ADIA ⁝ Lab

WHY HAS FACTOR INVESTING FAILED?: THE ROLE OF SPECIFICATION ERRORS

Marcos López de Prado *
Vincent Zoonekynd §

ADIA Lab Research Paper Series, No. 7

First version (v0.1): November 1, 2023
Current version (v4.4): July 20, 2024

* Global Head, Quantitative Research & Development, Abu Dhabi Investment Authority (ADIA); Board Member, ADIA Lab; Professor of Practice, School of Engineering, Cornell University; Research Fellow, Applied Mathematics & Computational Research Department, Lawrence Berkeley National Laboratory. E-mail: marcos.lopezdeprado@adia.ae.
§ Quantitative Research & Development Lead, Abu Dhabi Investment Authority (ADIA); Research Affiliate, ADIA Lab.

The views expressed in this paper are the authors', and do not necessarily represent the opinions of the organizations they are affiliated with. We would like to thank our ADIA colleagues, especially Pascal Blanqué, Alexander Lipton, Anders Svennesen, and Jean-Paul Villain for their suggestions. The paper has also benefited from conversations with Patrick Cheridito (ETH Zürich), Gerald Cubbin (Invesco), Frank Fabozzi (EDHEC), Campbell Harvey (Duke University), Miguel Hernán (Harvard University), Guido Imbens (Stanford University), Alessia López de Prado Rehder (ETH Zürich), Riccardo Rebonato (EDHEC), Alessio Sancetta (Royal Holloway University of London), Luis Seco (University of Toronto), Horst Simon (ADIA Lab), and Josef Teichmann (ETH Zürich).

WHY HAS FACTOR INVESTING FAILED?: THE ROLE OF SPECIFICATION ERRORS

ABSTRACT

P-hacking is a well-understood cause of false positives in factor investing. A far less studied cause is factor model specification choices. We prove that specification errors cause factor strategies to underperform and potentially yield systematic losses, even if *all* risk premia remain constant and are estimated with the correct sign. The erratic performance of factor investing strategies is better explained by specification errors than by time-varying risk premia. The implication is that specification errors are more common and dangerous to factor investors than previously thought. We also show that standard econometric practices cause researchers to over-control for colliders, increasing the likelihood of adverse outcomes. To our knowledge, this is the first study that connects specification errors, factor model selection practices, underperformance and systematic losses through a causal mechanism. These findings challenge the scientific soundness and long-term profitability of the current (associational, casual, non-causal) multi-trillion-dollar factor investing industry. To overcome these pitfalls, academics and practitioners should rebuild the financial economics literature on the more scientifically rigorous grounds of causal factor investing.

Keywords: Causal inference, causal discovery, confounder, collider, factor investing, *p*-hacking, underperformance, systematic losses.

JEL Classification: G0, G1, G2, G15, G24, E44.
AMS Classification: 91G10, 91G60, 91G70, 62C, 60E.

TABLE OF CONTENTS

1. INTRODUCTION

Factor investing can be defined as the investment approach that aims to monetize the exposure to measurable risk characteristics (called "factors") that presumably explain differences in the performance of a set of securities. Its origins can be traced back to the seminal work of Fama and MacBeth [1973] and Schipper and Thompson [1981], among others. Since then, thousands of academic papers have claimed the discovery of hundreds of investment factors (Harvey et al. [2016]), propelling the growth of a multi-trillion-dollar industry. In 2019, J.P. Morgan estimated that over USD 2.5 trillion (more than 20 percent of the US equity market capitalization) was managed by quant-style funds (Neuberger Berman [2019]). BlackRock estimates that the factor investing industry managed USD 1.9 trillion in 2017, and it projected that amount would grow to USD 3.4 trillion by 2022 (BlackRock [2017]).

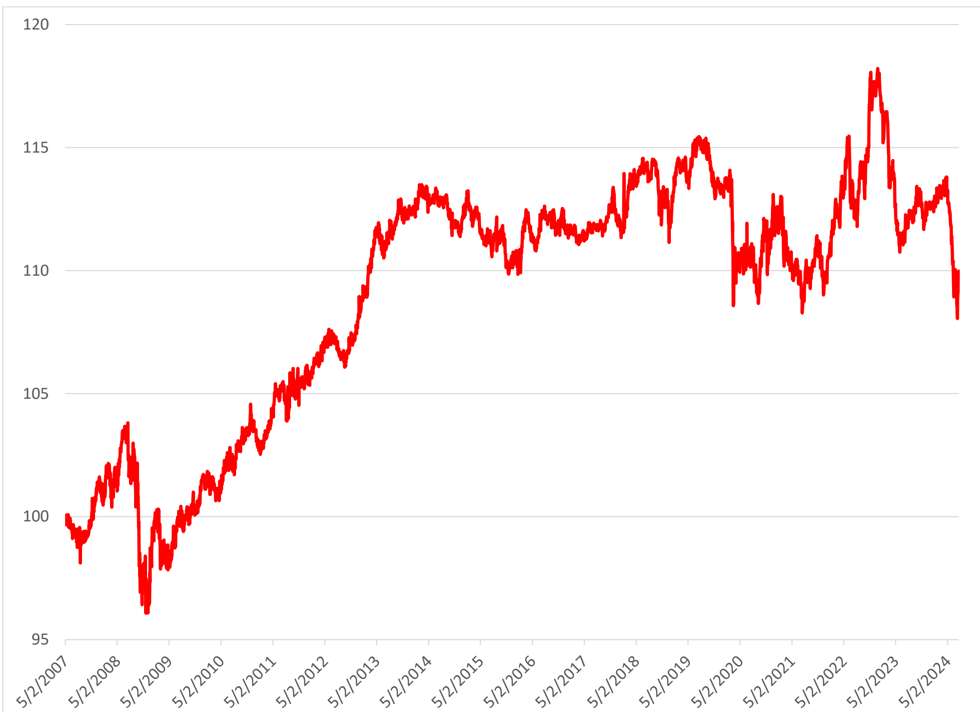

Figure 1 - Performance of US multi-factor equity strategies. Source: Bloomberg

Unfortunately for investors, these academically-endorsed investment products have failed to perform as expected. Figure 1 plots the performance of one of the broadest factor investing indices, the Bloomberg – Goldman Sachs Asset Management US Equity Multi-Factor Index (BBG code: BGSUSEMF <Index>). This Multi-Factor index tracks the long/short performance of the momentum, value, quality, and low-risk factors in U.S. stocks (Bloomberg [2021]). Its annualized Sharpe ratio from May 2, 2007 (the index's inception date) to July 19, 2024 (this paper's completion date) has been 0.17 (t-stat=0.69, *p*-value=0.25), and the average annualized excess return has been 0.62%. This performance does not include: (a) transaction costs; (b) market impact of order execution; (c) cost of borrowing stocks for shorting positions; (d) management and incentive fees. Including those costs would yield a negative total investment return. After almost 17 years of out-of-sample performance, factor investing's Sharpe ratio is statistically insignificant at any reasonable confidence level.

The primary goal of this paper is to elucidate why the performance of factor investing strategies has disappointed, and what can be done to correct this situation. Factor investing has failed to perform as expected, not for the lack of stringent peer-review (many of these failed factors have been published in the top journals of financial economics, including The Journal of Finance), but because the econometric canon used to make and peer-review factor claims is flawed. Until now, much of the blame has been placed on multiple testing and *p*-hacking (for example, see Bailey et al. [2014], and Harvey and Liu [2020]), and very little attention has been paid to the role of specification errors. We show how canonical econometric practices favor misspecified factor strategies that underperform (compared to the correctly specified model) and potentially yield systematic losses, even if all risk premia remain constant and are estimated with the correct sign. Unless the econometric canon is rewritten, these errors will not be addressed, and factor investors have no reason to expect that performance will improve in the future.

The rest of the paper is organized as follows: Section 2 outlines the contributions to the literature. Section 3 shows how factor model misspecification causes underperformance and systematic losses in factor investing. Section 4 proves that models that over-control for a collider always exhibit a higher R-squared, and under some conditions lower *p*-values. Section 5 exposes the misguided standard econometric practice of choosing a model's specification on the basis of its higher explanatory power. Section 6 explains the causal mechanism responsible for the underperformance and systematic losses delivered by factor investing. Section 7 summarizes our conclusions. The Appendix contains mathematical proofs of propositions used throughout the paper.

2. INNOVATIONS

When a factor is itself a portfolio, and therefore it is tradable, its risk premium can be simply estimated as the average excess return of that portfolio. The best-known example of a tradable factor is the market portfolio. In most cases, however, economic theories point to non-tradable factors, such as liquidity, and many variables associated with macroeconomic regimes, intangibles, or the business cycle. To estimate the risk premium of non-tradable factors, researchers typically implement a two-pass estimator (Jensen et al. [1972], and Fama and Macbeth [1973]). In the first pass, time series regressions estimate the exposures of securities to a factor, while controlling for all other risks to which the securities are exposed. In the second pass, cross-sectional regressions estimate the risk premium associated with exposure to the factor of interest, while controlling for all other exposures. All two-pass estimators are afflicted by two potential specification errors: (a) missing a confounder (also known as the omitted variable problem); and (b) including a collider (i.e., a variable that is correlated both with the right and left hand side variables in the regression). To address the specification error due to a missing confounder, Giglio and Xiu [2021] have proposed a three-pass method to estimate the risk premium of an observable factor, which is valid even when not all factors in the model are specified or observed. Virtually all factor investing papers and products predate the development of methods resilient to omitted factors, and hence likely suffer from confounder bias. The second type of specification error, due to the inclusion of a collider, remains an unsolved investment problem, as no econometric procedure is known to produce unbiased estimates of risk premia in that scenario.[1] This paper studies the effects of these

[1] Note that the inclusion of a collider is different from the inclusion of an irrelevant factor, because irrelevant factors are uncorrelated with the test asset returns. See Gospodinov et al. [2014] for a discussion of irrelevant factors.

two specification errors on the performance of factor investing strategies, and it proposes ways to address those errors.

Our contributions to the financial economics and econometrics literatures can be grouped in four themes. First, we show that the assumption of correct model specification is a defining feature that sets factor investing apart from other types of investment strategies, such as forecasting strategies. We study the effect of specification errors on the performance of factor and forecasting strategies. We prove that both types of strategies can underperform and yield systematic losses even if the risk premia remain constant and are estimated with the correct sign.

