\quad -A\mathcal{W}(1-\phi+\phi\psi_f)\sigma_f^2 \\ &\overset{\phi\to 1}{\to} -A\mathcal{W}\left(\frac{\sigma_\epsilon^2}{\sigma_f^2+\sigma_\epsilon^2}\right)\psi_f\sigma_f^2 \text{ , and} \end{aligned} \tag{OA.27}$$

$$\alpha_\epsilon^f = A\mathcal{W}(1-\phi)\zeta_m^2\sigma_g^2 + A\mathcal{W}(1-\phi)\sigma_\eta^2 \overset{\phi\to 1}{\to} 0. \tag{OA.28}$$

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

Proposition 4 (Two factor representation and spanning) *Let $r_{smb} \equiv \frac{1}{N-K}\sum_{i=K+1}^N r_i - \frac{1}{K}\sum_{i=1}^K r_i$ be the return on a portfolio that is long small-cap stocks and short large-cap stocks, and suppose r_{smb} and r_m are linearly independent. For arbitrary ϕ, expected returns are linear in the bivariate betas on (r_m, r_{smb}) or (f, r_{smb}). In the limit as $\phi \to 1$, r_{smb} is spanned by f*

and unnecessary to include when f is used as the market factor; by contrast, r_{smb} is required as a pricing factor if r_m is used as the market factor. Furthermore, if $\beta_S > 1$ and $\phi \to 1$, $E[r_{smb}] \geq 0$.

Proposition 4 is quite straightforward to establish given results above. First, notice that, the formula in (OA.10) is consistent with a SDF which is a linear combination of f and ϵ, where the weight on $\epsilon \to 0$ as $\phi \to 1$. $r_{smb} = (\beta_S - \beta_L)f - \frac{1}{\bar{v}}\epsilon$. Given our linear independence assumption, we can rewrite f and ϵ as linear combinations of r_m and r_{smb}, so the SDF can also be written as an affine function of r_m and r_{smb}. From the theorems of section 6.3 of Cochrane (2009), it follows that there will be a linear relationship between expected returns on any asset and multivariate betas on r_m and r_{smb}. The last result holds since $\beta_S > 1$ implies that $\beta_L < 1$ (since the loading of the market return is normalized to one and equal to a weighted average of β_S and β_L), and the risk premium on f is weakly positive, while prediction 4 implies that the risk premium on ϵ approaches zero as $\phi \to 1$.

Suppose the return on asset i are given by $r_i = \beta_i f + \gamma_i \epsilon + u_i$. To derive equation (16), start with expressions for r_{smb} and r_m from above:

$$r_{smb} = \beta_{smb} f - \frac{\epsilon}{\bar{v}},$$
$$r_m = f + \epsilon.$$

From the second equation:

$$\epsilon = r_m - f.$$

Substitute ϵ into the first equation:

$$r_{smb} = \beta_{smb} f - \frac{r_m - f}{\bar{v}} = \beta_{smb} f - \frac{r_m}{\bar{v}} + \frac{f}{\bar{v}}.$$

Combine like terms:

$$r_{smb} + \frac{r_m}{\bar{v}} = f\left(\beta_{smb} + \frac{1}{\bar{v}}\right).$$

Solve for f:

$$f = \frac{r_m + r_{smb}\bar{v}}{\beta_{smb}\bar{v} + 1}.$$

Compute ϵ:

$$\epsilon = r_m - f = r_m - \frac{r_m + r_{smb}\bar{v}}{\beta_{smb}\bar{v} + 1} = \frac{\beta_{smb}\bar{v}r_m - r_{smb}\bar{v}}{\beta_{smb}\bar{v} + 1}.$$

Next, substitute f and ϵ into r_i.

$$r_i = \beta_i f + \gamma_i \epsilon + u_i = \beta_i \left(\frac{r_m + r_{smb}\bar{v}}{\beta_{smb}\bar{v} + 1} \right) + \gamma_i \left(\frac{\beta_{smb}\bar{v}r_m - r_{smb}\bar{v}}{\beta_{smb}\bar{v} + 1} \right) + u_i.$$

Simplifying this expression yields the result.

OA.3 Theoretical model: Derivation details

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:

$$\begin{aligned}
&- E[\exp(-AC_j)] \\
&= -E[\exp(-A[Y_j + W(1 + r_p(\theta^j))])] \\
&= -E[\exp(-A[Y + \varepsilon_j + W(1 + r_p(\theta^j))])] \\
&= -E\left[E[\exp(-A[Y + \varepsilon_j + W(1 + r_p(\theta^j))])|f, \eta, \xi]\right] \\
&= -E\left[\underbrace{\exp(-AY) \cdot E[\exp(-A\varepsilon_j)|f, \eta, \xi]}_{=\exp(-A\tilde{Y})} \cdot \exp(-AW(1 + r_p(\theta^j)))\right] \\
&= -E\left[\exp(-A[\tilde{Y} + W(1 + r_p(\theta^j))])\right]
\end{aligned} \tag{OA.29}$$

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 $r_p(\theta^j)$:

$$\begin{aligned}\tilde{Y} + W(1 + r_p(\theta^j)) &= \underbrace{\left(E\left[\tilde{Y}\right] + W\right)}_{\equiv \mathcal{W}} \times \left(\underbrace{\frac{E\left[\tilde{Y}\right]}{E\left[\tilde{Y}\right] + W}}_{\equiv \phi} \times \underbrace{\frac{\tilde{Y}}{E\left[\tilde{Y}\right]}}_{\equiv 1 + r_h} + \underbrace{\frac{W}{E\left[\tilde{Y}\right] + W}}_{\equiv 1 - \phi} \times (1 + r_p(\theta^j)) \right) \\ &= \mathcal{W}\left(\phi(1 + r_h) + (1 - \phi)(1 + r_p(\theta^j))\right)\end{aligned} \tag{OA.30}$$

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

$$\begin{aligned}&- E\left[\exp(-A[\tilde{Y} + W(1 + r_p(\theta^j))])\right] \\ &= -E\left[\exp(-A\mathcal{W}[\phi(1 + r_h) + (1 - \phi)(1 + r_p(\theta^j))])\right] \\ &= -\exp\left(-A\mathcal{W}(1 - \phi)\left(1 + E[r_p(\theta^j)] - \frac{A\mathcal{W}}{2}(1 - \phi)Var(r_p(\theta^j)) + A\mathcal{W}\phi Cov(r_h, r_p(\theta^j))\right)\right) \dots \\ &\quad \cdots \times \exp\left(-A\mathcal{W}\phi(1 + E[r_h]) + \frac{A^2\mathcal{W}^2}{2}\phi^2 Var(r_h)\right)\end{aligned} \tag{OA.31}$$

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.9) of the main text:

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

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

$$\begin{aligned} r_p(\theta^j) &= \sum_{i=1}^{N} \theta_i^j r_i + \left[1 - \sum_{i=1}^{N} \theta_i^j\right] r_f \\ &= r_f + \sum_{i=1}^{N} \theta_i^j (r_i - r_f) \\ &= r_f + \sum_{i=1}^{N} \theta_i^j (E[r_i] - r_f) + \sum_{i=1}^{N} \theta_i^j (r_i - E[r_i]) \\ &= E[r_p(\theta^j)] + \sum_{i=1}^{N} \theta_i^j \left(\beta_i (f - E[f]) + \zeta_i (g - E[g]) + \eta_i\right) \\ &= E[r_p(\theta^j)] + \left(\sum_{i=1}^{N} \theta_i^j \beta_i\right) \cdot (f - E[f]) + \left(\sum_{i=1}^{N} \theta_i^j \zeta_i\right) \cdot (g - E[g]) + \left(\sum_{i=1}^{N} \theta_i^j \eta_i\right) \\ &\equiv E[r_p(\theta^j)] + \beta_p(\theta^j)(f - E[f]) + \zeta_p(\theta^j)(g - E[g]) + \eta_p(\theta^j) \end{aligned} \tag{OA.32}$$

The variance of the realized return is given by

$$Var(r_p(\theta^j)) = \left[\beta_p(\theta^j)\right]^2 \sigma_f^2 + \left[\zeta_p(\theta^j)\right]^2 \sigma_g^2 + \sum_{i=1}^{K} \left[\theta_i^j\right]^2 \sigma_i^2 \tag{OA.33}$$

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

$$Cov(r_p(\theta^j), r_h) = \begin{cases} \beta_p(\theta^j)\psi_f \sigma_f^2 + \zeta_p(\theta^j)\psi_g \sigma_g^2 + \frac{\overline{v}}{K}\psi_\eta \sum_{i=1}^{K} \theta_i^j \sigma_i^2 & \text{if } i = 1, \ldots, K \\ \beta_p(\theta^j)\psi_f \sigma_f^2 + \zeta_p(\theta^j)\psi_g \sigma_g^2 & \text{if } i = K+1, \ldots, N \end{cases} \tag{OA.34}$$

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

$$\frac{\partial}{\partial \theta_i^j} E[r_p(\theta^j)] = E[r_i] - r_f \tag{OA.35}$$

$$\frac{\partial}{\partial \theta_i^j} Var(r_p(\theta^j)) = 2\left[\beta_p(\theta^j)\beta_i \sigma_f^2 + \zeta_p(\theta^j)\zeta_i \sigma_g^2 + \theta_i^j \sigma_i^2\right] \tag{OA.36}$$

$$\frac{\partial}{\partial \theta_i^j} Cov(r_p(\theta^j), r_h) = \begin{cases} \beta_i \psi_f \sigma_f^2 + \zeta_i \psi_g \sigma_g^2 + \frac{\overline{v}}{K} \psi_\eta \sigma_i^2 & \text{for } i = 1, \ldots, K \\ \beta_i \psi_f \sigma_f^2 + \zeta_i \psi_g \sigma_g^2 & \text{for } i = K+1, \ldots, N \end{cases} \tag{OA.37}$$

For large stocks $i = 1, \ldots, K$, the first-order condition for θ_i^j in the portfolio choice problem (OA.9) is

$$\begin{aligned} 0 &= \frac{\partial}{\partial \theta_i^j} \left[E[r_p(\theta^j)] - A\mathcal{W} \left(\frac{1-\phi}{2} Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j)) \right) \right] \\ &= E[r_i] - r_f - A\mathcal{W} \left[(1-\phi) \left(\beta_p(\theta^j) \beta_i \sigma_f^2 + \zeta_p(\theta^j) \zeta_i \sigma_g^2 + \theta_i^j \sigma_i^2 \right) + \phi \left(\beta_i \psi_f \sigma_f^2 + \zeta_i \psi_g \sigma_g^2 + \frac{\overline{v}}{K} \psi_\eta \sigma_i^2 \right) \right] \\ &= E[r_i] - r_f - A\mathcal{W} \left[(1-\phi) Cov(r_i, r_p(\theta^j)) + \phi Cov(r_i, r_h) \right] \\ &= E[r_i] - r_f - A\mathcal{W} \left[\beta_i \left((1-\phi)\beta_p(\theta^j) + \phi \psi_f \right) \sigma_f^2 + \zeta_i \left((1-\phi)\zeta_p(\theta^j) + \phi \psi_g \right) \sigma_g^2 + \left((1-\phi)\theta_i^j + \phi \frac{\overline{v}}{K} \psi_\eta \right) \sigma_i^2 \right] \end{aligned} \tag{OA.38}$$

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

$$\begin{aligned} 0 &= \frac{\partial}{\partial \theta_i^j} \left[E[r_p(\theta^j)] - A\mathcal{W} \left(\frac{1-\phi}{2} Var(r_p(\theta^j)) + \phi Cov(r_h, r_p(\theta^j)) \right) \right] \\ &= E[r_i] - r_f - A\mathcal{W} \left[(1-\phi) \left(\beta_p(\theta^j) \beta_i \sigma_f^2 + \zeta_p(\theta^j) \zeta_i \sigma_g^2 + \theta_i^j \sigma_i^2 \right) + \phi \left(\beta_i \psi_f \sigma_f^2 + \zeta_i \psi_g \sigma_g^2 \right) \right] \\ &= E[r_i] - r_f - A\mathcal{W} \left[(1-\phi) Cov(r_i, r_p(\theta^j)) + \phi Cov(r_i, r_h) \right] \\ &= E[r_i] - r_f - A\mathcal{W} \left[\beta_i \left((1-\phi)\beta_p(\theta^j) + \phi \psi_f \right) \sigma_f^2 + \zeta_i \left((1-\phi)\zeta_p(\theta^j) + \phi \psi_g \right) \sigma_g^2 + (1-\phi)\theta_i^j \sigma_i^2 \right] \end{aligned} \tag{OA.39}$$

Therefore, the first-order condition (OA.10) holds for all stocks. To obtain the equilibrium risk-return relationship stated in equation (OA.11), we plug the equilibrium portfolio shares $\theta_1^{j*} = \cdots = \theta_K^{j*} = \overline{v}/K$ and $\theta_{K+1}^{j*} = \cdots = \theta_N^{j*} = (1-\overline{v})/(N-K)$ into the above first-order conditions. The normalization (OA.2) implies that $\beta(\theta^{j*}) = 1$, and equation (OA.11) then follows immediately.

