So what do we learn from Li and Wang (2022)?[a]

Michael Ungeheuer[b] Martin Weber,[c]

December 2022

[a]We are grateful to Zwetelina Iliewa, Heiko Jacobs, Petri Jylhä, and Esad Smajlbegovic for valuable comments.

[b]Michael Ungeheuer: Aalto University, michael.ungeheuer@aalto.fi.

[c]Martin Weber: University of Mannheim and the Centre for Economic Policy Research, London EC1V 3PZ, United Kingdom, weber@bank.BWL.uni-mannheim.de.

Li and Wang (2022) replicate our asset pricing result (Ungeheuer and Weber, 2021) on the return premium for stocks that frequently comove with the market (high-*Comove* stocks). Their successful replication includes our robustness check where the premium remains significant when controlling for the linear association between idiosyncratic volatility (*IVOL*) and stock returns. They then go beyond our tests by controlling for a non-linear association between *IVOL*, *Comove*, and stock returns. They find that this step turns the return premium we document statistically insignificant, with the 95% confidence interval for the annualized *Comove* return premium [–0.37%, 3.13%] covering both zero and an economically relevant positive return premium. It is difficult to say what we learn from one of many specifications leading to a confidence interval that covers zero, especially if that specification is unusual and purely motivated by a statistical association with the variable of interest.

In this response, we start by arguing that Li and Wang (2022)'s critique illustrates a weakness of empirical asset pricing tests, which are not motivated by ex-ante hypotheses based on theory or evidence on portfolio choice. Such ad-hoc tests provide too many degrees of freedom to result in persuasive evidence. We then continue by discussing how we believe convincing progress can be made, by using experiments. Causal evidence based on experiments makes up the bulk of our study, in contrast to the correlational evidence scrutinized by Li and Wang (2022). It rules out any confounds based on marginal distributions including *IVOL* by design. Better yet, building on our experimental design allows us to provide insights on the economic mechanism that explains causal effects of dependence on portfolio choice. Such a deep analysis of the validity of fundamental assumptions we use to model portfolio choice and asset prices is difficult to achieve when analyzing archival data with naturally correlated competing explanatory variables.

1 A dead end: Why we learn little from Li and Wang (2022)

Li and Wang (2022)'s critique illustrates a weakness of common empirical asset pricing tests based on ad-hoc return predictors. Ad-hoc predictors, which are neither justified by theory nor evidence on portfolio choice leave too many degrees of freedom (Harvey, 2017). This methodological issue is

amplified for archival data tests, where the sample is naturally limited, so that analyses are exposed to misleading conclusions that cannot be tested via a clean replication.

In the context of Li and Wang (2022)'s test, we see two important issues. First, theory and evidence on portfolio choice and asset pricing point at a positive or zero return premium for high-IVOL stocks, not a negative association between IVOL and subsequent returns. Established models that link past returns to prices and future returns do not suggest that higher IVOL should be associated with lower returns, neither linearly nor non-linearly (starting with Sharpe, 1964). If anything, the association should be positive under realistic levels of investor recognition (e.g., Merton, 1987) and diversification (e.g., Lewis, 1999; Grinblatt and Keloharju, 2001). Evidence for actual behavior suggests that investors indeed aim to avoid high-volatility investments (e.g., Chinco et al., 2022) and that they are not drawn to more volatile assets when volatility is varied in controlled experiments (e.g., Ungeheuer and Weber, 2022a,b).

Second, theory and evidence on the low returns of high-*IVOL* stocks do not suggest that a (linear or non-linear) association between *IVOL* and stock returns adequately describes asset prices. Instead, different constructs are proposed, which are empirically associated with and thus difficult to disentangle from *IVOL* and *Comove* using historical return data. For example, lottery characteristics might explain the high prices and low returns of stocks with recently extreme returns. Associated return predictors include maximum daily returns (Bali et al., 2011), expected idiosyncratic skewness (Boyer et al., 2010), and cumulative prospect theory values (Barberis et al., 2016). Many other models have been proposed to explain the low returns of stocks with historically extreme returns, like attention- and salience-based mechanisms (Barber and Odean, 2008; Cosemans and Frehen, 2021) or expected coskewness risk (Schneider et al., 2020). In line with *IVOL* capturing a latent omitted variable, the *IVOL* puzzle varies strongly with seemingly arbitrary estimation choices (Bali and Cakici, 2008). Candidate explanations for the *IVOL* puzzle have in common that it is not *IVOL* itself that drives asset prices. However, it remains unclear which candidate(s) solve(s) the puzzle (see Hou and Loh, 2016, for a test and a long list of competing explanations).