Second, several authors have argued that the erratic performance of factor investing strategies is due to time-varying risk premia (see, for example, Evans [1994], Anderson [2011], and Cochrane [2011]). Our findings show that this erratic performance is better explained (in simpler, more parsimonious terms) by specification errors, without requiring changes in the market's compensation for risk. The risk premia can remain constant, and specification errors will still cause factor investing strategies to perform erratically.

Third, we prove sufficient conditions for a misspecified factor model to exhibit a higher R-squared and lower *p*-values than the correctly specified model, despite producing biased parameter estimates.

Fourth, we describe a causal mechanism whereby the econometric canon enables *p*-hacking within a few trials, by favoring over-controlled (and possibly money-losing) models with higher R-squared and lower *p*-values than the correctly specified model. The implication is that the proliferation of spurious factors that Cochrane [2011] vividly described as "the factor zoo" may be more harmful to investors than previously thought, because misspecified canon-compliant factor models can yield systematic losses even if the risk premia remain constant and are estimated with the correct sign.

To our knowledge, this is the first study that connects specification errors, factor model selection practices, underperformance and systematic losses through a causal mechanism.

3. THE EFFECT OF SPECIFICATION ERRORS ON STRATEGY PERFORMANCE

Factor strategies decide whether a security i is bought or sold based on the value of $\hat{\beta}X_{i,t}$ at time t, where $X_{i,t}$ measures i's exposure at time t to a risk characteristic that the market rewards with an estimated linear premium $\hat{\beta}$, as determined by a cross-sectional factor model with $|\hat{\beta}| \gg 0$.[2] The goal of the analysis is not to predict the future return of i, but to establish whether the market rewards on average holding securities with exposure X (a common non-diversifiable rewarded

[2] The factor investing and econometrics literatures use different notations. Some factor investing authors use $\beta_{i,t}$ to denote the exposure of security i at time t to a rewarded risk characteristic. The exposure $\beta_{i,t}$ could be measured (e.g., book-to-price ratio) or estimated through a time series regression. In both cases, the premium λ_t associated with $\beta_{i,t}$ is estimated through a cross-sectional regression. See Ferson [2019, chapter 22] for a detailed explanation. In this paper we use the prevailing econometric notation, as it is more relevant to misspecification problems.

risk).[3] Estimating the risk premium β without bias requires that the factor model is correctly specified. For instance, a value strategy may buy a security i when its book-to-price ratio is high, and hedge other risks by selling a security j whose book-to-price ratio is low, where i and j are similar in everything but their different exposure to book-to-price ratio. Provided that the factor model is correctly specified and β's estimate is unbiased and sufficiently positive, the strategy profits from i's outperformance over j. However, should the factor model be misspecified, the parameter estimates will be biased ($E\left[\hat{\beta}|X\right] \not\approx \beta$), the strategy will have exposures other than X, and the investor will likely experience underperformance, including potentially systematic losses.

In contrast, forecasting strategies decide whether a security i is bought or sold based on the forecasted return of i at time t. These strategies do not attempt to profit from holding securities exposed to a rewarded risk characteristic while hedging other exposures. Instead, forecasting strategies attempt to monetize associations, regardless of their source. For instance, a pairs trading strategy (a type of statistical arbitrage) may buy or sell the spread i between two cointegrated securities, where $\hat{\varepsilon}_{i,t}$ measures the value of the residual of an error-correction model, a type of time series analysis (López de Prado and Leinweber [2012]). That strategy profits from the forecast that $\hat{\varepsilon}_{i,t}$ will converge to zero over time, even if the forecasting model is misspecified and the parameter estimates are biased.

The above explanation highlights that the assumption of correct model specification sets factor investing strategies apart, as other investment styles do not make decisions based on premia attributed to risk characteristics. The violation of this assumption (correct model specification) has practical consequences in terms of profitability that factor academics and practitioners need to understand.

Factor model misspecification has two primary consequences for investors: Type-1 errors (false strategies) and type-2 errors (missed strategies). Out of the type-1 errors, some of them result in systematic losses (a factor strategy with an expected negative return). The next section studies in detail the effect of factor model misspecification on strategy performance.

3.1. MISSING A CONFOUNDER

Consider a data-generating process with equations:

$$\begin{aligned} X &:= Z\delta + v \\ Y &:= X\beta + Z\gamma + u \end{aligned} \tag{1}$$

[3] Factor investing models do not typically lag the explanatory variables, so the explanatory and explained variables are observed at the same time. For example, see the equations in Tables 3-11 of Fama and French [1993], and equations (4)-(6) in Fama and French [2015]. This modelling choice makes it possible for reverse causation to occur.

where $\gamma \neq 0$, $\delta \neq 0$, and variables (u, v, Z) are independent and identically distributed as a standard Normal, $(u, v, Z) \sim N[0, I]$.[4] The causal graph for this process is displayed in Figure 2.[5] This is the graph that represents through nodes all the variables involved in the process, and through arrows the direction of dependences. In the language of causal inference, Z is called a confounder, because it influences both the cause (X) and the effect (Y), thus obfuscating the true effect of X on Y.[6] The arrow in green denotes the causal path from X to Y, and the arrows in red denote a non-causal path from X to Y through Z, i.e. $X \leftarrow Z \rightarrow Y$. In presence of a confounder, this is also called a backdoor path, because it has an arrow pointing into the cause (X).

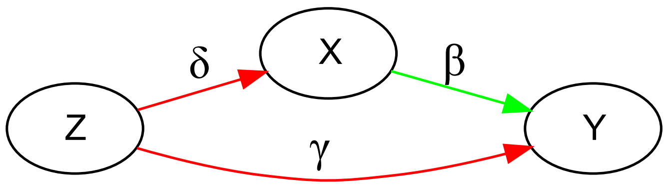

Figure 2 – Variable Z as confounder

The data-generating process is unknown to observers. With proper knowledge of the causal graph, observers would estimate the effect of X on Y by fitting the equation $Y = X\beta + Z\gamma + \varepsilon$. Controlling for confounder Z blocks the backdoor path, thus achieving that $E[Y|X, Z] = E[\hat{Y}|X, Z]$, i.e. the true value of Y (from the data-generating process) is on average equal to the value of Y expected by observers, and the parameters are estimated without bias, $E[\hat{\beta}|X, Z] = \beta$, $E[\hat{\gamma}|X, Z] = \gamma$.

Suppose that observers incorrectly attempt to estimate the causal effect of X on Y by fitting the equation $Y = X\beta + \varepsilon$ on a sample produced by the process. Then, the backdoor path remains unblocked, and the expected value of the estimated β is,

$$E[\hat{\beta}|X] = \beta + \frac{\gamma\delta}{1 + \delta^2} \tag{2}$$

See Section A.1 for a proof of the above statement. For appropriate values of (δ, γ), the confounder bias $\frac{\gamma\delta}{1+\delta^2}$ can be so large that it flips the sign of the estimated β (a phenomenon known as Simpson's paradox, see Pearl [2014]).[7] In this case, observers expect that the value of Y is

[4] Serial independence, linearity, stationarity and Normality are highly idealized assumptions commonly adopted in the factor investing literature. One goal of this paper is to prove that, even under this ideal scenario, factor strategies are structurally flawed and will not perform as designed.
[5] In this paper, we focus on causal graphs that are acyclic. Acyclicality allows us to represent joint distributions as the product of conditional probabilities between ancestors and descendants only. This choice does not limit the analysis to less realistic scenarios, because acyclicality can usually be imposed through: (a) redefinition of variables, e.g. adding lagged variables as causes; and (b) targeted sampling frequency, such that the effect on Y is measured before feedback loops come into play.
[6] For a general introduction to causal inference, causal discovery and do-calculus, see Neal [2020], Glymour et al. [2019], Pearl [2009], Pearl et al. [2016]. For an introduction to causal inference in finance, see López de Prado [2023].
[7] This bias in the parameter estimation is not in contradiction with the least-squares property whereby $E[Y - E[\hat{Y}|X]] = 0$.

$$E[\hat{Y}|X] = XE[\hat{\beta}|X] \tag{3}$$

whereas the true value (from the data-generating process) of Y has expectations

$$\begin{aligned} E[Y|X] &= X\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right) \\ E[Y|Z] &= Z(\gamma + \beta\delta) \\ E[Y|X,Z] &= X\beta + Z\gamma \end{aligned} \tag{4}$$

3.1.1. CONSTANT CORRELATIONS

From an investment standpoint, variables X and Z are factor exposures, because both are a cause of Y. However, not all system parameters are risk premia. This can be seen in the correctly specified model, $E[\hat{Y}|X,Z] = XE[\hat{\beta}|X,Z] + ZE[\hat{\gamma}|X,Z] = X\beta + Z\gamma = E[Y|X,Z]$, thus Y is only influenced by two parameters (β, γ). Therefore, (β, γ) are risk premia, and δ is not a risk premium. Parameter δ is a confounding effect that may bias the estimate of risk premium β (note that the bias is zero when $\delta = 0$). Next, we study the implications of under-controlling for a confounder on the performance of factor strategies and forecasting strategies, when system parameters (β, γ, δ) do not change, hence all correlations remain constant.

3.1.1.1. FACTOR STRATEGIES

Consider an investor who wishes to harvest the risk premia offered by securities with returns Y, where the data-generating process is the one described in Section 3.1. With the correctly specified factor model, the estimates of the factor premia are unbiased, and the factor strategy controls for X and confounder Z. A factor strategy that sizes positions by $E[\hat{Y}|X,Z]$ produces expected returns

$$\begin{aligned} E[YE[\hat{Y}|X,Z]|X,Z] &= E[Y(X\beta + Z\gamma)|X,Z] \\ &= E[Y|X,Z](X\beta + Z\gamma) \\ &= (X\beta + Z\gamma)^2 \geq 0 \end{aligned} \tag{5}$$

The above equation shows that the factor strategy based on correctly specified factor model extracts the risk premia β and γ, as it was designed to do. The equation excludes the confounding effect δ, because controlling for Z blocks the backdoor path. Alternatively, an investor with a correctly specified model may decide to harvest only X, for instance if Z is a confounding but non-investable variable. In this case, a factor strategy that sizes positions by $E[\hat{Y}|X]$ produces expected returns

$$\begin{aligned} E[YE[\hat{Y}|X]|X] &= E\left[(X\beta + Z\gamma + u)X\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)|X\right] \\ &= (\beta X + \gamma E[Z|X] + E[u|X])X\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right) \\ &= X\left(\beta X + \gamma\frac{\delta}{1+\delta^2}X + 0\right)X\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right) \\ &= X^2\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)^2 \geq 0 \end{aligned} \tag{6}$$

Now the confounding effect δ is also present, and the strategy failed to produce as expected return $(X\beta)^2$. This is an important result, because it highlights that using a factor model with the correct specification is not enough: the investment strategy must control for the confounder when choosing securities and sizing positions. It also shows that the standard industry practice of creating single-factor funds is misguided, as those funds are likely exposed to the confounding effect δ. Furthermore, pooling multiple single-factor strategies into one fund does not eliminate this confounding effect δ.