The proof of Proposition 1 is straightforward. The expected return on large stocks $i = 1, \ldots, K$ does not change as N increases, so we only need to prove that the expected excess return on each small stock $i = K+1, K+2, \ldots$ is given by $E[r_i] - r_f = \beta_i E[f]$.

Taking limits in the expression from equation (OA.16) directly gives the desired result:

$$E[r_i] - r_f = \beta_i E[f] + A\mathcal{W}(1-\phi) \underbrace{\frac{1-\overline{v}}{N-K}\sigma_i^2}_{\to 0 \text{ as } N\to+\infty} \to \beta_i E[f] \tag{OA.40}$$

To get the result for the market risk premium, we combine equation (OA.15) 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.3.1 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 that of all small-cap stocks in our model economy. Equation (OA.1) implies that the return on a value-weighted portfolio of all large-cap stocks in the economy is given by

$$r_L - \mathbb{E}[r_L] \;=\; \beta_L (f - \mathbb{E}[f]) + \zeta_L (g - \mathbb{E}[g]) + \eta_L, \tag{OA.41}$$

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, the return on a value-weighted portfolio of all small-cap stocks in the economy is given by

$$r_S - \mathbb{E}[r_S] \;=\; \beta_S (f - \mathbb{E}[f]) + \eta_S, \tag{OA.42}$$

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

$$\begin{aligned}
\beta_L^m &= \frac{\mathbb{COV}\left(r_L - \mathbb{E}\left[r_L\right], r_m - \mathbb{E}\left[r_m\right]\right)}{\mathbb{VAR}\left(r_m - \mathbb{E}\left[r_m\right]\right)} \\
&= \frac{\mathbb{COV}\left(\beta_L f + \zeta_L g + \eta_L, f + \zeta_m g + \eta\right)}{\mathbb{VAR}\left(f + \zeta_m g + \eta\right)} \\
&= \frac{\mathbb{COV}\left(\beta_L f + \frac{\zeta_m}{\overline{\nu}} g + \frac{1}{K}\sum_{i=1}^{K}\eta_i, f + \zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i + \frac{1-\overline{\nu}}{N-K}\sum_{i=K+1}^{N}\eta_i\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]} \\
&= \frac{\beta_L\sigma_f^2 + \frac{\zeta_m^2}{\overline{\nu}}\sigma_g^2 + \frac{\overline{\nu}}{K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}.
\end{aligned} \tag{OA.43}$$

The small-cap portfolio beta with respect to the value-weighted market return in equation (OA.4) is given by

$$\begin{aligned}
\beta_S^m &= \frac{\mathbb{COV}\left(r_S - \mathbb{E}\left[r_S\right], r_m - \mathbb{E}\left[r_m\right]\right)}{\mathbb{VAR}\left(r_m - \mathbb{E}\left[r_m\right]\right)} \\
&= \frac{\mathbb{COV}\left(\beta_S f + \eta_S, f + \zeta_m g + \eta\right)}{\mathbb{VAR}\left(f + \zeta_m g + \eta\right)} \\
&= \frac{\mathbb{COV}\left(\beta_S f + \frac{1}{N-K}\sum_{i=K+1}^{N}\eta_i, f + \zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i + \frac{1-\overline{\nu}}{N-K}\sum_{i=K+1}^{N}\eta_i\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]} \\
&= \frac{\beta_S\sigma_f^2 + \frac{1-\overline{\nu}}{N-K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \sigma_i^2\left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K}\right]}.
\end{aligned} \tag{OA.44}$$

It is straightforward to show that the betas of the large- and small-cap portfolios 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

$$\begin{aligned}
\beta_L^m - \beta_L^f &= \frac{\beta_L \sigma_f^2 + \zeta_m^2 \sigma_g^2 + \frac{\overline{\nu}}{K} \sigma_i^2}{\sigma_f^2 + \zeta_m^2 \sigma_g^2 + \sigma_i^2 \left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K} \right]} - \beta_L \\
&= \frac{(1 - \beta_L) \zeta_m^2 \sigma_g^2 + \left(\frac{\overline{\nu}}{K} - \beta_L \right) \overline{\nu} \sigma_i^2}{\sigma_f^2 + \zeta_m^2 \sigma_g^2 + \sigma_i^2 \left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K} \right]} \frac{1}{\overline{\nu}} \\
&= \frac{\zeta_m^2 \sigma_g^2 + \frac{\overline{\nu}^2}{K} \sigma_i^2 - \beta_L \left(\zeta_m^2 \sigma_g^2 - \overline{\nu} \sigma_i^2 \right)}{\sigma_f^2 + \zeta_m^2 \sigma_g^2 + \sigma_i^2 \left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K} \right]} \frac{1}{\overline{\nu}}.
\end{aligned} \tag{OA.45}$$

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

$$\lim_{N \to \infty} \beta_L^m - \beta_L^f = \frac{1}{\overline{\nu}} \left[\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2} - \beta_L \frac{\zeta_m^2 \sigma_g^2 - \overline{\nu} \sigma_i^2}{\sigma_f^2 + \sigma_\epsilon^2} \right] \tag{OA.46}$$

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

$$\begin{aligned}
\beta_S^m - \beta_S^f &= \frac{\beta_S \sigma_f^2 + \frac{1-\overline{\nu}}{N-K} \sigma_i^2}{\sigma_f^2 + \zeta_m^2 \sigma_g^2 + \sigma_i^2 \left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K} \right]} - \beta_S \\
&= \frac{\left(1 - \beta_S \left(1 - \overline{\nu}\right)\right) \frac{1-\overline{\nu}}{N-K} \sigma_i^2 - \beta_S \left(\sigma_i^2 \frac{\overline{\nu}^2}{K} + \zeta_m^2 \sigma_g^2 \right)}{\sigma_f^2 + \zeta_m^2 \sigma_g^2 + \sigma_i^2 \left[\frac{\overline{\nu}^2}{K} + \frac{(1-\overline{\nu})^2}{N-K} \right]}.
\end{aligned} \tag{OA.47}$$

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

$$\begin{aligned}
\lim_{N \to \infty} \beta_S^m - \beta_S^f &= -\left(\frac{\zeta_m^2 \sigma_g^2 + \sigma_i^2 \frac{\overline{\nu}^2}{K}}{\sigma_f^2 + \zeta_m^2 \sigma_g^2 + \sigma_i^2 \frac{\overline{\nu}^2}{K}} \right) \beta_S \\
&\equiv -\left(\frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2} \right) \beta_S,
\end{aligned} \tag{OA.48}$$

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

In this subsection, we derive the two-factor alphas of the small-cap portfolios and the IFF (i.e., $\epsilon \equiv \zeta_m g + r_\eta$) when either the value-weighted market factor, $r_m - r_f$, or f is used as the level factor, with g included as the second factor. 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 the expected returns on each asset under the two-factor model when either $r_m - r_f$ or f is used as the level factor. The difference between the true expected excess return and the model-based expected excess return is the implied alpha under each factor model.

Equation (OA.22) implies the expected excess return on the small-cap portfolio is

$$\mathbb{E}\left[r_S\right] - r_f \;=\; \beta_S \mathbb{E}\left[f\right]. \tag{OA.49}$$

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

$$\mathbb{E}\left[\epsilon\right] \;=\; \zeta_m \mathbb{E}\left[g\right] + \mathbb{E}\left[r_\eta\right] \tag{OA.50}$$

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

$$\begin{aligned} \alpha_S^f &= \mathbb{E}\left[r_S\right] - r_f - \beta_S\left(\mathbb{E}\left[f\right]\right) \\ &= \beta_S \mathbb{E}\left[f\right] - \beta_S \mathbb{E}\left[f\right] \\ &= 0 \end{aligned} \tag{OA.51}$$

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.44) 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.15). Making use of the expressions for $\mathbb{E}[f]$, $\mathbb{E}[g]$, and $\mathbb{E}[r_\eta]$ in equations (OA.19)-(OA.21) 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.19)-(OA.21) and noting the definition $\sigma_\epsilon^2 \equiv \zeta_m^2 \sigma_g^2 + \sigma_\eta^2$, and simplifying:

$$\begin{aligned}\alpha_S^m &= \beta_S \mathbb{E}[f] \left[1 - \frac{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi)\zeta_m^2\sigma_g^2 + (1-\phi)\sigma_\eta^2}{(1-\phi+\phi\psi_f)\sigma_f^2 + \left((1-\phi)\zeta_m^2\sigma_g^2 + (1-\phi)\sigma_\eta^2\right) \frac{\left(1-\phi+\phi\psi_f\right)}{(1-\phi)}} \right] \\ &= \beta_S \mathbb{E}[f] \left[1 - \frac{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi)\sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi)\sigma_\epsilon^2 \frac{\left(1-\phi+\phi\psi_f\right)}{(1-\phi)}} \right] \\ &= \beta_S \mathbb{E}[f] \left[1 - \frac{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi)\sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi+\phi\psi_f)\sigma_\epsilon^2} \right] \\ &= \beta_S \mathbb{E}[f] \left[\frac{\phi\psi_f\sigma_\epsilon^2}{(1-\phi+\phi\psi_f)\sigma_f^2 + (1-\phi+\phi\psi_f)\sigma_\epsilon^2} \right] > 0. \qquad \text{(OA.52)}\end{aligned}$$

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

$$\lim_{\phi \to 1} \alpha_S^m = \frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2} \beta_S \mathbb{E}[f] > 0. \tag{OA.53}$$

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.48) (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

$$\begin{aligned} \beta_\epsilon^f &= \frac{\mathbb{COV}(\epsilon, f)}{\mathbb{VAR}(f)} \\ &= \frac{\mathbb{COV}(\zeta_m g + r_\eta, f)}{\mathbb{VAR}(f)} \\ &= 0. \end{aligned} \tag{OA.54}$$

Similarly, the IFF exposure to g is given by

$$\begin{aligned} \beta_\epsilon^g &= \frac{\mathbb{COV}(\epsilon, g)}{\mathbb{VAR}(g)} \\ &= \frac{\mathbb{COV}(\zeta_m g + r_\eta, g)}{\mathbb{VAR}(g)} \\ &= \zeta_m. \end{aligned} \tag{OA.55}$$

Note that since $f \perp g$, the univariate and multivariate betas are equivalent. When r_m is used as the market factor, the measured IFF exposures to r_m and g must be calculated jointly, since r_m is correlated with g. Alternatively, the two-factor model will have the same pricing implications if we instead use a version of g that is orthogonalized with respect to r_m. Since this significantly simplifies the beta calculations, we follow this approach. The market beta

is given by

$$\begin{aligned}
\beta_{\epsilon}^{m} &= \frac{\mathbb{COV}\left(\epsilon, r_m\right)}{\mathbb{VAR}\left(r_m\right)} \\
&= \frac{\mathbb{COV}\left(\zeta_m g + r_\eta, r_m\right)}{\mathbb{VAR}\left(r_m\right)} \\
&= \frac{\mathbb{COV}\left(\zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i, f + \zeta_m g + \frac{\overline{\nu}}{K}\sum_{i=1}^{K}\eta_i\right)}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2} \\
&= \frac{\zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2}{\sigma_f^2 + \zeta_m^2\sigma_g^2 + \frac{\overline{\nu}^2}{K}\sigma_i^2} \\
&= \frac{\sigma_\epsilon^2}{\sigma_f^2 + \sigma_\epsilon^2}
\end{aligned} \tag{OA.56}$$