In short, controlling for a non-linear negative association between *IVOL* and stock returns is not justified by theory or evidence on portfolio choice and asset pricing. A justification of control

variables purely based on a mechanically or statistically close association with a variable of interest is bound to result in 'statistically insignificant' results at some point and thus methodologically problematic. In contrast, our asset pricing test in Ungeheuer and Weber (2021) is based on an ex-ante hypothesis, with our *Comove* measure being justified by evidence on portfolio choice from randomized experiments (the core contribution of the paper) combined with well-known theory (the CAPM, Sharpe, 1964).

2 Making progress: How to learn about drivers of portfolio choice in Ungeheuer and Weber (2021)

Well designed experiments can help us test fundamental assumptions about investor behavior, where yet another correlation in historical stock returns leaves us in limbo. As stressed in Ungeheuer and Weber (2021), the mechanism driving a rejection of narrow framing in our experiments remains unclear. Standard models on portfolio choice and asset pricing (like Markowitz, 1952; Sharpe, 1964) suggest that perceived benefits of diversification for portfolio risk drive the effect. This conventional mechanism motivates asset pricing tests based on the *Comove* measure. However, there is some evidence suggesting a search for alternative mechanisms.

Next to the investigation of subtle alternative mechanisms, experiments can also be used to rule out some of the boilerplate objections brought forward against lab experiments. Li and Wang (2022) shrug off the core contribution of our study—the experimental results—with references to external validity issues like differences in experience and sophistication of real investors, as well as the 'non-natural' return distributions. We believe that this broad-brush criticism of experiments is outdated and tosses the baby out with the bathwater. First, the models we test do not specify that they are only valid for certain subsets of the population or specific return distributions. Second, there is no need to replace 'gold standard' causal estimates from randomized experiments with exploratory correlations from archival data to counter such external validity concerns. External validity can be tested within the experimental framework.

We now use recent research to show how experiments can help us uncover alternative portfolio choice mechanisms and disentangle them from conventional portfolio return optimization. Along

the way, the following studies also show how external validity concerns can be refuted. The *Comove* measure in our asset pricing test is based on evidence on portfolio choice from randomized experiments. The experimental design rules out many confounds including Li and Wang (2022)'s non-linear effect of *IVOL*. Participants' beliefs about return dependence in the experiments are driven by the frequency of equally-signed returns (*Comove*) and their choices are closely linked to these beliefs. Thus, results are consistent with a beliefs-based channel, where participants select portfolios like a mean-variance optimizer based on biased beliefs about dependence. Laudenbach et al. (2022) find consistent evidence in different variations of the experimental design, including replications with real investors. The consistency of results from the lab and with real investors refutes Li and Wang (2022)'s external validity concerns about investment experience. However, some evidence casts doubt on a mechanism based on portfolio return optimization.

First, participants' beliefs about portfolio risk are insensitive to large changes in dependence. The probability of a loss varies strongly between conditions in our study and Laudenbach et al. (2022), but participants' beliefs about portfolio risk remain virtually the same. Although participants select portfolios *as if* they incorporate diversification benefits, this casts doubt on a mechanism based on portfolio return optimization ('broad framing'). Second, other researchers find that investors' preferences and beliefs are inconsistent with textbook portfolio return optimization. Chinco et al. (2022) find that most investors claim to not consider correlation at all when making investment decisions. The few investors who consider correlation seem to prefer investing in high-correlation assets with low diversification benefits, although the overwhelming majority of investors prefer a high-mean and low-volatility portfolio. This result holds for retail investors, as well as more experienced and sophisticated professionals. Reinholtz et al. (2021) find that investors do not understand basic effects of diversification on portfolio risk and return.

So what could explain a rejection of narrow framing besides diversification benefits for the portfolio return distribution? Research from decision science suggests that dependence might also matter via anticipated regret of underperformance (Loomes and Sugden, 1982). A low-return asset may become less attractive when it frequently co-moves with another asset (or a benchmark index) because frequent co-movement is naturally associated with infrequent outperformance. Building on

our experimental design, Ungeheuer and Weber (2022b) show that investors are drawn to frequently outperforming assets. Their evidence supports a beliefs-based mechanism, where investors become overoptimistic for such assets. Marginal return distributions in these experiments are based on 'natural' historical returns of the S&P 500. This counters an external validity concern used by Li and Wang (2022) to discount our causal evidence from experiments. Systematically selecting boom and bust subsamples, or periods with high or low volatility, skewness, or kurtosis, does not affect findings. Results also hold for both retail investors and professional asset managers. This counters the other external validity concern mentioned by Li and Wang (2022) and confirms the frequent finding that more experience and sophistication do not change results.