With the incorrect model specification $Y = X\beta + \varepsilon$, the investor aims to harvest X, but without controlling for Z's confounding effect on Y. If positions are sized by $E\left[\hat{Y}|X\right]$, then the expected returns of this strategy are

$$\begin{aligned} E[YE[\hat{Y}|X]|X] &= E\left[YX\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)|X\right] \\ &= E[Y|X]\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)X \\ &= X^2\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)^2 \geq 0 \end{aligned} \tag{7}$$

Again, the confounding effect δ is present, and the strategy failed to produce expected return $(X\beta)^2$. A factor strategy that under-controls for a confounder produces the same performance as a factor strategy that used the correct model specification but failed to harvest the confounder.

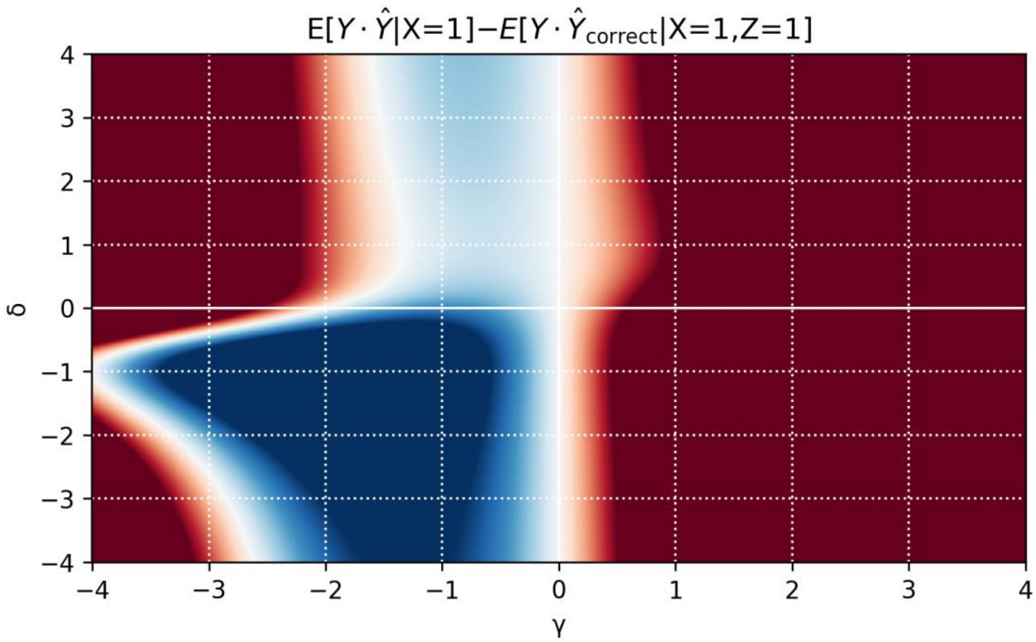

Figure 3 – Underperformance (red) and outperformance (blue) of a factor strategy that under-controls for a confounder, where $X = 1$, $Z = 1$, $\beta = 1$ and different combinations of (γ, δ)

The conclusion from the above equations is that under-controlling for a confounder causes factor strategies to underperform, for two reasons: (a) the correctly specified model allows the strategy to harvest two factors, whereas the under-controlled model limits the harvesting to only one; and (b) the non-causal association created by the open backdoor path cannot compensate for leaving Z unharvested. For instance, given $\beta = 1$, $X = 1$, $Z = 1$ and $\gamma > 0$, Figure 3 shows the underperformance (colored in red) of factor strategies that under-control for a confounder, relative to correctly specified factor strategies. As an example, for $(X, Z, \beta, \gamma, \delta) = (1,1,1,3,1)$, then $E\left[YE\left[\hat{Y}\middle|X,Z\right]\middle|X,Z\right] = 16$ while $E\left[YE\left[\hat{Y}\middle|X\right]\middle|X\right] = 6.25$, even though $E\left[\hat{\beta}\middle|X\right] = 2.5$ is estimated with the correct sign ($\beta = 1$). The second and third quadrants contain regions where under-controlling for a confounder causes factor strategies to outperform (colored in blue), however the reason is the misuse of the correctly specified strategy: for $\gamma < 0$, a rational investor would not target securities with $Z = 1$. Those two quadrants turn red for $(X, Z) = (1, -1)$. We merely report results of $\gamma < 0$ for completeness.

3.1.1.2. FORECASTING STRATEGIES

Consider an investor who wishes to monetize the forecasting power of a correctly specified factor model. In order to do so, he develops a strategy that buys or sells securities based solely on the forecasted returns $E\left[\hat{Y}|X,Z\right]$, regardless of that security's factor exposure X. Since the investor is not targeting factor exposures, there is no conditioning on X or Z. A forecasting strategy that sizes positions by $E\left[\hat{Y}|X,Z\right]$ produces (unconditional) expected returns

$$\begin{aligned}
E[YE[\hat{Y}|X,Z]] \quad &= E[Y(X\beta + Z\gamma)] \\
&= \beta E[XY] + \gamma E[YZ] \\
&= \beta(\beta(1+\delta^2) + \delta\gamma) + \gamma(\gamma + \beta\delta) \\
&= \beta^2 + \beta^2\delta^2 + 2\beta\delta\gamma + \gamma^2 \\
&= \beta^2 + (\beta\delta + \gamma)^2 \geq 0
\end{aligned} \tag{8}$$

However, with the incorrect model specification $Y = X\beta + \varepsilon$, the expected returns of a forecasting strategy that sizes positions by $E\left[\hat{Y}|X\right]$ are

$$\begin{aligned}
E[YE[\hat{Y}|X]] \quad &= E\left[YX\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)\right] \\
&= (\beta(1+\delta^2) + \gamma\delta)\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right) \\
&= (1+\delta^2)\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)^2 \geq 0
\end{aligned} \tag{9}$$

The conclusion from the above equations is that under-controlling for a confounder always causes forecasting strategies to underperform.

3.1.1.3. CONSTANT CORRELATIONS SCORECARD

Figure 4 summarizes the earlier results, which debunk two widespread myths: (a) that specification errors are not dangerous to factor investors as long as correlations remain constant and the factor model estimates β with the correct sign; and (b) that specification errors do not impact the performance of forecasting strategies. In presence of a confounder Z, an under-controlling model will cause the underperformance of both, factor strategies and forecasting strategies.

Confounder Case	Correct Specification	Under-Controlling
Factor Strategy	$(X\beta + Z\gamma)^2$	$X^2\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)^2$
Forecasting Strategy	$\beta^2 + (\beta\delta + \gamma)^2$	$(1+\delta^2)\left(\beta + \frac{\gamma\delta}{1+\delta^2}\right)^2$

Figure 4 – Scorecard of performance for factor and forecasting strategies in case of a confounder with constant correlations

3.1.2. CONSTANT RISK PREMIA WITH CONFOUNDING EFFECT SHIFT

Consider the data-generating process described in Section 3.1. After a sufficient sample size has been generated, an investor can estimate parameters with sufficient confidence, and design an investment strategy. Suppose that, sometime afterwards, the confounding effect shifts from δ_0 to δ_1, while keeping constant the risk premia (β, γ). For a long period of time, there are not enough observations to detect that shift, thus the investor continues to use the original parameter estimates. Next, we study whether that shift in $\delta_0 \neq \delta_1$, while keeping constant the risk premia (β, γ), can be enough to cause systematic losses.

3.1.2.1. FACTOR STRATEGIES

After the shift in parameters, a factor strategy based on the correctly specified model that sizes positions by $E_0\left[\hat{Y}|X,Z\right]$ delivers expected returns

$$\begin{aligned} E_1[YE_0[\hat{Y}|X,Z]|X,Z] &= E_1[Y(X\beta + Z\gamma)|X,Z] \\ &= E_1[Y|X,Z](X\beta + Z\gamma) \\ &= E_1[X\beta + Z\gamma + u|X,Z](X\beta + Z\gamma) \\ &= (X\beta + Z\gamma)^2 \geq 0 \end{aligned} \tag{10}$$

This equation does not include δ_1, therefore the performance of the factor strategy based on the correctly specified model is not affected by the change in $\delta_0 \neq \delta_1$.

In contrast, for a factor strategy based on a model that under-controls for a confounder and sizes positions by $E_0\left[\hat{Y}|X\right]$, the expected returns after the parameter shift are

$$\begin{aligned} E_1[YE_0[\hat{Y}|X]|X] &= E_1\left[Y\left(\beta + \frac{\gamma\delta_0}{1+\delta_0^2}\right)X|X\right] \\ &= E_1[Y|X]\left(\beta + \frac{\gamma\delta_0}{1+\delta_0^2}\right)X \\ &= X^2\left(\beta + \frac{\gamma\delta_1}{1+\delta_1^2}\right)\left(\beta + \frac{\gamma\delta_0}{1+\delta_0^2}\right) \end{aligned} \tag{11}$$

For appropriate combinations of real values of the above variables, the strategy yields systematic losses, i.e. $E_1\left[YE_0\left[\hat{Y}|X\right]\middle|X\right] < 0$. For example, the strategy yields systematic losses under $(\beta, \gamma, \delta_0, \delta_1) = (1,3,1,-1)$, even though the risk premia remain constant and $E_0\left[\hat{\beta}|X\right] = 2.5$ is estimated with the correct sign ($\beta = 1$).

3.1.2.1. FORECASTING STRATEGIES

After the shift in parameters, a forecasting strategy based on the correctly specified model that sizes positions by $E_0\left[\hat{Y}|X,Z\right]$ delivers expected returns

$$\begin{aligned} E_1[YE_0[\hat{Y}|X,Z]] &= E_1[Y(X\beta + Z\gamma)] \\ &= \beta E_1[XY] + \gamma E_1[YZ] \\ &= \beta(\beta(1+\delta_1^2) + \delta_1\gamma) + \gamma(\gamma + \beta\delta_1) \\ &= \beta^2 + (\beta\delta_1 + \gamma)^2 \geq 0 \end{aligned} \tag{12}$$

This equation does include δ_1, therefore the performance of the forecasting strategy based on the correctly specified model is affected by the change in $\delta_0 \neq \delta_1$, however not to the point of experiencing systematic losses.