The orthogonalized version of g is given by

$$g^{\perp} = \left(1 - \zeta_m\right) g, \tag{OA.57}$$

and the beta is

$$\beta_{\epsilon}^{g,\perp} = \zeta_m\left(1 - \zeta_m\right). \tag{OA.58}$$

Given the expression for the expected return on ϵ in equation (OA.50), the expression for β_{ϵ}^{f} in equation (OA.54), the expression for the expected return on f in equation (OA.19), the expression for β_{ϵ}^{g} in equation (OA.55), and the expression for the expected return on g in equation (OA.19), the alpha implied when f is used as the market factor is

$$\begin{aligned}
\alpha_{\epsilon}^{f} &= \mathbb{E}\left[\epsilon\right] - \beta_{\epsilon}^{f}\mathbb{E}\left[f\right] - \beta_{\epsilon}^{g}\mathbb{E}\left[g\right] \\
&= \zeta_m\mathbb{E}\left[g\right] + \mathbb{E}\left[r_\eta\right] - \zeta_m\mathbb{E}\left[g\right] \\
&= A\mathcal{W}\left(1 - \phi\right)\sigma_\eta^2 > 0
\end{aligned} \tag{OA.59}$$

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

$$\lim_{\phi \to 1} \alpha_\epsilon^f = 0. \tag{OA.60}$$

Given the expression for the expected return on ϵ in equation (OA.50), the expression for β_ϵ^m in equation (OA.56), the expression for the expected excess return on r_m in equation (OA.15), the expression for $\beta_\epsilon^{g,\perp}$ in equation (OA.58), and the expression for the expected return on g in equation (OA.19), the alpha implied when $r_m - r_f$ is used as the market factor is

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

Making use of the expressions for $\mathbb{E}[f]$, $\mathbb{E}[g]$, and $\mathbb{E}[r_\eta]$ in equations (OA.19)-(OA.21) and simplifying,

$$\begin{aligned}
\alpha_\epsilon^m &= A\mathcal{W} \left(\frac{\sigma_f^2}{\sigma_f^2 + \sigma_\epsilon^2}\right) \left((1 - \phi + \phi\psi_f) \sigma_f^2 + (1 - \phi) \zeta_m^2 \sigma_g^2 + (1 - \phi) \sigma_\eta^2\right) \\
&\quad - A\mathcal{W} \left[(1 - \phi + \phi\psi_f) \sigma_f^2 + (1 - \phi) \zeta_m^2 (1 - \zeta_m)^2 \sigma_g^2\right].
\end{aligned} \tag{OA.61}$$

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

$$\lim_{\phi \to 1} \alpha_{\epsilon}^{m} = -A\mathcal{W}\left(\frac{\sigma_{\epsilon}^{2}}{\sigma_{f}^{2}+\sigma_{\epsilon}^{2}}\right)\psi_{f}\sigma_{f}^{2} < 0. \tag{OA.62}$$

OA.4 Evidence on the share of nonpublic-equity market wealth

According to the intuition developed in our theoretical model from Section (OA.2), we expect the IFF 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 (given our assumptions), the IFF'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. In this section, we highlight some evidence from the US economy that suggests that 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 US economy. For example, using US Census data, Davis et al. (2006) estimate that public firms comprised only approximately 30% of total employment in the US over their sample. For many investors, news about expected labor and/or entrepreneurial income (and the 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 approximately one fifth of its 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 that 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 firm net payout has never exceeded 20% of US personal consumption expenditures.

Next, as we show in Figure 2, the equity market industry distribution of large-cap stocks (here defined as those in the top 2.5% of the market value distribution using NYSE market cap breakpoints) is not representative of the industry distribution of the rest of the equity market, and its variation drives much of that in the overall equity market. Furthermore, the industry distribution within the equity market is not representative of that within the broader economy (as measured using employment shares in Figure 2). Therefore, any variation in the VMF's return that is driven by such compositional changes (i.e., via the IFF) is likely priced differently by the marginal utility of the representative investor than the broader market risk reflected in the DMF.

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.2) in which changes in the value of public equity 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 (2024), 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 large-cap stocks), it is likely that any IFF generated by these large-cap stocks has a low price of risk that differs substantially from that of the dominant market factor, as in our theoretical model, a prediction for which we find empirical evidence.

OA.5 PCA: Additional details

In this section, we discuss some additional details related to our PCA described in Section 4.1 and the use of PCA more generally as it applies to the study of financial economics.

OA.5.1 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, 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.[28] 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, the 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.[29]

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 use a broad cross-section of portfolios. 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 asset pricing tests (as per standard convention). Third, we would

[28]Some notable recent extensions of the basic PCA framework include Kelly et al. (2019), who develop an instrumented PCA that helps provide a link between factors constructed from characteristics to those constructed from PCs. Lettau and Pelger (2020) 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 nontraded factors.

[29]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 that use estimated PCs have the same asymptotic properties as if the true unobserved factors were used instead. This property holds when $\frac{T}{N} \to 0$ as $N, T \to \infty$. The sample sizes that are used to construct estimates of factors and loadings usually involve a number of stocks that is an order of magnitude larger than the number of time periods, so this condition approximately holds. Even when this condition fails, the 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.

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 Section 6.1. Note that these are a superset of the equity portfolios used for the PCA analysis in Giglio and Xiu (2021).

Consistent with our interpretation of the first PC as a level factor, Figure OA.1 plots a histogram of the portfolio weights of this PC on each of our 372 equal-weighted portfolios used in the PCA. Each portfolio weight is positive, and the distribution is relatively tight. Next, we provide evidence in Figure OA.2 that the PCs approximately conform to the assumptions about loadings on f_{t+1} and g_{t+1} in equations (7) and (8). Namely, the value-weighted average loading of the value-weighted portfolios on the DMF is 0.993, and the value-weighted average loading of these portfolios on each higher PC is approximately zero.

OA.5.2 Estimating the DMF and IFF: 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 (7). In other words, we run the following regression

$$VMF_{t+1} = a + b \cdot PC_{1,t+1} + \varepsilon_{VMF,t+1}, \tag{OA.63}$$

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

$$DMF_{t+1} \equiv b \cdot PC_{1,t+1}. \tag{OA.64}$$

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

$$IFF_{t+1} \equiv VMF_{t+1} - DMF_{t+1}, \tag{OA.65}$$

which is consistent with the definition of the IFF in equation (9).

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}.[30] 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 (9). 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 OA.8.

OA.6 The DMF and the equal-weighted market

If we have a balanced panel of asset returns ($N_t = N$ for all t), then our assumptions outlined in Section 3.2.1 imply that returns on an equal-weighted portfolio of assets satisfy

$$r_{ew,t+1} = \frac{1}{N}\sum_{i=1}^{N}\left(\beta_i \cdot f_{t+1} + \zeta_i' \cdot g_{t+1} + \eta_{i,t+1}\right) = \beta_{ew} \cdot f_{t+1} + 0 \cdot g_{t+1} + \frac{1}{N}\sum_{i=1}^{N}\eta_{i,t+1}, \tag{OA.66}$$

where $\beta_{ew} \equiv \frac{1}{N}\sum_{i=1}^{N}\beta_i$. The factor loadings here are time invariant as in the single-stock case due the somewhat restrictive assumption here that the composition of stocks does not change over time. Because $E_t[\eta_{i,t+1}] = 0$, $r_{ew,t+1}$ is an *unbiased* estimator of f_{t+1}. Because 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 holds almost exactly, and $r_{ew,t+1}$ is also *consistent* for f_{t+1} up to a constant of proportionality, i.e.,

$$\operatorname*{plim}_{N\to\infty} r_{ew,t+1} = \beta_{ew} f_{t+1} \iff \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\eta_{i,t+1} = 0, \tag{OA.67}$$

in which case $r_{ew,t+1}$ reveals a scaled version of f_{t+1}. Thus, we use the equal-weighted market portfolio to help identify an alternate version of the DMF in robustness tests.

[30] 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.5.1 and in Online Appendix Figure OA.2.

OA.7 The VMF is priced by the DMF

Here, we test an additional implication of our model related to Prediction 4 in Section 3.1, which is that the DMF prices the VMF. When the wealth share of nonfinancial assets is high (i.e., $\phi \approx 1$), the risk premium on ϵ is small, and f should price r_m.

Table OA.2 provides results from regressions of the VMF on the DMF and higher PCs from the equal-weighted portfolio PCA. Consistent with the above implication, the DMF prices the VMF, yielding a statistically and economically insignificant annualized alpha of -0.53%. 84 percent of the variance in the VMF is explained by its projection on the DMF (where the R-squared value of 0.84 equals the ratio of the DMF's variance to the VMF's variance). Furthermore, adding more PCs to the regression does not change the alpha appreciably and only marginally increases the R-squared values. Although adding PCs helps explain more variation in the VMF, either the PCs themselves are not priced, as is the case for PCs 2 and 4 (i.e., their Sharpe ratios are not different from zero), the VMF loads negatively on a PC with a positive risk price (PC 3), or the PC does not contribute to the VMF risk premium in a statistically significant manner (PC 5). These results conform with our hypothesis that the main source of risk in the VMF for which investors demand a positive risk premium is captured by the DMF.

OA.8 The VMF has a stable loading on the first PC but time-varying loadings on higher PCs

While our theoretical model described in Section 3.1 is static, an additional source of VMF instability appears in the dynamic setting described in Section 3.2 due to the VMF's time-varying composition, which may cause time-varying tilts away from the DMF. This potential instability shows up as time-varying loadings, $\zeta_{m,t}$, on g in equation (6), which primarily reflect loadings on large-cap stocks (since the DMF is constructed from equal-weighted portfolios) that we argue are disconnected from the broader economy.

To explore this possibility, we plot coefficient estimates from rolling regressions of the VMF on the first five PCs constructed using equal-weighted portfolios in Figure OA.3. Co-

efficients on the first PC are relatively stable, implying that time variation in the $\beta_{m,t} - 1$ from equation (9) is unlikely to contribute much to variation in the IFF. Indeed, our analysis in Online Appendix OA.9 indicates that allowing the VMF loading on the first PC to be time-varying generates an IFF with similar volatility as our baseline IFF. However, the VMF generates 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. Thus, paired with evidence from Section 4.3 that the granular residual is a small component of the IFF, most IFF variation is driven by time-varying loadings of the VMF on higher factors (i.e., the second IFF component described in equation (9)).

OA.9 How large is the granular residual?

In this section, we estimate the granular residual (GR) "size" relative to the VMF and the IFF. Recall that, according to equation (9), the GR is only one component of the IFF. Specifically, we estimate the ratios of the GR variance and standard deviation to those from the VMF and the IFF.

Equation (9) 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 (f_{t+1}), 2) time-varying loadings on other factors (g_{t+1}), and 3) the GR (η_{t+1}). Since, according to our maintained assumptions, f_{t+1} is spanned by the first PC, and g_{t+1} are spanned by the remaining PCs, we can estimate η_{t+1} by projecting the IFF onto the 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 that the 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 that the 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.[31]

Table OA.3 provides the GR summary statistics for the full-sample analysis (Panel A)

[31] 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 $t + 29$.

and the rolling analysis (Panel B). Focusing 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 estimated 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 15 PCs indicate that the GR standard deviation is 2.21% (annualized) and that its variance (standard deviation) is approximately 2.08% (13.23%) of the VMF variance (standard deviation) and 14.42% (36.38%) 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 f_{t+1}. The fact that $\sigma^2_{GR}/\sigma^2_{IFF} < 100\%$ in this case (it is 89.63%) implies that the IFF has some time-varying loading on 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 we again limit the maximum number of PCs in the regressions to 15. Results from the rolling regressions with 15 PCs indicate that the GR standard deviation is approximately 1.61% (annualized) and that its variance (standard deviation) is approximately 1.10% (7.02%) of the VMF variance (standard deviation) and 10.51% (26.50%) of the IFF variance (standard deviation).

OA.10 Impulse responses: Additional details

Suppose the DMF or IFF goes up by one standard deviation unit. By how much would an econometrician revise her forecast about future macroeconomic conditions or returns on other asset classes on the basis of that information? To answer this question, we follow the method in Schmidt (2024), which is itself related to Jordà (2005).[32] While we refer the

[32]The identifying assumptions proposed in Schmidt (2024) rely on correct specification of the mean of returns and allow the conditional expectation of the macro variable Y_{t+h} to potentially be misspecified. This

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

$$Y_{t+h} = \alpha_h + \beta_h z_{t+1} + \gamma_h Y_t + \phi_h(L) Y_t + \epsilon_{t+h} \tag{OA.68}$$

where, in our context, z_{t+1} is the market return component of interest (i.e., the DMF or the IFF), $\phi(L)$ is a lag polynomial, and ϵ_{t+h} is the forecast error. The LPIR is simply $\{\beta_h\}_{h=1}^{H}$.