Could our results be driven by both, the frequent outperformance *and* the diversification benefits of low-*Comove* assets? Ungeheuer and Weber (2022a) 'turn off' the frequency of outperformance mechanism to test just that. The result is a precise zero-effect, where participants do not react to significant variation in diversification benefits despite strong variation in correlation and the frequency of co-movement. This study also rules out alternative explanations for a non-result, including low risk aversion (e.g., due to a lack of 'skin in the game,' Rabin, 2000).

Overall, the two follow-up studies of Ungeheuer and Weber (2021) demonstrate how experiments can be used to make constructive progress on our understanding of fundamental assumptions about portfolio choice and asset pricing. We believe that there is great potential in such research (see Chinco et al., 2022, for a more general discussion), in contrast to inconclusive stand-alone analyses of correlations in historical stock returns.

References

Bali, Turan G, and Nusret Cakici, 2008, Idiosyncratic Volatility and the Cross Section of Expected Returns, *Journal of Financial and Quantitative Analysis* 43, 29–58.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427–446.

Barber, Brad M., and Terrance Odean, 2008, All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *Review of Financial Studies* 21, 785–818.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect theory and stock returns: An empirical test, *Review of Financial Studies* 29, 3068–3107.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected Idiosyncratic Skewness, *Review of Financial Studies* 23, 170–202.

Chinco, Alex, Samuel M. Hartzmark, and Abigail B. Sussman, 2022, A New Test of Risk Factor Relevance, *Journal of Finance* 77, 2183–2238.

Cosemans, Mathijs, and Rik Frehen, 2021, Salience Theory and Stock Prices: Empirical Evidence, *Journal of Financial Economics* 140, 460–483.

Grinblatt, Mark, and Matti Keloharju, 2001, How Distance, Language, and Culture Influence Stockholdings and Trades, *Journal of Finance* 56, 1053–1073.

Harvey, Campbell R., 2017, Presidential address: The scientific outlook in financial economics, *Journal of Finance* 72, 1399–1440.

Hou, Kewei, and Roger K. Loh, 2016, Have we solved the idiosyncratic volatility puzzle?, *Journal of Financial Economics* 121, 167–194.

Laudenbach, Christine, Michael Ungeheuer, and Martin Weber, 2022, How to Alleviate Correlation Neglect in Investment Decisions, *Management Science* (forthcoming).

Lewis, Karen K, 1999, Trying to Explain Home Bias in Equities and Consumption, *Journal of Economic Literature* 37, 571–608.

Li, Peixin, and Baolian Wang, 2022, The Ungeheuer and Weber (2021) Comove and Stock Returns Effect Disappears with Control for Idiosyncratic Volatility, *Critical Finance Review* (forthcoming).

Loomes, Graham, and Robert Sugden, 1982, Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty, *The Economic Journal* 92, 805–824.

Markowitz, Harry M, 1952, Portfolio Selection, *Journal of Finance* 7, 77–91.

Merton, Robert C., 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance* 42, 483–510.

Rabin, Matthew, 2000, Risk Aversion and Expected-Utility Theory: A Calibration Theorem, *Econometrica* 68, 1281–1292.

Reinholtz, Nicholas, Philip Fernbach, and Bart De Langhe, 2021, Do People Understand the Benefit of Diversification?, *Management Science* 67, 7291–7950.

Schneider, Paul, Christian Wagner, and Josef Zechner, 2020, Low-Risk Anomalies?, *Journal of Finance* 75, 2673–2718.

Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance* 19, 425–442.

Ungeheuer, Michael, and Martin Weber, 2021, The Perception of Dependence, Investment Decisions, and Stock Prices, *Journal of Finance* 76, 797–844.

Ungeheuer, Michael, and Martin Weber, 2022a, Diversification Benefit Neglect.

Ungeheuer, Michael, and Martin Weber, 2022b, The Frequency of Outperformance and Investment Decisions.