In contrast, for a forecasting strategy based on a model that under-controls for a confounder and sizes positions by $E_0\left[\hat{Y}|X\right]$, the expected returns after the parameter shift are

$$\begin{aligned} E_1[YE_0[\hat{Y}|X]] &= E_1[YX]\left(\beta + \frac{\gamma\delta_0}{1+\delta_0^2}\right) \\ &= (\beta(1+\delta_1^2) + \gamma\delta_1)\left(\beta + \frac{\gamma\delta_0}{1+\delta_0^2}\right) \\ &= (1+\delta_1^2)\left(\beta + \frac{\gamma\delta_1}{1+\delta_1^2}\right)\left(\beta + \frac{\gamma\delta_0}{1+\delta_0^2}\right) \end{aligned} \tag{13}$$

For appropriate combinations of real values of the above variables, the strategy yields systematic losses, i.e. $E_1[YE_0[\hat{Y}|X]] < 0$. For example, the strategy yields systematic losses under $(\beta, \gamma, \delta_0, \delta_1) = (1,3,1,-1)$, even though the risk premia remain constant and $E_0\left[\hat{\beta}|X\right] = 2.5$ is estimated with the correct sign ($\beta = 1$).

From the above results, we conclude that a shift in the confounding effect δ: (a) can cause systematic losses in factor and forecasting strategies that fail to control for a collider; (b) does not affect the performance of correctly specified factor strategies, which remains non-negative; and (c) affects the performance of correctly specified factor strategies, however not to the point of yielding systematic losses.

3.2. CONTROLLING FOR A COLLIDER

Consider a data-generating process with equations:

$$\begin{aligned} Y &:= X\beta + u \\ Z &:= Y\gamma + X\delta + v \end{aligned} \tag{14}$$

where $\gamma \neq 0$, $\delta \neq 0$, and variables (u, v, X) are independent and identically distributed as a standard Normal, $(u, v, X) \sim N[0, I]$. The causal graph for this process is displayed in Figure 5. In the language of causal inference, Z is a collider, because it is influenced by both the cause (X) and the effect (Y). The non-causal path from X to Y through Z (i.e., $X \rightarrow Z \leftarrow Y$) is blocked, unless the model controls for collider Z.

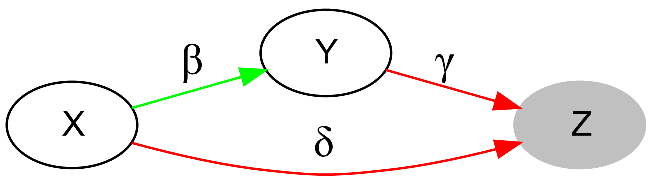

Figure 5 – Variable Z as a controlled collider

The data-generating process is unknown to observers. With proper knowledge of the causal graph, observers would estimate the effect of X on Y (green arrow in Figure 5) by fitting the equation $Y = X\beta + \varepsilon$. With that model specification, $E[Y|X] = E\left[\hat{Y}|X\right]$, i.e. the true value of Y (from the data-generating process) is on average equal to the value of Y expected by observers, and the parameter is estimated without bias, $E\left[\hat{\beta}|X\right] = \beta$.

Suppose that observers incorrectly attempt to estimate the causal effect of X on Y by fitting the equation $Y = X\beta + Z\theta + \varepsilon$ on a sample produced by the process. Controlling for collider Z (shaded node in Figure 5) unblocks the non-causal path (red arrows in Figure 5). Then, β and θ are estimated with a bias, with expected values,[8]

$$E\left[\hat{\beta}|X, Z\right] = \frac{\beta - \delta\gamma}{1 + \gamma^2} \tag{15}$$

$$E\left[\hat{\theta}|X, Z\right] = \frac{\gamma}{1 + \gamma^2} \tag{16}$$

See Section A.2 for a proof of the above statements. Parameters (δ, γ) may bias the estimate of risk premium β (note that the bias is zero when $(\delta, \gamma) = (0,0)$). For appropriate values of (δ, γ), the bias can be so large that it flips the sign of the estimated β and θ (a phenomenon that biostatisticians call Berkson's fallacy, named after the physician who discovered it, see Berkson [1946]).[9] Accordingly, observers expect that the value of Y is

[8] Since θ is absent from the correctly specified model, its unbiased estimate should be $E\left[\hat{\theta}|X, Z\right] = 0$.

[9] This bias in the parameter estimation is not in contradiction with the least-squares property whereby $E[Y - E[\hat{Y}|X, Z]] = 0$.

$$
\begin{aligned}
E[\hat{Y}|X] &= E[XE[\hat{\beta}|X,Z] + ZE[\hat{\theta}|X,Z]|X] \\
&= XE[\hat{\beta}|X,Z] + E[Z|X]E[\hat{\theta}|X,Z] \\
&= X\frac{\beta - \delta\gamma}{1+\gamma^2} + X\frac{\gamma}{1+\gamma^2}(\beta\gamma + \delta) \\
&= X\beta
\end{aligned} \tag{17}
$$

$$
E[\hat{Y}|X,Z] \quad = X\frac{\beta - \delta\gamma}{1+\gamma^2} + Z\frac{\gamma}{1+\gamma^2} \tag{18}
$$

whereas the true value (from the data-generating process) of Y has expectation

$$
E[Y|X] = X\beta \tag{19}
$$

Notably, the model that over-controls for a collider estimates Y with lower variance of the error than the correctly specified model, i.e. $Var[\varepsilon] < Var[u]$ (see Section A.3.1 for a proof).

3.2.1. CONSTANT CORRELATIONS

From an investment standpoint, only variable X is a factor exposure, because only X is a cause of Y. The correctly specified model is $E\left[\hat{Y}\middle|X\right] = XE\left[\hat{\beta}\middle|X\right] = X\beta = E[Y|X]$, thus Y is only influenced by parameter β. Therefore, only β is a risk premium, and (δ, γ) are not risk premia. Next, we study the implications of over-controlling for a collider on the performance of factor strategies and forecasting strategies, when system parameters (β, γ, δ) do not change, hence all correlations remain constant.

3.2.1.1. FACTOR STRATEGIES

Consider an investor who wishes to harvest the risk premia offered by securities with returns Y, where the data-generating process is the one described in Section 3.2. With the correctly specified factor model, the factor strategy that sizes positions by $E\left[\hat{Y}\middle|X\right]$ produces expected returns

$$
\begin{aligned}
E\left[YE\left[\hat{Y}|X\right]\middle|X\right] &= E[Y|X]X\beta \\
&= E[X\beta + u|X]X\beta \\
&= (X\beta)^2 \geq 0
\end{aligned} \tag{20}
$$

The above equation shows that the factor strategy based on the correctly specified factor model extracts the only risk premium, β, as designed. The collider parameters (γ, δ) are absent from the equation.

However, with the incorrect model specification $Y = X\beta + Z\theta + \varepsilon$, the investor attempts to design a factor strategy that profits from securities' exposures X and Z, instead of from the only factor exposure, X. One problem with this attempt is that collider Z is a function of Y, and not the other way around. This model specification is an instance of reverse causation. By the time Z is known and it is possible to condition on its value, the value of Y has already been set. In other words, it is not possible to estimate $E[\hat{Y}|X,Z]$ *before* the value of Y has been set. If the investor knew that Z is a function of Y, he could at least condition on $E\left[\hat{Z}\middle|X\right] = X(\beta\gamma + \delta)$, but this is inconsistent with the investor's (false) assumption that Z is a cause of Y. Therefore, attempting to condition on a value of Z before the value of Y is set is tantamount to conditioning on a variable $\tilde{Z}$ that adds no information beyond what is known when the value of X is set. Formally, $E[\cdot\,|X,\tilde{Z}] = E[\cdot\,|X]$, thus $E[u\big|X,\tilde{Z}] = E[u|X] = 0$. The expected returns of a strategy that sizes positions by $E\big[\hat{Y}|X,\tilde{Z}\big]$ are

$$\begin{aligned} E[YE[\hat{Y}|X,\tilde{Z}]|X,\tilde{Z}] \quad &= E[(X\beta+u)(E[\hat{\beta}|X,Z]X+E[\hat{\theta}|X,Z]\tilde{Z})|X,\tilde{Z}] \\ &= (X\beta+E[u|X,\tilde{Z}])(E[\hat{\beta}|X,Z]X+E[\hat{\theta}|X,Z]\tilde{Z}) \\ &= (X\beta+E[u|X,\tilde{Z}])\frac{\beta X+\gamma(\tilde{Z}-\delta X)}{1+\gamma^2} \\ &= X\beta\frac{X\beta+\gamma(\tilde{Z}-\delta X)}{1+\gamma^2} \end{aligned} \tag{21}$$

This factor strategy has failed to achieve the intended expected return, $(X\beta+Z\theta)^2$, which would have been achieved if Z had been a confounder or an independent cause of Y. Controlling for collider Z has unblocked the non-causal path $X \rightarrow Z \leftarrow Y$, with the effect of creating a non-causal association between X and Y. Accordingly, now the collider parameters (γ,δ) appear in the equation for expected strategy returns, despite of not being risk premia. Even if $\beta=0$, this over-controlled factor strategy is still exposed to that non-causal association.

In the above equation, we have assumed that $E[u|X,\tilde{Z}]=0$, however this is not a critical assumption. Regardless of the value of $E[u|X,\tilde{Z}]$, the conclusion holds that for appropriate combinations of real values of $\left(X,\tilde{Z},\beta,\gamma,\delta\right)$, the strategy yields systematic losses, i.e. $E[YE[\hat{Y}|X,\tilde{Z}]|X,\tilde{Z}]<0$. For example, the strategy yields systematic losses under $\left(X,\tilde{Z},\beta,\gamma,\delta\right)=$ (1,1,1,1,3).