OA.11 GARCH models: Additional details

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

is advantageous given that returns are much closer to random walks than many of the macro variables in the consideration set.

OA.11.1 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 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

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 u_t^2 + \alpha_2 \sigma_t^2 + \alpha_3 I_t u_t^2, \tag{OA.69}$$

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

$$\ln(\sigma_{t+1}^2) = \alpha_0 + \alpha_1 \left(|\widetilde{u}_t| - E\,|\widetilde{u}_t|\right) + \alpha_2 \ln(\sigma_t^2) + \alpha_3 \widetilde{u}_t, \tag{OA.70}$$

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).[33]

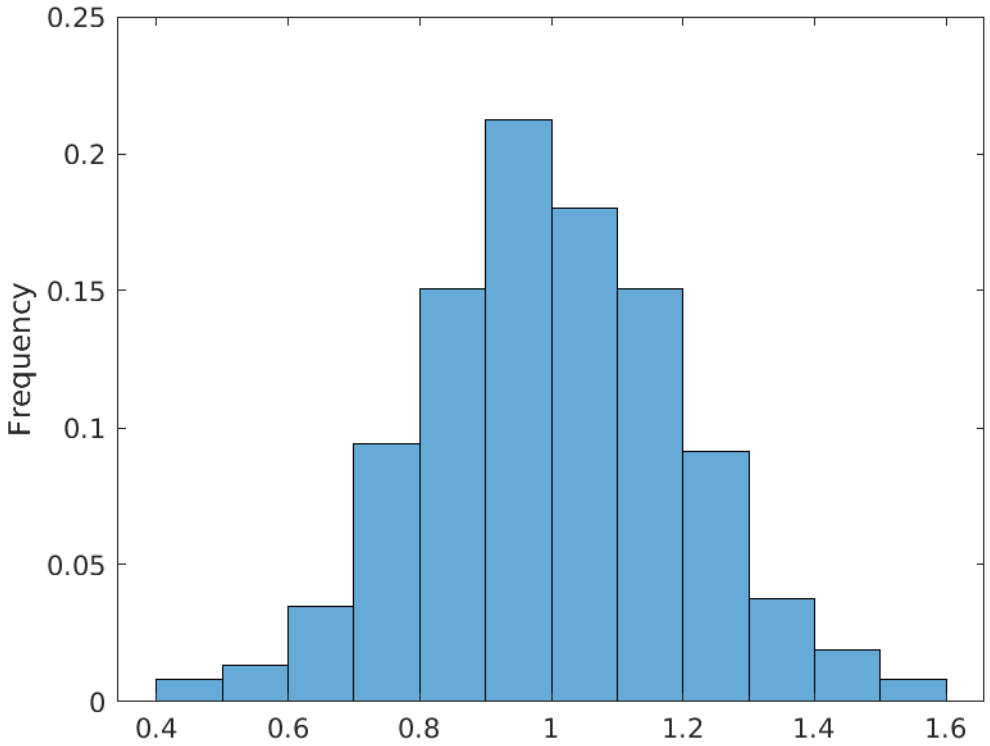
OA.12 Bootstrap procedure for estimating standard errors and p-values

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 and then recompute the PCA, the DMF, and the IFF within each simulated sample). We then sample

[33]The recognition of the leverage effect goes back to Black (1976), who noted 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.

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 the testing criterion in which we are interested (as a fraction of the total number of simulations).

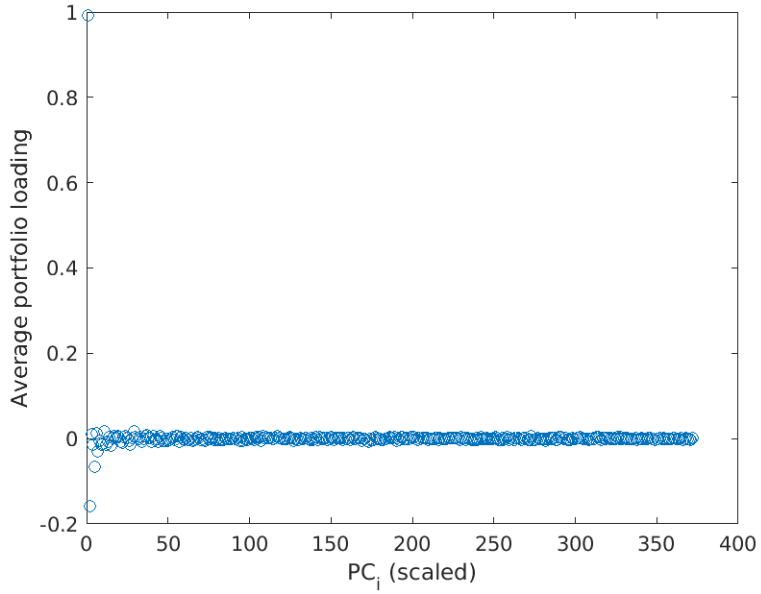
Figure OA.1: PC1 portfolio weights



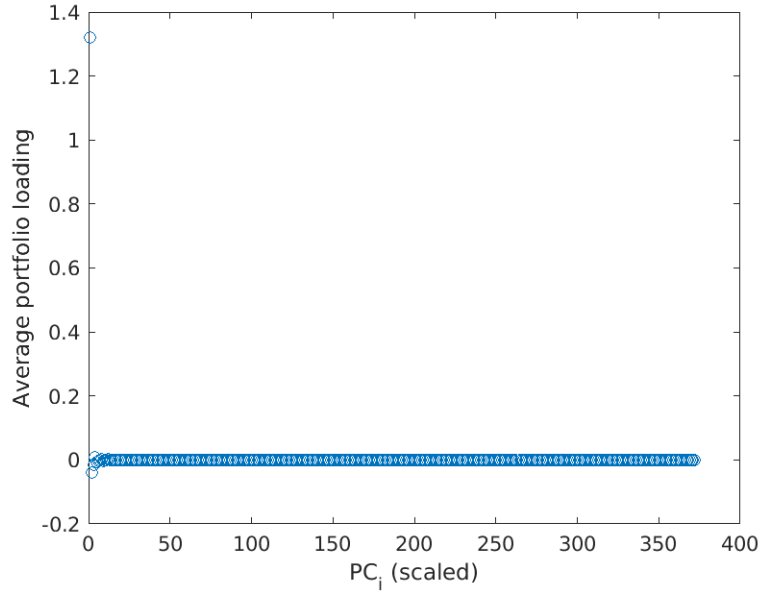
Note. This figure plots a histogram of the portfolio weights PC1 places on each of the 372 portfolios we use as test assets (described in Online Appendix Table OA.1). The weights are scaled such that the average weight is equal to 1. All data are from 1963–2021.

Figure OA.2: Average portfolio loadings on PCs

Panel 1: Value-weighted portfolios

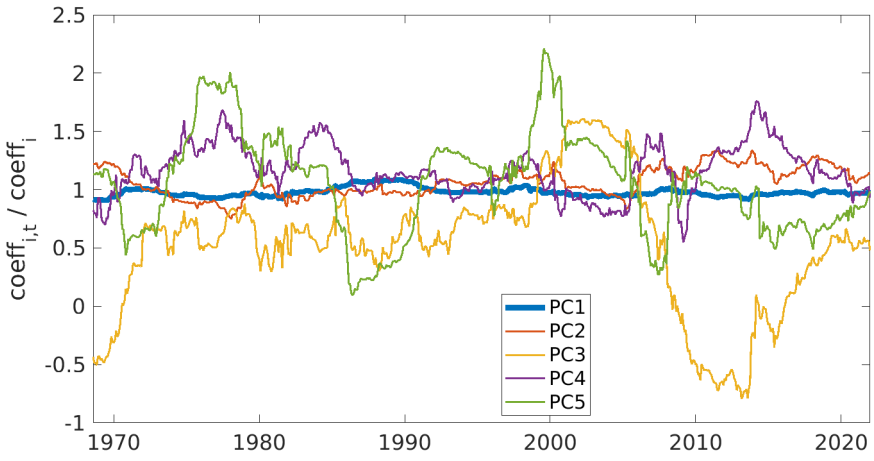


Panel 2: Equal-weighted portfolios



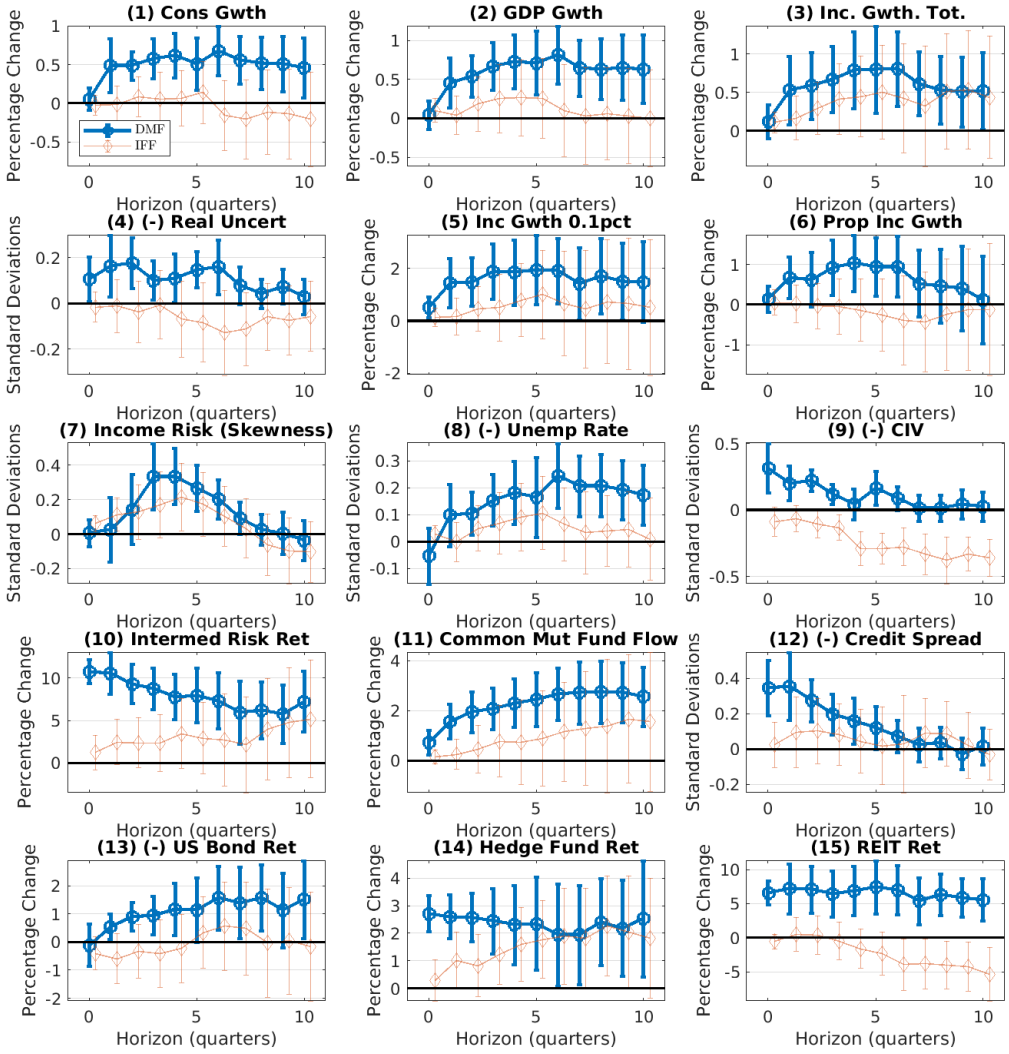
Note. This figure plots the average loadings of all 372 portfolios we use as test assets (described in Online Appendix Table OA.1 on scaled PCs. Panel 1 (Panel 2) presents results using value-weighted (equal-weighted) portfolios. The first PC is simply the DMF. Higher PCs are scaled to have the same volatility as the VMF. The average loading of portfolios on the first (scaled) PC is computed according to equation (7), where the weights w_i are based on the average market capitalizations of each portfolio over the entire sample (Panel 1) or equal weights (Panel 2). Average portfolio loadings on higher (scaled) PCs are computed as equal-weighted averages according to equation (8). All data are from 1963–2021.

Figure OA.3: Rolling coefficients of the VMF on the first five PCs



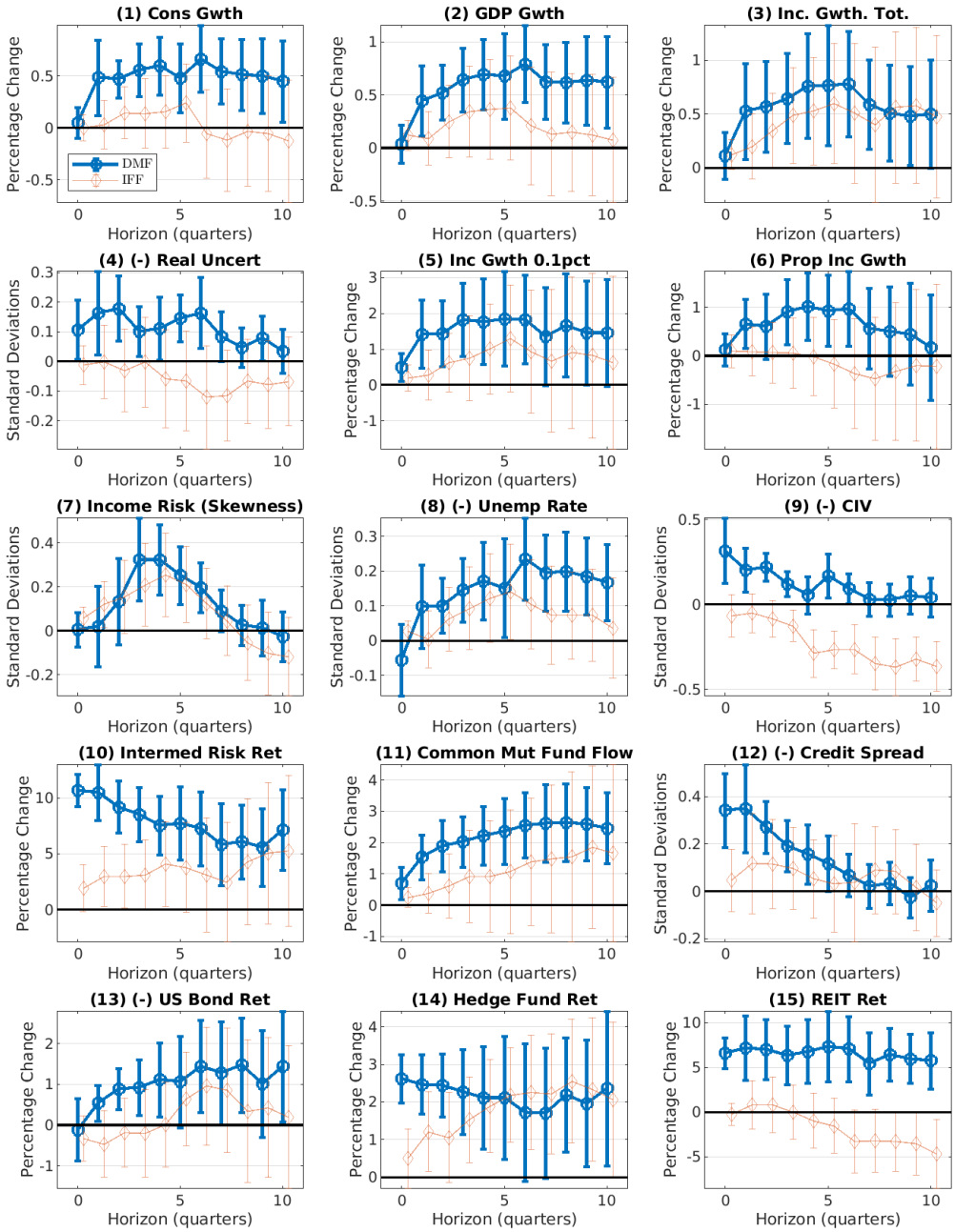
Note. This figure plots normalized coefficients from rolling regressions of the prior 60 months of VMF returns on five PCs constructed from our set of 372 equal-weighted, characteristics-sorted portfolios. For ease of interpretation, coefficients are scaled by their full-sample OLS counterparts. All data are from 1963-2021.

Figure OA.4: Local projection impulse responses to the DMF and IFF (robustness, capped VW PCA)



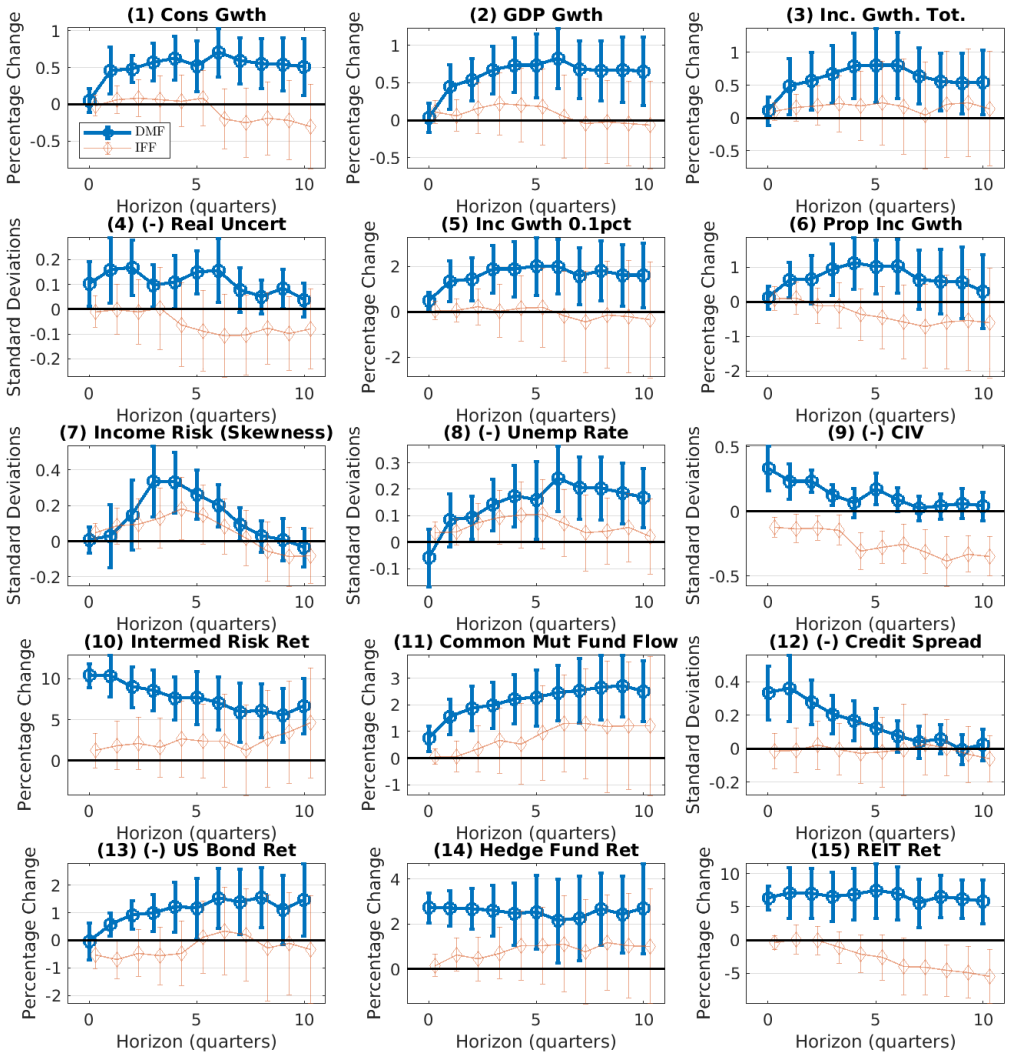
Note. This figure plots the local projection impulse responses of selected macroeconomic aggregates and financial return time series to the DMF and IFF when the DMF is identified using a PCA on our 372 capped value-weighted portfolios (see description in Section 4.2). The indexes are defined in Section 5 and are signed so that increases in the variable value correspond to decreases in marginal utility in standard models. A "(-)" in the panel title indicates that the negative value of a time series is used to conform to this logic. Monthly returns are compounded to 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 response variable. Response variables are in units of percent in the case of return or growth time series (e.g., consumption growth, US bond returns, etc.) or units of standard deviation in the case of stationary time series (e.g., unemployment rate, real uncertainty, etc.). The initial point estimates (i.e., when the horizon is equal to zero) 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. All data are from 1963–2021 (conditional on variable availability). Additional details on the estimation procedure can be found in Online Appendix OA.10.

Figure OA.5: Local projection impulse responses to the DMF and IFF (robustness, no Sm/Lg EW PCA)



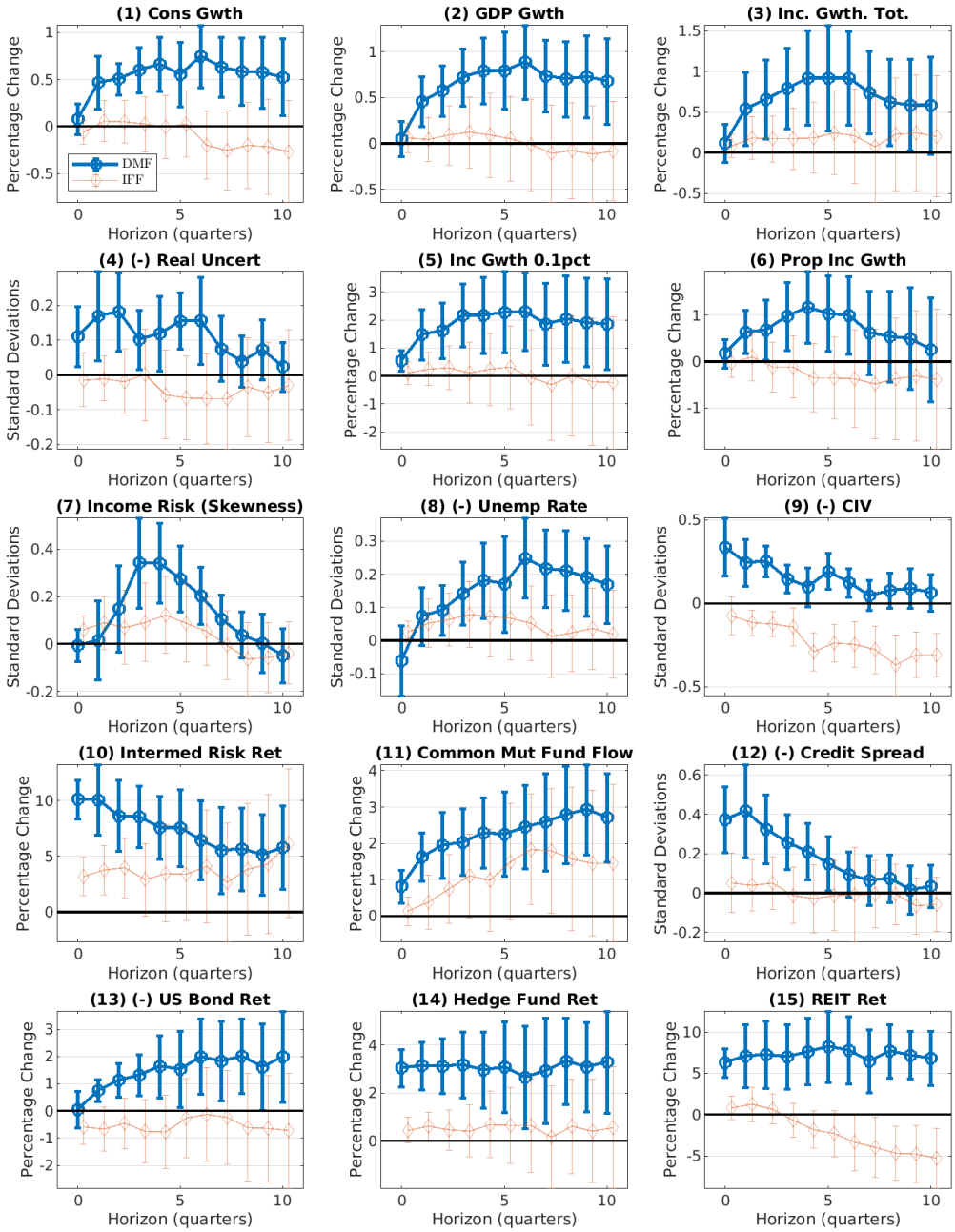
Note. This figure plots local projection impulse responses of selected macroeconomic aggregates and financial return time series to the DMF and IFF when the DMF is identified from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints (see description in Section 4.2). The indexes are defined in Section 5 and are signed so that increases in the variable value correspond to decreases in marginal utility in standard models. A "(-)" in the panel title indicates that the negative value of a time series is used to conform to this logic. Monthly returns are compounded to 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 response variable. Response variables are in units of percent in the case of return or growth time series (e.g., consumption growth, US bond returns, etc.) or units of standard deviation in the case of stationary time series (e.g., unemployment rate, real uncertainty, etc.). The initial point estimates (i.e., when the horizon is equal to zero) 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. All data are from 1963–2021 (conditional on variable availability). Additional details on the estimation procedure can be found in Online Appendix OA.10.