3.2.1.2. FORECASTING STRATEGIES

With the correctly specified factor model, a forecasting strategy that sizes positions by $E[\hat{Y}|X]$ has expected returns

$$\begin{aligned} E[YE[\hat{Y}|X]] \quad &= E[(X\beta+u)X\beta] \\ &= \beta^2E[X^2]+\beta E[Xu] \\ &= \beta^2 \geq 0 \end{aligned} \tag{22}$$

Let us turn our attention to a model that over-controls for a collider Z. As argued earlier, attempting to condition on a value of Z before the value of Y is set is tantamount to conditioning on a variable $\tilde{Z}$ that adds no information beyond what is known when the value of X is set. For instance, this would be the case if $\tilde{Z}$ represents the previous value of Z, in which case $E[Y\tilde{Z}]=0$. Then, the forecasting strategy that sizes positions by $E[\hat{Y}|X,\tilde{Z}]$, has expected returns

$$
\begin{aligned}
E[YE[\hat{Y}|X,\tilde{Z}]] &= E[(X\beta+u)(E[\hat{\beta}|X,Z]X+E[\hat{\theta}|X,Z]\tilde{Z})] \\
&= E[(X\beta+u)E[\hat{\beta}|X,Z]X]+E[(X\beta+u)E[\hat{\theta}|X,Z]\tilde{Z}] \\
&= \beta E[\hat{\beta}|X,Z]E[X^2]+E[\hat{\beta}|X,Z]E[Xu]+E[\hat{\theta}|X,Z]E[Y\tilde{Z}] \\
&= \beta\frac{\beta-\gamma\delta}{1+\gamma^2}+\frac{\gamma}{1+\gamma^2}E[Y\tilde{Z}] \\
&= \beta\frac{\beta-\gamma\delta}{1+\gamma^2}
\end{aligned} \tag{23}
$$

In the above equation, we have assumed that $E[Y\tilde{Z}]=0$, however this is not a critical assumption. Regardless of the value of $E[Y\tilde{Z}]$, the conclusion holds that for appropriate combinations of real values of (β,γ,δ), the strategy yields systematic losses, i.e. $E[YE[\hat{Y}|X,\tilde{Z}]]<0$. For example, the strategy yields systematic losses under $(\beta,\gamma,\delta)=(1,1,2)$.

3.2.1.3. CONSTANT CORRELATIONS SCORECARD

Figure 6 summarizes the above results.

Collider Case	Correct Specification	Over-Controlling
Factor Strategy	$(X\beta)^2$	$X\beta\frac{X\beta+\gamma(\tilde{Z}-\delta X)}{1+\gamma^2}$
Forecasting Strategy	β^2	$\beta\frac{\beta-\gamma\delta}{1+\gamma^2}$

Figure 6 – Scorecard of performance for factor and forecasting strategies in case of a collider with constant correlations

These results evidence that over-controlling for a collider opens the possibility for systematic losses, even if correlations remain constant and the risk premium is estimated with the correct sign. In contrast, the downside of under-controlling for a confounder was limited to underperforming the strategy based on the correctly specified model, without the possibility of systematic losses (see Figure 4). The implication is that over-controlling for a collider is more dangerous to investors than under-controlling for a confounder.

3.2.2. CONSTANT RISK PREMIUM WITH COLLIDER PARAMETERS SHIFT

Consider the data-generating process described in Section 3.2. After a sufficient sample size has been generated, an investor can estimate parameters with sufficient confidence, and design an investment strategy. Suppose that, some time afterwards, the collider parameters shift from (γ_0,δ_0) to (γ_1,δ_1), while keeping constant the risk premium β. For a long period of time, there are not enough observations to detect that shift, thus the investor continues to use the original parameter estimates. Next, we study whether that shift in $\gamma_0\neq\gamma_1$ or $\delta_0\neq\delta_1$, while keeping constant the only risk premium β, can be enough to cause systematic losses.[10]

[10] Strictly speaking, a shift in the confounding effect or collider parameters could increase profitability, however investors are generally more concerned about unforeseen losses than they are about unforeseen profits.

3.2.2.1. FACTOR STRATEGIES

After the shift in parameters, a factor strategy based on the correctly specified model that sizes positions by $E_0\left[\hat{Y}|X\right]$ delivers expected returns

$$\begin{aligned} E_1\left[YE_0\left[\hat{Y}|X\right]\middle|X\right] &= E_1[Y|X]X\beta \\ &= E_1[X\beta+u|X]X\beta \\ &= (X\beta)^2 \geq 0 \end{aligned} \tag{24}$$

This equation does not include γ_1 or δ_1, therefore the shift in collider parameters has no effect on the performance of a factor strategy based on the correctly specified model. Intuitively, the performance of the correctly specified strategy is unaffected by a shift in collider parameters because: (a) the collider parameters have no effect on neither X or Y; and (b) the estimates of the correctly specified model are not influenced by the collider parameters.

In contrast, for a factor strategy based on a model that over-controls for a collider and sizes positions by $E_0\left[\hat{Y}|X,\tilde{Z}\right]$, the expected returns after the parameter shift are

$$\begin{aligned} E_1[YE_0[\hat{Y}|X,\tilde{Z}]|X,\tilde{Z}] &= E_1[(X\beta+u)(XE_0\left[\hat{\beta}|X,\tilde{Z}\right]+\tilde{Z}E_0\left[\hat{\theta}|X,\tilde{Z}\right])|X,\tilde{Z}] \\ &= (X\beta+E_1[u|X,\tilde{Z}])(XE_0\left[\hat{\beta}|X,\tilde{Z}\right]+\tilde{Z}E_0\left[\hat{\theta}|X,\tilde{Z}\right]) \\ &= X\beta\frac{X\beta+\gamma_0(\tilde{Z}-\delta_0X)}{1+\gamma_0^2} \end{aligned} \tag{25}$$

This is the same result that was derived under constant correlation (Section 3.2.1.1). The reason is, a change in the collider parameters affects Z, but it has no effect on $\tilde{Z}$. Like before, for appropriate combinations of real values of the above variables, the strategy yields systematic losses, i.e. $E_1[YE_0[\hat{Y}|X,\tilde{Z}]|X,\tilde{Z}] < 0$.

3.2.2.2. FORECASTING STRATEGIES

After the shift in parameters, a forecasting strategy based on the correctly specified model that sizes positions by $E_0\left[\hat{Y}|X\right]$ delivers expected returns

$$\begin{aligned} E_1[YE_0[\hat{Y}|X]] &= E_1[(\beta X+u)\beta X] \\ &= \beta^2E_1[X^2]+\beta E_1[uX] \\ &= \beta^2 \geq 0 \end{aligned} \tag{26}$$

This equation does not include γ_1 or δ_1, therefore the shift in collider parameters has no effect on the performance of a forecasting strategy based on the correctly specified model.

However, for a forecasting strategy based on a model that over-controls for a collider and sizes positions by $E_0\left[\hat{Y}|X,\tilde{Z}\right]$, the expected returns after the parameter shift are

$$
\begin{aligned}
E_1[YE_0[\hat{Y}|X,\tilde{Z}]] &= E_1[Y(XE_0[\hat{\beta}|X,\tilde{Z}] + \tilde{Z}E_0[\hat{\theta}|X,\tilde{Z}])] \\
&= E_0[\hat{\beta}|X,\tilde{Z}]E_1[XY] + E_0[\hat{\theta}|X,\tilde{Z}]E_1[Y\tilde{Z}] \\
&= \beta\frac{\beta - \gamma_0\delta_0}{1+\gamma_0^2}
\end{aligned}
\tag{27}
$$

This is the same result that was derived under constant correlation (Section 3.2.1.2). The reason is, a change in the collider parameters affects Z, but it has no effect on $\tilde{Z}$. Like before, for appropriate combinations of real values of the above variables, the strategy yields systematic losses, i.e. $E_1[YE_0[\hat{Y}|X,\tilde{Z}]] < 0$.

From the above results, we conclude that a shift in the collider parameters (γ, δ): (a) does not affect the performance of strategies based on the correctly specified model, whether factor strategies or forecasting strategies, which remains non-negative; and (b) does not affect the performance of strategies based on a model that over-controls for a collider, whether factor strategies or forecasting strategies, which can be negative. The intuition behind this conclusion is that a change in collider parameters has no effect on Y, hence it does not impact the strategy's performance, unlike in the confounder case.

3.3. ROBUSTNESS AS A VIRTUE OF CAUSALITY

The results derived in Sections 3.1 and 3.2 allow us to draw three important conclusions. First, even if all correlations remain constant, misspecified factor and forecasting strategies: (a) underperform the correctly specified strategy in the confounder case; and (b) can yield systematic losses in the collider case. This is true regardless of whether the risk premia were estimated with the correct sign.

Second, the performance of factor strategies that under-control for a confounder is not robust to parameter shift, and these strategies can yield systematic losses, even if the risk premia remain constant and are estimated with the correct sign.

Third, the performance of correctly specified factor and forecasting strategies is robust to parameter shifts. The performance of correctly specified factor strategies is unaffected by changes in both, the confounding effect and the collider parameters. The performance of correctly specified forecasting strategies is unaffected by changes in the collider parameters, but it is impacted (positively or negatively) by changes in the confounding effect, although not to the point of yielding systematic losses. Accordingly, robustness is neither a virtue of a model's complexity nor simplicity, but of the causal content injected by the modeler.

Several authors have argued that the erratic performance of factor investing strategies is due to time-varying risk premia (see, for example, Evans [1994], Anderson [2011], Cochrane [2011]). Our findings show that this erratic performance is better explained (in simpler, more parsimonious terms) by specification errors, without requiring changes in the market's compensation for risk, or changes in investor behavior. The risk premia can remain constant, and specification errors will still cause factor investing strategies to perform erratically.

Factor strategies often involve only a handful of parameters, but even then the risk of over-controlling for a collider is high. The next sections analyze how the econometric canon favors over-controlled models, which in turn explains the underperformance and systematic losses experienced by factor investors.

4. EXPLANATORY POWER AND P-VALUES UNDER OVER-CONTROLLING

In statistics, the explanatory power of a model indicates how well that model fits a set of observations. By far the most common way to express a factor model's explanatory power is through its R-squared (also known as its coefficient of determination), or its adjustment for the number of variables, known as the adjusted R-squared.

Section A.3 contains proofs to the following statements: (a) the R-squared of a factor model that over-controls for a collider always exceeds the R-squared of the correct model; (b) given enough observations, the adjusted R-squared of a factor model that over-controls for a collider exceeds the adjusted R-squared of the correct model; and (c) when the collider effects are strong and of similar magnitude, or the causal effect is weak, the *p*-values of estimated coefficients in a factor model that over-controls for a collider are smaller than those of the correct model.