Figure OA.6: Local projection impulse responses to the DMF and IFF (robustness, rolling EW PCA)



Note. This figure plots the local projection impulse responses of selected macroeconomic aggregates and financial return time series to the DMF and IFF when the DMF is identified using a rolling PCA on our 372 equal-weighted portfolios so that we identify the DMF and IFF in an out-of-sample fashion (see description in Section 4.2). The indexes are defined in Section 5 and are signed so that increases in the variable value correspond to decreases in marginal utility in standard models. A "(-)" in the panel title indicates that the negative value of a time series is used to conform to this logic. Monthly returns are compounded to 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 response variable. Response variables are in units of percent in the case of return or growth time series (e.g., consumption growth, US bond returns, etc.) or units of standard deviation in the case of stationary time series (e.g., unemployment rate, real uncertainty, etc.). The initial point estimates (i.e., when the horizon is equal to zero) 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. All data are from 1963–2021 (conditional on variable availability). Additional details on the estimation procedure can be found in Online Appendix OA.10.

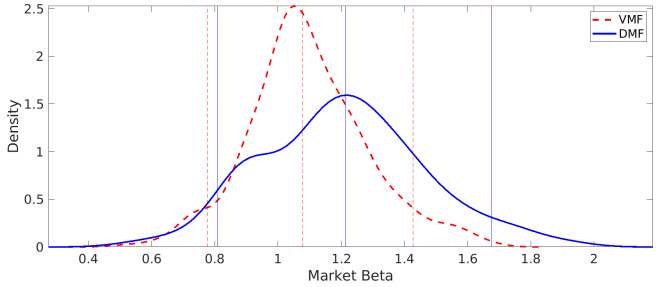
Figure OA.7: Local projection impulse responses to the DMF and IFF (robustness, EW MKT)



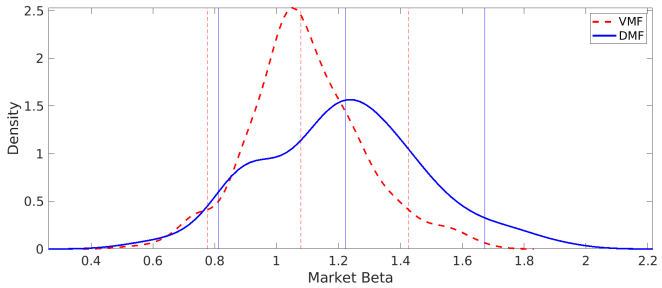
Note. This figure plots the local projection impulse responses of selected macroeconomic aggregates and financial return time series to the DMF and IFF when the DMF is identified using a projection of the VMF onto excess returns on the CRSP equal-weighted market index (see description in Section 4.2). The indexes are defined in Section 5 and are signed so that increases in the variable value correspond to decreases in marginal utility in standard models. A "(-)" in the panel title indicates that the negative value of a time series is used to conform to this logic. Monthly returns are compounded to 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 response variable. Response variables are in units of percent in the case of return or growth time series (e.g., consumption growth, US bond returns, etc.) or units of standard deviation in the case of stationary time series (e.g., unemployment rate, real uncertainty, etc.). The initial point estimates (i.e., when the horizon is equal to zero) 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. All data are from 1963–2021 (conditional on variable availability). Additional details on the estimation procedure can be found in Online Appendix OA.10.

Figure OA.8: Distribution of β_p^{VMF} and β_p^{DMF} for a large number of test assets (robustness)

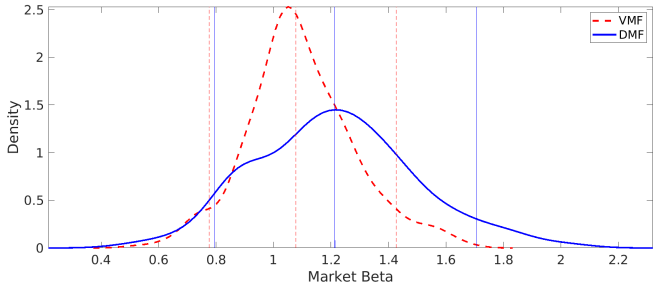
Panel 1: Capped VW PCA



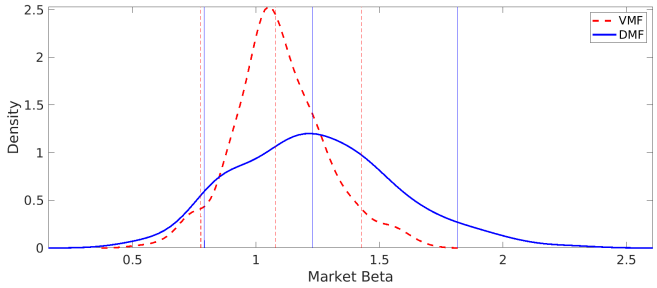
Panel 2: No Sm/Lg EW PCA



Panel 3: Rolling EW PCA



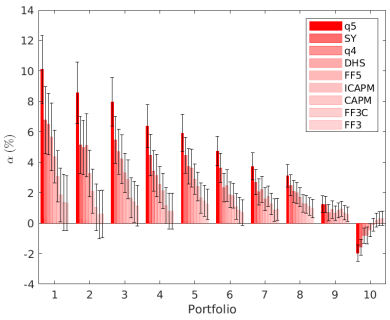
Panel 4: EW MKT



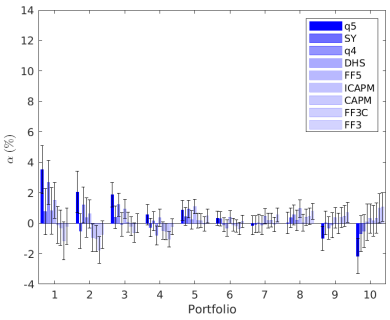
Note. This figure plots kernel density estimates of market betas for our set of 372 value-weighted portfolios estimated using the VMF and DMF with the latter based on four robustness checks: 1) "Capped VW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio (Panel 1), 2) "No Sm/Lg EW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints (Panel 2), 3) "Rolling EW PCA" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify PCA portfolio loadings each month using the standard 372 equal-weighted portfolios (Panel 3), and 4) "EW MKT" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index (Panel 4). For more details on these constructions, see Section 4.2. See Online Appendix Table OA.1 for additional details about the 372 portfolios. 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 1963–2021.

Figure OA.9: Alphas for size-sorted decile portfolios using either the VMF or the DMF (robustness)

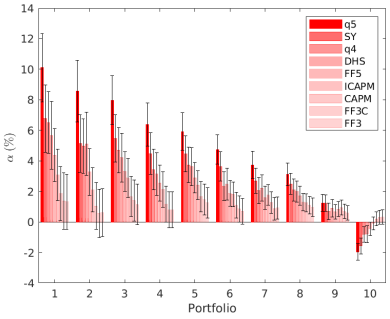
Panel 1: VMF



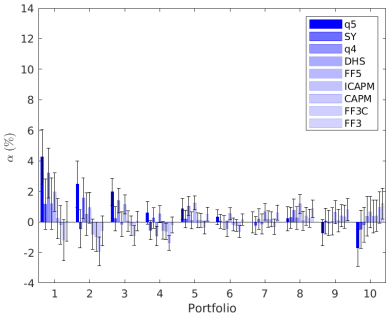
Panel 2: DMF, Capped VW PCA



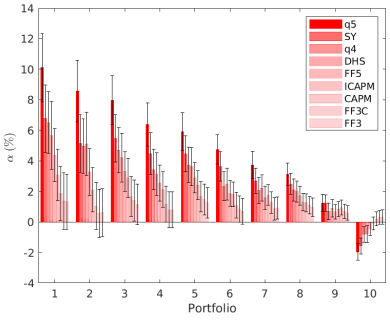
Panel 3: VMF



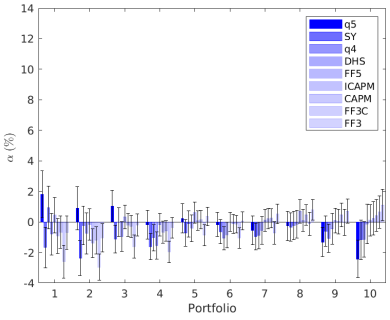
Panel 4: DMF, No Sm/Lg EW PCA



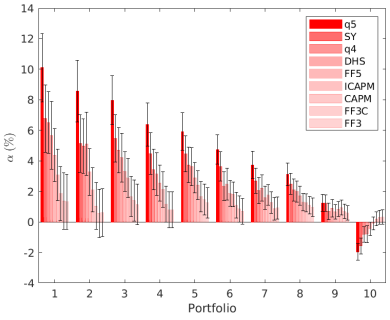
Panel 5: VMF



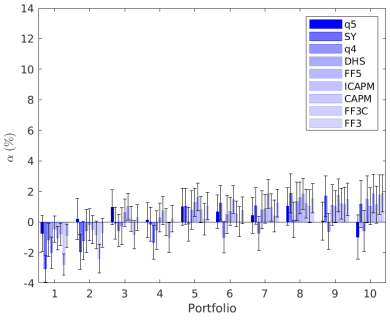
Panel 6: DMF, Rolling EW PCA



Panel 7: VMF



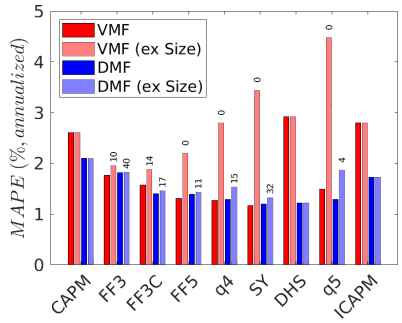
Panel 8: VMF, EW MKT



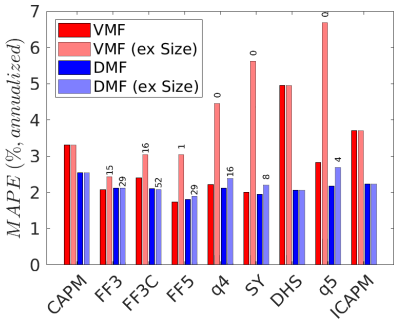
Note. This figure reports alphas for value-weighted size-sorted decile portfolios (obtained from our set of 372 characteristics-sorted portfolios) for each of the nine factor models described in Section 6.2. We drop size factors from the six models with size factors (FF3, FF3C, FF5, q4, SY, and q5) in this analysis because, consistent with our theoretical model and our empirical results, the primary role of these factors is help correct for attenuation in the market beta estimates that we study here. The left panels (1, 3, 5, and 7) use the VMF as the market factor in each model (repeated for comparison with the DMF-based panels). The right panels (2, 4, 6, and 8) use the DMF as the market factor in each model based on four robustness checks: 1) "Capped VW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio (Panel 2), 2) "No Sm/Lg EW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints, 3) "Rolling EW PCA" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify PCA portfolio loadings each month (using the standard 372 equal-weighted portfolios) (Panel 6), and 4) "EW MKT" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index (Panel 8). For more details on these constructions, see Section 4.2. The error bars represent 95% confidence intervals on each alpha estimate based on the bootstrap procedure described in Online Appendix OA.12. All data are from 1963–2021 (conditional on factor model availability).

Figure OA.10: Mean absolute pricing errors for a large number of test assets (robustness)

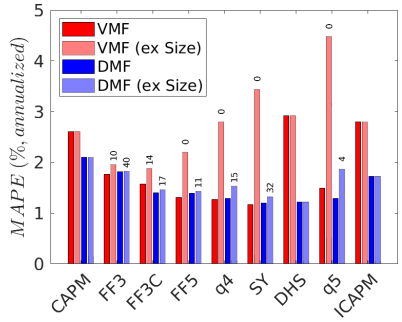
Panel 1: VW portfolios, Capped VW PCA



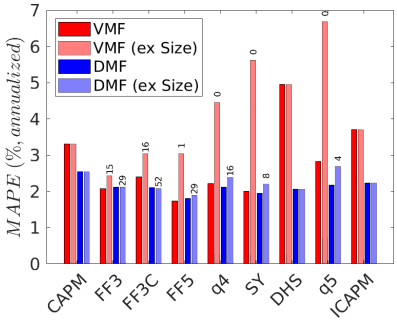
Panel 2: EW portfolios, Capped VW PCA



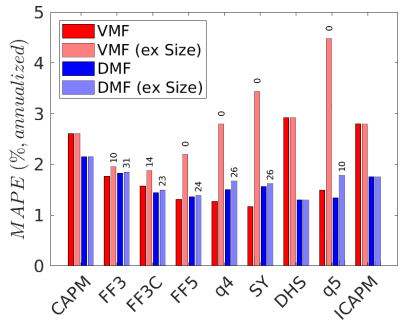
Panel 3: VW portfolios, No Sm/Lg EW PCA



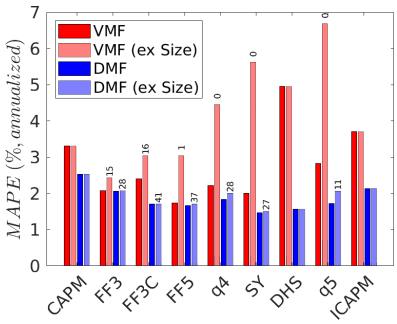
Panel 4: EW portfolios, No Sm/Lg EW PCA



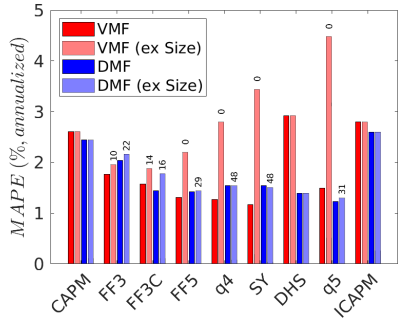
Panel 5: VW portfolios, Rolling EW PCA



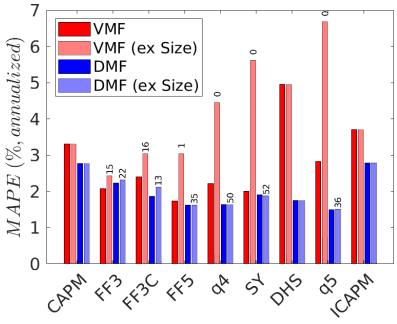
Panel 6: EW portfolios, Rolling EW PCA



Panel 7: VW portfolios, EW MKT

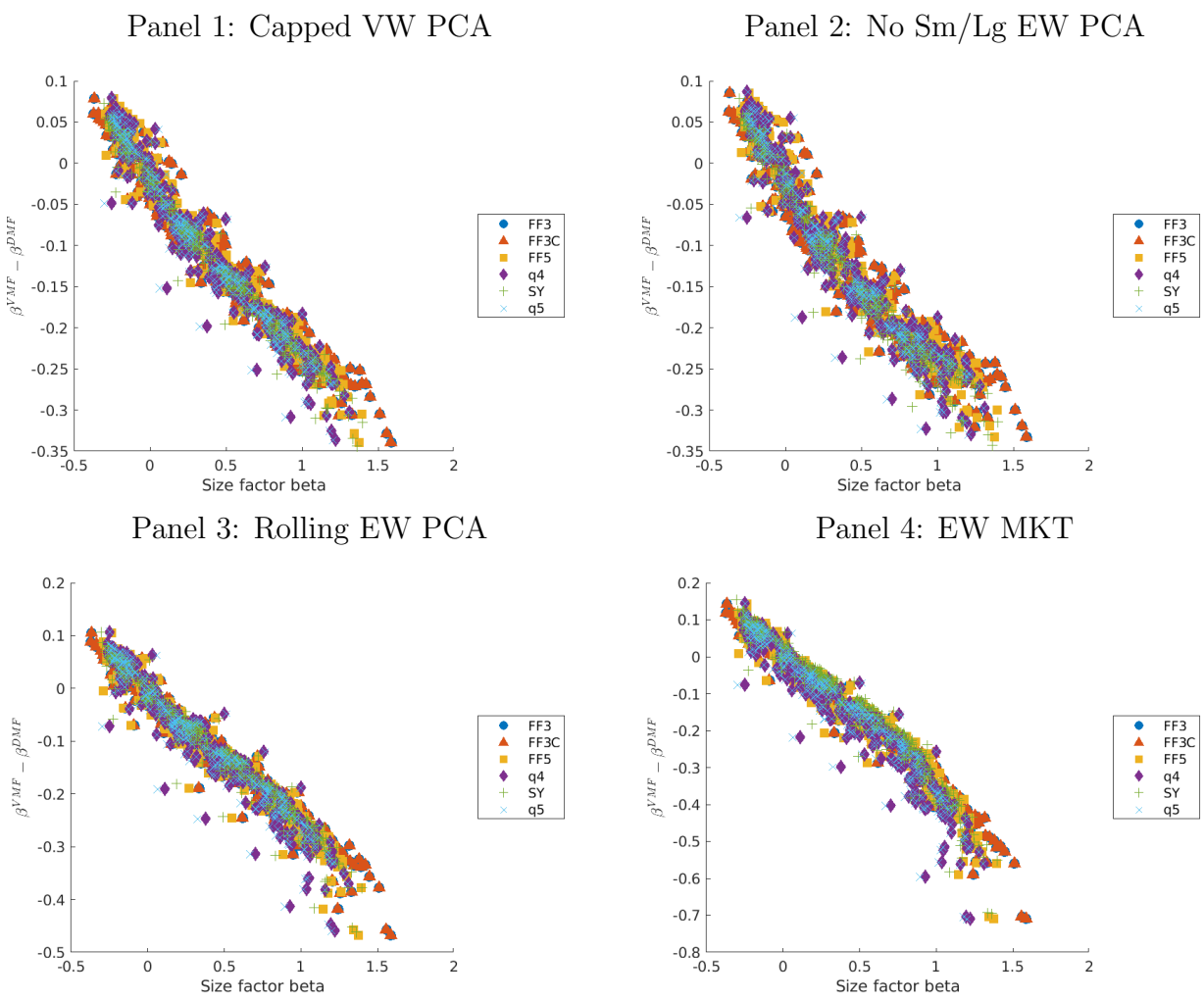


Panel 8: EW portfolios, EW MKT



Note. This figure plots the mean absolute pricing errors (α) for our set of 372 characteristics-sorted portfolios as test assets across the nine standard factor models described in Section 6.2 and listed in chronological order based on publication date. The left panels (1, 3, 5, and 7) use value-weighted portfolios. The right panels (2, 4, 6, and 8) use equal-weighted portfolios. In this set of robustness checks, the DMF is computed in four ways: 1) "Capped VW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio (Panels 1 and 2), 2) "No Sm/Lg EW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints (Panels 3 and 4), 3)"Rolling EW PCA" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify PCA portfolio loadings each month (using the standard 372 equal-weighted portfolios) (Panels 5 and 6), and 4) "EW MKT" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index (Panels 7 and 8). For more details on these constructions see Section 4.2. Portfolio definitions are described in Online Appendix Table OA.1. We study four versions of each model: 1) "VMF": Model that uses the VMF as the market factor and includes size factors when present, 2) "VMF (ex Size)": Model that uses the VMF as the market factor and excludes size factors when present, 3) "DMF": Model that uses the DMF as the market factor and includes size factors when present, and 4) "DMF (ex Size)": Model that uses the DMF as the market factor and excludes size factors when present. For models that include size factors, the numbers above the bars represent p-values (in percent) from testing the null hypothesis that models that exclude the size factors yield lower mean absolute pricing errors than the alternative model that excludes the size factor. P-values are based on the bootstrap procedure described in Online Appendix OA.12. All data are from 1963–2021 (conditional on factor model availability).

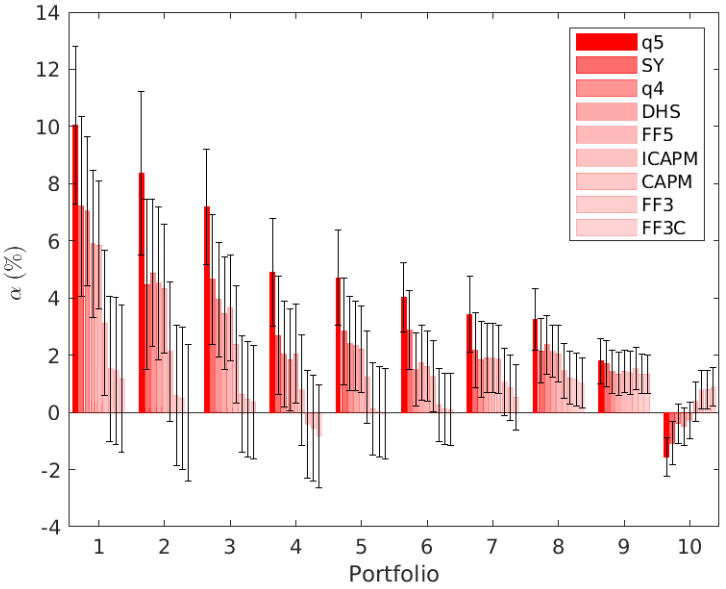
Figure OA.11: Beta gap vs. size factor betas (robustness)



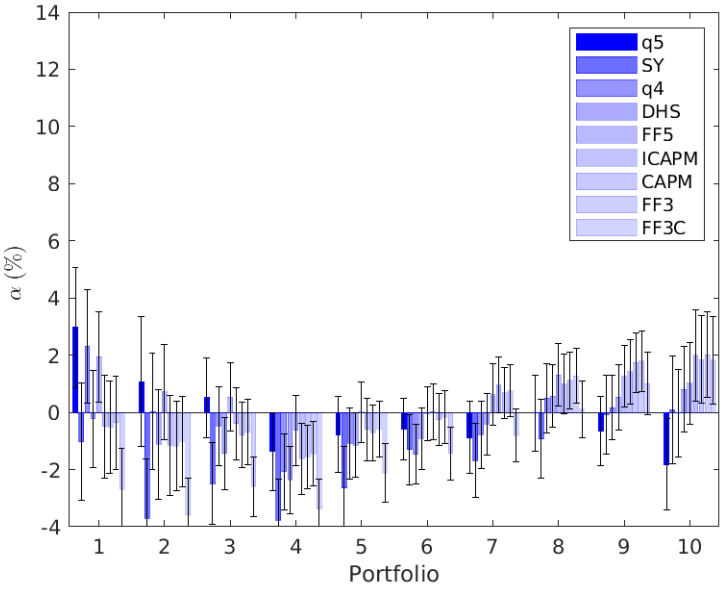
Note. This figure reports the beta gaps ($\beta_p^{VMF} - \beta_p^{DMF}$) versus size factor betas among the six factor models we study that include size factors (FF3, FF3C, FF5, q4, SY, and q5; see Section 6.2 for a description of these models) using value-weighted versions of our set of 372 characteristics-sorted portfolios as test assets. The panels report results using four alternative DMF constructions: 1) "Capped VW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio (Panel 1), 2) "No Sm/Lg EW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints (Panel 2), 3) "Rolling EW PCA" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify PCA portfolio loadings each month using the standard 372 equal-weighted portfolios (Panel 3), and 4) "EW MKT" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index (Panel 4). For more details on these constructions, see Section 4.2. Portfolio definitions are described in Online Appendix Table OA.1. β_p^{VMF} and β_p^{DMF} are estimated as multivariate betas within each model that use either the VMF or the DMF as the market factor. The size factor betas are estimated as multivariate betas within each model that uses the VMF as the market factor. All data are from 1963–2021 (conditional on factor model availability).