In the case of the adjusted R-squared, the number of observations (T) needed for choosing the over-controlled model is $|\gamma| > \frac{1}{\sqrt{T-3}}$, where $|\gamma|$ is a function of the strength of the collider (relative to the standard Gaussian noise). For instance, for the value of γ used in Section 3.2.1.1's example ($\gamma = 0.7$), the over-controlled model has a higher adjusted R-squared than the correct model when $T > 5$. Like the adjusted R-squared, alternative associational methods designed to penalize model complexity fail to address this problem, because only causal approaches can target the specific issue of controlling for a collider.

In the case of *p*-values, false positives are more likely when the collider effects in an over-controlled model are sufficiently strong and of similar magnitude ($|\delta| \gg 0$, $|\gamma| \gg 0$, $|\delta| \approx |\gamma|$) or when the causal effect is weak ($|\beta| \approx 0$). The implication is that the presence of colliders facilitates *p*-hacking under rather general conditions.

5. SPECIFICATION-SEARCHING

The econometric canon, as taught in the most widely accepted textbooks, advocates choosing the model specification that yields the highest explanatory power. For example, Greene [2012, section 7.4] favors the use of R-squared for choosing among competing specifications when the model is intended for within-sample prediction (as it is common in factor investing papers), and favors the use of adjusted R-squared for choosing among specifications when the model is intended for forecasting.

For the reasons presented in Section 4, this approach to specification selection is particularly misleading in the presence of colliders. However, the econometric canon is seemingly ignorant of the problem of collider bias, discussed in Section 3.2. For instance, Greene [2012, p. 136] indicates that the downside of including irrelevant variables is "the reduced precision of the estimates," and that "[o]mitting variables from the equation seems generally to be the worse of the two errors." Furthermore, Greene [2012, p. 137] recommends a "general-to-simple" approach, where a researcher starts with a "kitchen sink regression, which contains every variable that might conceivably be relevant." Those relevant variables certainly include every possible collider, all but guaranteeing that the selected model is misspecified.

This state of confusion is not particular to Greene [2012], but a misunderstanding prevalent throughout the econometrics literature. The root problem is that econometricians tend to choose their specifications through associational arguments (such as R-squared), rather than through the discovery and analysis of the causal structure of the data-generating process. Chen and Pearl [2013] reviewed six of the most popular textbooks in econometrics, concluding that they "deviate significantly from modern standards of causal analysis".[11]

These flawed specification selection procedures have been applied to factor investing, as epitomized by two of the most influential publications in this field: Fama and French [1993] and Fama and French [2015]. These authors' argument for adding two factors to their initial three-factor model specification was that "the five-factor model performs better than the three-factor model when used to explain average returns." Unfortunately, as proven in Section 4, if their models include a collider, it is guaranteed that they have chosen the wrong specification, with the potentially dire consequences for investors discussed in Section 3. Mounting empirical evidence points in that direction. Sadeghi et al. [2024] apply the PC algorithm for causal discovery to the variables used in Fama and French [1993], and find that the HML factor may be a collider (see figure 8b in that publication). Similarly, Gu et al. [2023] present empirical evidence that the factor model regression in Fama and French [2015] is also misspecified, with the presence of at least one collider (market returns net of the risk-free rate).

The burden of proof is on factor researchers promoting an investment opportunity, to present theoretical and empirical evidence that the chosen variables do not include potential colliders. Absent that evidence, there is a significant risk that the model is over-controlled, and that investors are exposed to adverse outcomes.

6. THE CAUSAL MECHANISM FOR ADVERSE OUTCOMES

Virtually all factor researchers apply econometric procedures that fundamentally date back to the 1930s (when the Econometric Society and the Cowles commission were launched) and even the late 18th century (when Gauss developed the least squares method, see Stigler [1981]), before the development of causal inference. For the reasons explained earlier, the current econometric canon misleads researchers into selecting a certain class of misspecified factor models, in place of correctly specified models. Factor researchers (who selected the misspecified factor models) build investment portfolios that are more likely to underperform and yield systematic losses, to the benefit of more informed investors (who applied causal discovery to select correctly specified models). Consequently, factor investors and informed investors tend to take opposite sides of the same trade.

[11] There are remarkable exceptions to econometrics' state of causal confusion. In the year 2021, Guido Imbens and Joshua Angrist received (in conjunction with David Card) the Nobel Memorial Prize in Economics in recognition "for their methodological contributions to the analysis of causal relationships." These two authors have fought for decades to instill the rigor of causal inference into econometrics, or as Leamer [1983] put it, "take the con out of e*con*ometrics."

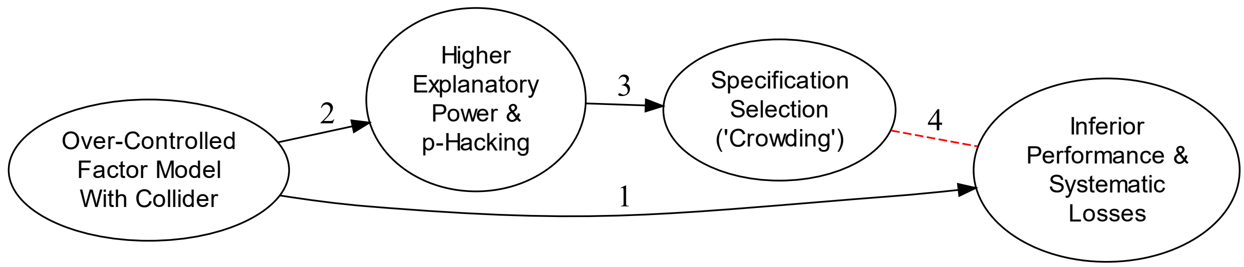

Figure 7 – The causal mechanism behind the failure of factor investing

Figure 7 outlines the causal mechanism that enables adverse outcomes in factor investing. In Section 3, we proved that factor and forecasting strategies based on factor models that over-control for a collider can yield systematic losses, even if correlations remain constant and the risk premium was estimated with the correct sign (link 1 in the causal mechanism). In Section 4, we proved that misspecified factor models that over-control for a collider exhibit a higher R-squared than correctly specified models. We also proved that, under general conditions, colliders make it easier to *p*-hack a factor (link 2 in the causal mechanism). In Section 5, we explained how the canonical approaches for specification-selection in econometrics favor over-controlled models, crowding out correctly specified models (link 3 in the causal mechanism). The combined effect is an increased association between selected (either published or deployed) factor investing strategies and underperformance, including systematic losses (link 4 in the causal mechanism).

There are two reasonable counter-arguments to Figure 7. First, factor investing papers usually publish long historical backtests with a positive expected return. Could those positive backtests indicate that those strategies' models are not over-controlled? The answer is no. As the False Strategy Theorem has proven (Bailey et al. [2014], López de Prado and Bailey [2021]), it is trivial to overfit a backtest to simulate any desired performance. For example, a researcher can run thousands of backtests by trying alternative investment universes and time periods, introducing arbitrary profit-taking and stop-out rules, or imposing *ad hoc* constraints to the portfolio optimization, to cite a few common variations. The expected Sharpe ratio of the best backtest rises with the number of trials, even if the true Sharpe ratio is zero. Historical backtests do not model the data-generating process, hence they cannot tell us much about whether the factor model is correctly specified. A better tool for assessing specification errors is causal discovery algorithms. A second counter-argument to Figure 7 is that some popular factor investing strategies have delivered positive out-of-sample performance for multiple years, before experiencing losses or flattening out. Could this be evidence that these strategies' models are not over-controlled? Again, the answer is no. The sheer number of factors, many of which are noisy (Akey et al. [2023]), has resulted in thousands of alternative implementations. Some of those implementations will perform well out-of-sample for a while (a statistical fluke). Also, a strategy's popularity can push up the prices of securities with targeted characteristics, driving up the strategy's live performance for several years (a self-fulfilling prophecy), before losses offset the initial trend. The factor investing industry continues to grow at sufficient pace to potentially mask its true expected returns for years.

Cochrane [2011] famously described the proliferation of spurious factor claims as "the factor zoo." Until now, authors have explained this proliferation as the result of brute-force *p*-hacking via multiple testing (e.g., Harvey et al. [2016]). While this remains a very plausible explanation, the adverse outcomes mechanism indicates that this *p*-hacking may not be a brute-force exercise. The econometric canon enables "few-shot *p*-hacking", by favoring over-controlled underperforming (including money-losing) models with higher R-squared and lower *p*-values than the correctly specified money-making model. Factor researchers complying with the current econometric canon produce *p*-hacked factor strategies after a few trials, without ill-intent, and despite their best efforts to avoid selection bias or to control for multiple testing.

The causal mechanism for adverse outcomes is enforced by canon-complying factor researchers (a self-inflicted wound), not by the informed investors taking the other side of the trades. The solution is to correct the econometric canon such that the causal mechanism is disrupted. Figure 7 suggests two corrections. First, utilize the tools of causal inference. The correct way to justify a factor model's specification is to discover the causal graph associated with the data-generating process, and to apply do-calculus to block all non-causal paths, thus avoiding both, under-controlling for a confounder and over-controlling for a collider. Second, prevent that over-controlled models crowd out correctly specified factor models. Academics and practitioners should desist in their practice of choosing the factor model's specification with the highest explanatory power, as this practice all but guarantees over-controlling, with potentially dire consequences for investors. Investors can also take matters into their own hands, by defunding investment strategies that do not have a solid causal foundation (see López de Prado [2023]).

7. CONCLUSIONS

A common belief among researchers and investors is that specification errors are not dangerous, because an investment strategy based on a misspecified factor model will still perform well as long as the risk premia remain constant and are estimated with the correct sign. In this paper, we have proved that this belief is false, for two reasons. First, even if all correlations remain constant, misspecified factor and forecasting strategies: (a) underperform the correctly specified strategy in the confounder case; and (b) can yield systematic losses in the collider case. This is true regardless of whether the risk premia were estimated with the correct sign. Second, the performance of factor strategies that under-control for a confounder is not robust to parameter shift, and these strategies can yield systematic losses, even if the risk premia remain constant and are estimated with the correct sign.

Robustness is a key virtue of (correctly specified) causal models that associational (non-causal) models lack. Causal models are robust to parameter shift, in the sense that changes to parameters outside the causal path have no effect on the performance of those models. This gives a competitive advantage to causal models over associational models in finance. These are not arguments against machine learning or econometrics models in finance, but in favor of a better (causal) machine learning and econometrics models, that combine interdisciplinary research from statistics, economic theory, and experiment design.