Figure OA.12: Alphas for size-sorted decile portfolios using the VMF or the DMF (1993-2021 sample)

Panel 1: VMF-based models

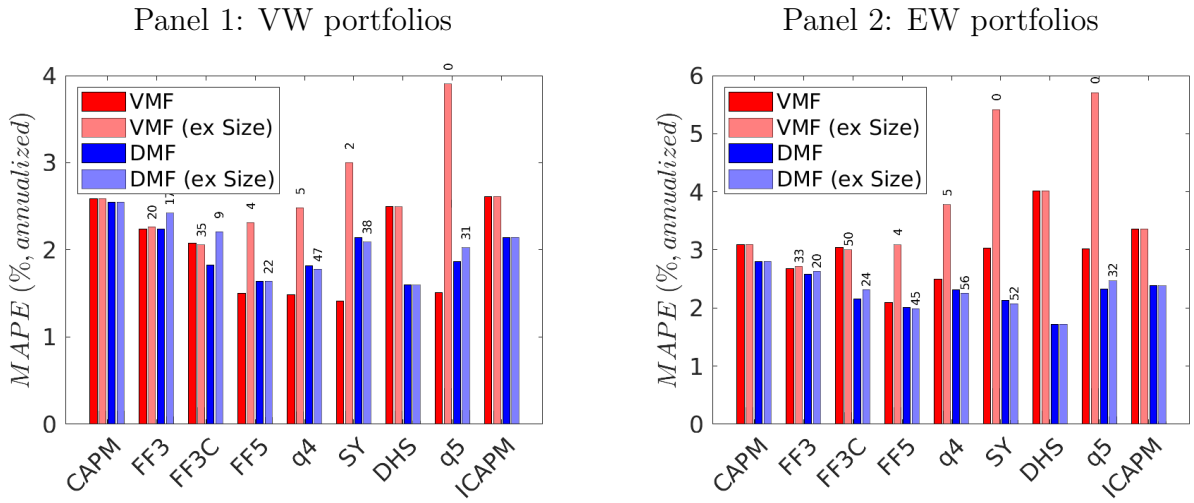


Panel 2: DMF-based models



Note. This figure reports alphas for value-weighted size-sorted decile portfolios (obtained from our set of 372 characteristics-sorted portfolios) for each of the nine factor models described in Section 6.2. We drop size factors from the six models with size factors (FF3, FF3C, FF5, q4, SY, and q5) in this analysis because, consistent with our theoretical model and our empirical results, the primary role of these factors is help correct for attenuation in the market beta estimates that we study here. All portfolio returns were obtained from Kenneth French's webpage and are described in Online Appendix Table OA.1. Panel 1 uses the VMF as the market factor in each model. Panel 2 uses the DMF as the market factor in each model. The error bars represent 95% confidence intervals on each alpha estimate based on the bootstrap procedure described in Online Appendix OA.12. All data are from 1993–2021 (conditional on factor model availability).

Figure OA.13: Mean absolute pricing errors for a large number of test assets (1993-2021 sample)



Note. This figure plots the mean absolute pricing errors (α) for our set of 372 characteristics-sorted portfolios as test assets across the nine standard factor models described in Section 6.2 and listed in chronological order based on publication date. Panel 1 uses value-weighted portfolios. Panel 2 uses equal-weighted portfolios. All portfolio returns were obtained from Kenneth French's webpage and are described in Online Appendix Table OA.1. We study four versions of each model: 1) "VMF": Model that uses the VMF as the market factor and includes size factors when present, 2) "VMF (ex Size)": Model that uses the VMF as the market factor and excludes size factors when present, 3) "DMF": Model that uses the DMF as the market factor and includes size factors when present, and 4) "DMF (ex Size)": Model that uses the DMF as the market factor and excludes size factors when present. For models that include size factors, the numbers above the bars represent p-values (in percent) from testing the null hypothesis that models that exclude the size factors yield a lower mean absolute pricing error that the alternative model that excludes the size factor. P-values are based on the bootstrap procedure described in Online Appendix OA.12. All data are from 1993–2021 (conditional on factor model availability).

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 PCA (equal-weighted versions of the portfolios) and 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: Explaining the VMF with PCs

Variable	Sharpe ratio	(1)	(2)	(3)	(4)	(5)	$\beta_{PC_i} \cdot \mathbb{E}[PC_i]$
VMF	0.44***						6.79***
	(0.14)						(2.00)
DMF (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	

Note: This table provides results from regressions of the VMF on the DMF and higher PCs (note that the DMF is just a scaled version of the first PC). The DMF and the higher PCs are constructed using our set of 372 equal-weighted portfolios; they are signed to have positive average returns and scaled to have unit standard deviation. All portfolio returns were obtained from Kenneth French's webpage and are described in Online Appendix Table OA.1. 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 VMF 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 fourth to last row 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 parentheses and are computed using the bootstrap procedure described in Online Appendix OA.12. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from 1963–2021.

Table OA.3: 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 statistics for the granular residual (GR) estimated as the residual from regressing the IFF on progressively more PCs. It also provides summary statistics for the VMF and the IFF for comparison. As highlighted in equation (9), the IFF is the sum of three components: 1) time-varying market exposure to the first PC, 2) time-varying market exposure to higher PCs, and 3) the GR. We estimate the GR by regressing the IFF on various numbers of the 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 the results from full-sample regressions, effectively imposing constant IFF loadings on PCs. 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 the 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 1963–2021.

Table OA.4: Size factor spanning tests (robustness)

	Factor model					
Alphas	FF3	FF3C	FF5	q4	SY	q5
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]
α^{DMF} (Capped VW PCA)	-0.03	-0.62	0.75	1.60	2.09	2.66*
	[-0.03]	[-0.55]	[0.65]	[1.10]	[1.64]	[1.74]
α^{DMF} (No Sm/Lg EW PCA)	-0.05	-0.76	0.68	1.52	1.78	2.48
	[-0.05]	[-0.69]	[0.59]	[1.03]	[1.40]	[1.58]
α^{DMF} (Rolling EW PCA)	-0.26	-1.18	0.35	1.02	1.20	2.03
	[-0.24]	[-1.08]	[0.30]	[0.70]	[0.95]	[1.31]
α^{DMF} (EW MKT)	-0.60	-1.63	-0.35	-0.01	0.16	0.63
	[-0.57]	[-1.60]	[-0.32]	[0.00]	[0.13]	[0.44]

Note: This table presents factor spanning tests for the size factors in each factor model when the VMF (reproduced for comparison) or the DMF is used as the market factor, with the DMF being based on our four alternative constructions: 1) "Capped VW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio, 2) "No Sm/Lg EW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints, 3) "Rolling EW PCA" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify the PCA portfolio loadings each month (using the standard 372 equal-weighted portfolios), and 4) "EW MKT" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index. For more details on these constructions, see Section 4.2. We only include models with size factors and test each model using its respective size factor. α^{VMF} (α^{DMF}) represents the alpha when the VMF (DMF) is used and is annualized and in percent. We omit β estimates for brevity, but results are similar to our main results from Table 3. T-statistics are reported in brackets and are based on the bootstrap procedure described in Online Appendix OA.12. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from 1963–2021 (conditional on factor model availability).

Table OA.5: IFF spanning tests (robustness)

	Factor model								
	CAPM	FF3	FF3C	FF5	q4	SY	DHS	q5	ICAPM
Panel A: Capped VW PCA									
α^{VMF} (w/ size)		0.11	-0.38*	0.18	-0.16	-0.03		-0.73**	
		[0.45]	[-1.69]	[0.79]	[-0.50]	[-0.10]		[-2.44]	
α^{VMF} (w/o size)	-1.06*	-0.49	-0.99	-1.23**	-1.99***	-2.82***	-2.39***	-3.64***	-1.64***
	[-1.70]	[-0.81]	[-1.61]	[-2.04]	[-2.98]	[-3.90]	[-3.09]	[-5.09]	[-2.76]
α^{DMF} (w/ size)		0.29	-0.26	0.42	0.15	0.51		-0.32	
		[0.99]	[-0.96]	[1.48]	[0.37]	[1.20]		[-0.80]	
α^{DMF} (w/o size)	-0.33	0.30	0.04	0.03	-0.59	-0.54	-0.34	-1.54**	-0.45
	[-0.50]	[0.48]	[0.06]	[0.05]	[-0.83]	[-0.68]	[-0.42]	[-2.00]	[-0.66]
Panel B: No Sm/Lg EW PCA									
α^{VMF} (w/ size)		0.08	-0.65**	0.25	-0.17	-0.31		-0.87**	
		[0.23]	[-2.11]	[0.78]	[-0.40]	[-0.74]		[-2.21]	
α^{VMF} (w/o size)	-1.18*	-0.53	-1.28*	-1.23*	-2.08***	-3.21***	-2.68***	-3.88***	-1.82***
	[-1.74]	[-0.82]	[-1.91]	[-1.85]	[-2.79]	[-4.14]	[-3.14]	[-5.04]	[-2.78]
α^{DMF} (w/ size)		0.39	-0.43	0.69*	0.43	0.52		0.00	
		[0.97]	[-1.13]	[1.74]	[0.75]	[0.96]		[0.00]	
α^{DMF} (w/o size)	-0.29	0.42	-0.05	0.33	-0.29	-0.43	-0.14	-1.18	-0.30
	[-0.39]	[0.60]	[-0.07]	[0.46]	[-0.36]	[-0.48]	[-0.15]	[-1.39]	[-0.40]
Panel C: Rolling EW PCA									
α^{VMF} (w/ size)		-0.19	-1.21***	-0.54	-1.29**	-1.36**		-1.73***	
		[-0.55]	[-3.26]	[-1.28]	[-2.34]	[-2.46]		[-2.99]	
α^{VMF} (w/o size)	-1.29*	-0.85	-1.88**	-2.01***	-3.17***	-4.26***	-3.49***	-4.74***	-2.11***
	[-1.84]	[-1.22]	[-2.49]	[-2.84]	[-3.98]	[-4.80]	[-3.58]	[-5.67]	[-3.20]
α^{DMF} (w/ size)		0.15	-1.11**	-0.08	-0.87	-0.59		-0.96	
		[0.34]	[-2.39]	[-0.15]	[-1.30]	[-0.85]		[-1.41]	
α^{DMF} (w/o size)	-0.29	0.29	-0.46	-0.26	-1.36	-1.24	-0.72	-1.95**	-0.46
	[-0.39]	[0.40]	[-0.59]	[-0.35]	[-1.59]	[-1.28]	[-0.75]	[-2.17]	[-0.59]
Panel D: EW MKT									
α^{VMF} (w/ size)		-0.22	-1.19**	-0.81	-2.05***	-1.53**		-2.44***	
		[-0.48]	[-2.31]	[-1.53]	[-3.00]	[-2.06]		[-3.35]	
α^{VMF} (w/o size)	-1.35	-1.01	-2.01**	-2.55***	-4.18***	-4.99***	-4.11***	-5.85***	-2.07***
	[-1.60]	[-1.19]	[-2.10]	[-2.98]	[-4.41]	[-4.46]	[-3.47]	[-5.76]	[-2.59]
α^{DMF} (w/ size)		0.56	-0.60	0.26	-1.11	0.87		-0.55	
		[0.88]	[-0.83]	[0.37]	[-1.25]	[0.88]		[-0.61]	
α^{DMF} (w/o size)	0.36	0.97	0.58	0.50	-1.10	0.76	0.62	-0.93	1.04
	[0.38]	[1.00]	[0.55]	[0.51]	[-1.03]	[0.60]	[0.50]	[-0.81]	[1.05]

Note: This table presents factor spanning tests for the IFF using each factor model when the VMF or the DMF is used as the market factor, with the DMF being based on our four alternative constructions: 1) "Capped VW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a capped value-weighted version of our 372 characteristics-sorted portfolio (Panel A), 2) "No Sm/Lg EW PCA" is constructed as a projection of the VMF onto the first PC from a PCA using a version of our 372 characteristics-sorted portfolios that drops stocks in the lowest and highest 20th percentiles of the size distribution based on NYSE breakpoints (Panel B), 3) "Rolling EW PCA" is an out-of-sample (tradable) version of the DMF constructed using a 60-month historical rolling window to identify PCA portfolio loadings each month (using the standard 372 equal-weighted portfolios) (Panel C), and 4) "EW MKT" is constructed as a projection of the VMF onto excess returns on the CRSP equal-weighted market index (Panel D). For more details on these constructions, see Section 4.2. Within each panel, we report alphas from models that use the VMF (DMF) as α^{VMF} (α^{DMF}). For each case, we also use models that include their size factors (designated as "w/ size") or that exclude their size factors (designated as "w/o size"). Note that the CAPM, DHS, and ICAPM models do not include size factors, and therefore we do not present any results "w/ size" in these cases. α is annualized and in percent. We omit β estimates for brevity, but the results are similar to our main results from Table 4. T-statistics are reported in brackets and are based on the bootstrap procedure described in Online Appendix OA.12. */**/*** represent statistical significance at 90%, 95%, and 99%, respectively. All data are from 1963–2021 (conditional on factor model availability).