The implication is that factor researchers should cease to dismiss as inconsequential concerns regarding specification errors. Authors should justify very carefully their specification choices, discuss potential confounders, and argue why the chosen controls cannot be colliders. Put bluntly, without some knowledge of the causal graph, it is possible (even likely) that an investment strategy will be fatally flawed.

Causal discovery and causal inference methods are being applied successfully to untangle the causal relationships in complex systems across various scientific fields (for example, see Runge et al. [2019], Shen at al. [2011]. In contrast, virtually no factor investing studies motivate their specification choices in terms of causal graphs derived from data; or develop refutable economic theories consistent with discovered causal graphs; or provide an analysis of potential colliders, confounders or mediators; or apply do-calculus to derive a sufficient adjustment set. The consequence is that factor researchers routinely make avoidable specification errors. A misspecified model that over-controls for a collider exhibits a higher explanatory power than a correctly specified model. Given the standard econometric practice of choosing a model's specification based on its explanatory power, it is natural for over-controlled models to crowd-out correctly specified models, making it more likely for underperforming and money-losing models to be selected.

Our findings challenge the scientific soundness and long-term profitability of the current multi-trillion-dollar (associational, casual, non-causal) factor investing industry. To overcome these pitfalls, academics and practitioners should rebuild the financial economics literature on the more scientifically rigorous grounds of causal factor investing.

APPENDIX

A.1. PROOF OF PROPOSITIONS IN SECTION 3.1

Consider a data-generating process with equations:

$$\begin{aligned} X &:= Z\delta + v \\ Y &:= X\beta + Z\gamma + u \end{aligned} \tag{28}$$

where $\gamma \neq 0$, $\delta \neq 0$, and variables (u, v, Z) are independent and identically distributed as a Gaussian. The process is unknown to observers, who attempt to estimate the causal effect of X on Y by fitting the equation $Y = X\beta + \varepsilon$ on a sample produced by the process. Then, the expression $E[\hat{\beta}|X]$ is,

$$E[\hat{\beta}|X] = (X'X)^{-1}X'E[Y|X] \tag{29}$$

First, note that the equation for Y can be rewritten as

$$Y := (Z\delta + v)\beta + Z\gamma + u = Z(\delta\beta + \gamma) + v\beta + u \tag{30}$$

The model can be written

$$\begin{bmatrix} Z \\ v \\ u \\ X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \delta & 1 & 0 \\ \gamma + \beta\delta & \beta & 1 \end{bmatrix} \begin{bmatrix} Z \\ v \\ u \end{bmatrix} = A \begin{bmatrix} Z \\ v \\ u \end{bmatrix} \tag{31}$$

where $(Z, v, u)' \sim N(0, I)$. The variance is then

$$\mathrm{Var}\begin{bmatrix} Z \\ v \\ u \\ X \\ Y \end{bmatrix} = AA' = \begin{bmatrix} 1 & 0 & 0 & \delta & \gamma + \beta\delta \\ 0 & 1 & 0 & 1 & \beta \\ 0 & 0 & 1 & 0 & 1 \\ \delta & 1 & 0 & \delta^2 + 1 & \beta(1 + \delta^2) + \delta\gamma \\ \gamma + \beta\delta & \beta & 1 & \beta(1 + \delta^2) + \gamma\delta & 1 + \beta^2 + (\gamma + \beta\delta)^2 \end{bmatrix} \tag{32}$$

We can read the covariances between X, Y, Z, u, v from this matrix. We can use it to compute conditional expectations (Eaton [1983, pp. 116-117]), for instance:

$$\begin{aligned} E[v|X] &= E[v] + \Sigma_{vX}\Sigma_{XX}^{-1}(X - E[X]) = \frac{1}{1 + \delta^2}X \\ E[Z|X] &= E[Z] + \Sigma_{ZX}\Sigma_{XX}^{-1}(X - E[X]) = \frac{\delta}{1 + \delta^2}X \\ E[Y|Z] &= (\gamma + \beta\delta)Z \\ E[Y|X, Z] &= X\beta + Z\gamma \end{aligned} \tag{33}$$

In particular,

$$E[Y|X] = E[Y] + \Sigma_{YX}\Sigma_{XX}^{-1}(X - E[X]) = \Sigma_{YX}\Sigma_{XX}^{-1}X = \frac{\delta(\delta\beta + \gamma)\sigma_Z^2 + \beta\sigma_v^2}{\delta^2\sigma_Z^2 + \sigma_v^2}X \tag{34}$$
$$= \left(\beta + \frac{\delta\gamma\sigma_Z^2}{\delta^2\sigma_Z^2 + \sigma_v^2}\right)X$$

$$E[\hat{\beta}|X] = \beta + \frac{\delta\gamma\sigma_Z^2}{\delta^2\sigma_Z^2 + \sigma_v^2} \tag{35}$$

This expression simplifies further when the variables are independent and identically distributed as a standard Normal, $(u, v, Z) \sim N[0, I]$, since

$$E[\hat{\beta}|X] = \beta + \frac{\gamma\delta}{1 + \delta^2} \tag{36}$$

When $\gamma = \delta = 0$, there is no confounder, thus $E[\hat{\beta}|X] = \beta$.

This completes the proof.

A.2. PROOF OF PROPOSITIONS IN SECTION 3.2

Consider a data-generating process with equations:

$$\begin{aligned} Y &:= X\beta + u \\ Z &:= Y\gamma + X\delta + v \end{aligned} \tag{37}$$

where $\gamma \neq 0$, $\delta \neq 0$, and variables (u, v, X) are independent and identically distributed as a Gaussian. The process is unknown to observers, who attempt to estimate the causal effect of X on Y by fitting the equation $Y = X\beta + Z\theta + \varepsilon$ on a sample produced by the process.

First, note that the equation for Z can be rewritten as

$$Z := (X\beta + u)\gamma + X\delta + v = X(\beta\gamma + \delta) + u\gamma + v \tag{38}$$

This can be written

$$\begin{bmatrix} X \\ u \\ v \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \beta & 1 & 0 \\ \delta + \beta\gamma & \gamma & 1 \end{bmatrix} \begin{bmatrix} X \\ u \\ v \end{bmatrix} = A \begin{bmatrix} X \\ u \\ v \end{bmatrix} \tag{39}$$

where $(X, u, v) \sim N(0, I)$. The variance is then

$$\operatorname{Var}\begin{bmatrix} X \\ u \\ v \\ Y \\ Z \end{bmatrix} = AA' = \begin{bmatrix} 1 & 0 & 0 & \beta & \beta\gamma+\delta \\ 0 & 1 & 0 & 1 & \gamma \\ 0 & 0 & 1 & 0 & 1 \\ \beta & 1 & 0 & 1+\beta^2 & \beta(\beta\gamma+\delta)+\gamma \\ \beta\gamma+\delta & \gamma & 1 & \beta(\beta\gamma+\delta)+\gamma & 1+\gamma^2+(\beta\gamma+\delta)^2 \end{bmatrix} \tag{40}$$

From this matrix, we can read the covariances between X, Y, Z, u, v. We can also use it to compute conditional expectations (Eaton [1983, pp. 116-117]), for instance,

$$\begin{aligned} E[Z|X] &= E[Z] + \Sigma_{ZX}\Sigma_{XX}^{-1}(X - E[X]) \\ &= (\beta\gamma+\delta)X \\ E[u|X,Z] &= E[u] + \Sigma_{u,XZ}\Sigma_{XZ,XZ}^{-2}\begin{pmatrix} X - E[X] \\ Z - E[Z] \end{pmatrix} \\ &= (0 \quad \gamma)\begin{pmatrix} 1 & \beta\gamma+\delta \\ \beta\gamma+\delta & 1+\gamma^2+(\beta\gamma+\delta)^2 \end{pmatrix}^{-1}\begin{pmatrix} X \\ Z \end{pmatrix} \\ &= -\frac{\gamma(\beta\gamma+\delta)}{1+\gamma^2}X + \frac{\gamma}{1+\gamma^2}Z \end{aligned} \tag{41}$$

In particular,

$$\begin{aligned} E[Y|X,Z] = E[Y] + \Sigma_{Y,XZ}\Sigma_{XZ,XZ}^{-1}\left(\begin{bmatrix} X \\ Z \end{bmatrix} - E\begin{bmatrix} X \\ Z \end{bmatrix}\right) \\ = [\Sigma_{YX} \quad \Sigma_{YZ}]\begin{bmatrix} \Sigma_{XX} & \Sigma_{XZ} \\ \Sigma_{ZX} & \Sigma_{ZZ} \end{bmatrix}^{-1}\begin{bmatrix} X \\ Z \end{bmatrix} \\ = \frac{1}{\Sigma_{XX}\Sigma_{ZZ} - \Sigma_{XZ}^2}[\Sigma_{YX} \quad \Sigma_{YZ}]\begin{bmatrix} \Sigma_{ZZ} & -\Sigma_{XZ} \\ -\Sigma_{ZX} & \Sigma_{XX} \end{bmatrix}\begin{bmatrix} X \\ Z \end{bmatrix} \\ = \frac{1}{\Sigma_{XX}\Sigma_{ZZ} - \Sigma_{XZ}^2}[\Sigma_{YX}\Sigma_{ZZ} - \Sigma_{YZ}\Sigma_{ZX} \quad \Sigma_{YZ}\Sigma_{XX} - \Sigma_{YX}\Sigma_{XZ}]\begin{bmatrix} X \\ Z \end{bmatrix} \end{aligned} \tag{42}$$

From the above, we obtain that

$$E[\hat{\beta}|X,Z] = \frac{\Sigma_{YX}\Sigma_{ZZ} - \Sigma_{YZ}\Sigma_{ZX}}{\Sigma_{XX}\Sigma_{ZZ} - \Sigma_{XZ}^2} = \frac{\beta\sigma_v^2 - \delta\gamma\sigma_u^2}{\sigma_v^2 + \gamma^2\sigma_u^2} \tag{43}$$

$$E[\hat{\theta}|X,Z] = \frac{\Sigma_{YZ}\Sigma_{XX} - \Sigma_{YX}\Sigma_{XZ}}{\Sigma_{XX}\Sigma_{ZZ} - \Sigma_{XZ}^2} = \frac{\gamma\sigma_u^2}{\sigma_v^2 + \gamma^2\sigma_u^2} \tag{44}$$

This expression simplifies further when the variables are independent and identically distributed as a standard Normal, $(u, v, X) \sim N[0, I]$, since

$$E[\hat{\beta}|X,Z] = \frac{\beta - \delta\gamma}{1+\gamma^2} \tag{45}$$

$$E\left[\hat{\theta}|X,Z\right] = \frac{\gamma}{1+\gamma^2} \tag{46}$$

When $\gamma = \delta = 0$, there is no collider, thus $E\left[\hat{\beta}|X,Z\right] = \beta$ and $E\left[\hat{\theta}|X,Z\right] = 0$.

This completes the proof.

A.3. PROOFS OF PROPOSITIONS IN SECTION 4

Consider the following collider model, where $\gamma \neq 0$, $\delta \neq 0$, and $(X, u, v) \sim N[0, I]$:

$$\begin{aligned} Y &\coloneqq X\beta + u \\ Z &\coloneqq Y\gamma + X\delta + v \end{aligned} \tag{47}$$

We can rewrite Z as

$$Z := (X\beta + u)\gamma + X\delta + v = X(\beta\gamma + \delta) + u\gamma + v \tag{48}$$

The model that over-controls for a collider is

$$\begin{aligned} Y \quad &= XE\left[\hat{\beta}|X,Z\right] + ZE\left[\hat{\theta}|X,Z\right] + \varepsilon \\ &= \frac{\beta - \delta\gamma}{1+\gamma^2}X + \frac{\gamma}{1+\gamma^2}Z + \varepsilon \end{aligned} \tag{49}$$

The residuals of the over-controlled model are

$$\begin{aligned} \varepsilon \quad &= Y - XE\left[\hat{\beta}|X,Z\right] - ZE\left[\hat{\theta}|X,Z\right] \\ &= (X\beta + u) - XE\left[\hat{\beta}|X,Z\right] - [X(\beta\gamma + \delta) + u\gamma + v]E\left[\hat{\theta}|X,Z\right] \\ &= X\left[\beta - E\left[\hat{\beta}|X,Z\right] - E\left[\hat{\theta}|X,Z\right](\beta\gamma + \delta)\right] + u\left(1 - E\left[\hat{\theta}|X,Z\right]\gamma\right) - E\left[\hat{\theta}|X,Z\right]v \\ &= X\left[\beta - \frac{\beta - \delta\gamma}{1+\gamma^2} - \frac{\gamma}{1+\gamma^2}(\beta\gamma + \delta)\right] + u\left[1 - \frac{\gamma}{1+\gamma^2}\gamma\right] - \frac{\gamma}{1+\gamma^2}v \\ &= X\frac{\beta(1+\gamma^2) - (\beta - \delta\gamma) - \gamma(\beta\gamma + \delta)}{1+\gamma^2} + u\frac{1}{1+\gamma^2} - \frac{\gamma}{1+\gamma^2}v \\ &= \frac{1}{1+\gamma^2}u - \frac{\gamma}{1+\gamma^2}v \end{aligned} \tag{50}$$

The variances of ε and Y in the over-controlled model are

$$Var[\varepsilon] = \left(\frac{1}{1+\gamma^2}\right)^2 + \left(\frac{\gamma}{1+\gamma^2}\right)^2 = \frac{1+\gamma^2}{(1+\gamma^2)^2} = \frac{1}{1+\gamma^2} \tag{51}$$

$$Var[Y] = \beta^2 + 1 \tag{52}$$

A.3.1. R-SQUARED

The R-squared (R^2) of the correctly specified model (if the estimated coefficients have their expected values) is

$$\begin{aligned} R^2 &= 1 - \frac{Var[u]}{Var[Y]} \\ &= 1 - \frac{1}{1+\beta^2} \end{aligned} \tag{53}$$

For the over-controlled model, the R^2 is

$$\begin{aligned} R^2 &= 1 - \frac{Var[\varepsilon]}{Var[Y]} \\ &= 1 - \frac{1}{(1+\gamma^2)(1+\beta^2)} \end{aligned} \tag{54}$$

Therefore, for R^2 we always have:

$$R^2_{\text{over-controlled}} > R^2_{\text{correct}} \tag{55}$$

This completes the proof.

A.3.2. ADJUSTED R-SQUARED

Since the number of estimated parameters in the correct model is $k = 1$, the adjusted R^2 for T observations is

$$\begin{aligned} R^2_{\text{adj}} &= 1 - \frac{Var[u]}{Var[Y]} \frac{T-1}{T-k-1} \\ &= 1 - \frac{1}{1+\beta^2} \frac{T-1}{T-2} \end{aligned} \tag{56}$$

Since we now have $k = 2$ estimated parameters, the adjusted R^2 of the over-controlled model is

$$\begin{aligned} R^2_{\text{adj}} &= 1 - \frac{Var[\varepsilon]}{Var[Y]} \frac{T-1}{T-k-1} \\ &= 1 - \frac{1}{(1+\gamma^2)(1+\beta^2)} \frac{T-1}{T-3} \end{aligned} \tag{57}$$

Therefore, if $|\gamma|$ is sufficiently large relative to the number of observations, for the adjusted R^2 we have:

$$
\begin{aligned}
R^2_{\text{adj, over-controlled}} > R^2_{\text{adj, correct}} \quad &\Leftrightarrow \quad \frac{1}{(1+\gamma^2)(1+\beta^2)}\frac{T-1}{T-3} < \frac{1}{1+\beta^2}\frac{T-1}{T-2} \\
&\Leftrightarrow \quad \frac{1}{1+\gamma^2} < \frac{T-3}{T-2} \\
&\Leftrightarrow \quad 1+\gamma^2 > \frac{T-2}{T-3} \\
&\Leftrightarrow \quad |\gamma| > \frac{1}{\sqrt{T-3}}
\end{aligned} \tag{58}
$$

This completes the proof.

A.3.3. P-VALUES

For a linear regression model $Y = X\beta + Z\theta + \varepsilon$, fitted on T observations, the covariance matrix of the coefficients is

$$
\begin{aligned}
\mathrm{Var}\begin{bmatrix} E[\hat{\beta}|X,Z] \\ E[\hat{\theta}|X,Z] \end{bmatrix} &= \mathrm{Var}[\varepsilon]\begin{bmatrix} T\ \mathrm{Var}[X] & T\ \mathrm{Cov}[X,Z] \\ T\ \mathrm{Cov}[X,Z] & T\ \mathrm{Var}[Z] \end{bmatrix}^{-1} \\
&= \frac{1}{T}\frac{\mathrm{Var}[\varepsilon]}{\mathrm{Var}[X]\ \mathrm{Var}[Z] - \mathrm{Cov}[X,Z]^2}\begin{bmatrix} \mathrm{Var}[Z] & -\mathrm{Cov}[X,Z] \\ -\mathrm{Cov}[X,Z] & \mathrm{Var}[X] \end{bmatrix}
\end{aligned} \tag{59}
$$

For individual tests on the regression coefficients, we only need the diagonal elements. In the case where the model over-controls for a collider, the variances of the estimated coefficients are:

$$
\begin{aligned}
\mathrm{Var}\left[E[\hat{\beta}|X,Z]\right] &= \frac{\mathrm{Var}[\varepsilon]}{T}\frac{\mathrm{Var}[Z]}{\mathrm{Var}[X]\ \mathrm{Var}[Z] - \mathrm{Cov}[X,Z]^2} \\
&= \frac{1}{T}\frac{1}{1+\gamma^2}\frac{(\beta\gamma+\delta)^2+\gamma^2+1}{(\beta\gamma+\delta)^2+\gamma^2+1-(\beta\gamma+\delta)^2} \\
&= \frac{(\beta\gamma+\delta)^2+\gamma^2+1}{T(1+\gamma^2)^2}
\end{aligned} \tag{60}
$$

$$
\begin{aligned}
\mathrm{Var}\left[E[\hat{\theta}|X,Z]\right] &= \frac{\mathrm{Var}[\varepsilon]}{T}\frac{\mathrm{Var}[X]}{\mathrm{Var}[X]\ \mathrm{Var}[Z] - \mathrm{Cov}[X,Z]^2} = \frac{1}{T}\frac{1}{1+\gamma^2}\frac{1}{1+\gamma^2} \\
&= \frac{1}{T(1+\gamma^2)^2}
\end{aligned} \tag{61}
$$

Under the null hypothesis, the t-values of the estimated coefficients in the over-controlled model follow a t-distribution with $T-2$ degrees of freedom,

$$t_{E[\hat{\beta}|X,Z]} = \frac{E\left[\hat{\beta}|X,Z\right]}{\sqrt{\mathrm{Var}\left[E\left[\hat{\beta}|X,Z\right]\right]}} = \frac{\frac{\beta-\delta\gamma}{1+\gamma^2}}{\sqrt{\frac{(\beta\gamma+\delta)^2+\gamma^2+1}{T(1+\gamma^2)^2}}} \tag{62}$$
$$= \sqrt{T}\frac{\beta-\delta\gamma}{\sqrt{(\beta\gamma+\delta)^2+\gamma^2+1}}$$

$$t_{E[\hat{\theta}|X,Z]} = \frac{E\left[\hat{\theta}|X,Z\right]}{\sqrt{\mathrm{Var}\left[E\left[\hat{\theta}|X,Z\right]\right]}} = \frac{\frac{\gamma}{1+\gamma^2}}{\frac{1}{\sqrt{T(1+\gamma^2)^2}}} = \sqrt{T}\gamma \tag{63}$$

For the correctly specified model, the computations are simpler. The variance of the coefficient is

$$\mathrm{Var}\left[E\left[\hat{\beta}|X,Z\right]\right] = \mathrm{Var}[u](T\,V\mathrm{ar}[X])^{-1} = \frac{1}{T} \tag{64}$$

Under the null hypothesis, the t-value of the correctly specified model follows a t-distribution with $T-1$ degrees of freedom,

$$t_{E[\hat{\beta}|X,Z]} = \frac{E\left[\hat{\beta}|X,Z\right]}{\sqrt{\mathrm{Var}\left[E\left[\hat{\beta}|X,Z\right]\right]}} = \frac{E\left[\hat{\beta}|X,Z\right]}{\sqrt{1/T}} = \sqrt{T}\beta \tag{65}$$

In conclusion, false positives become more likely in models that over-control for a collider, because: (a) $E\left[\hat{\theta}|X,Z\right]$ is statistically significant for $T \gg 0$; and (b) for some values of δ, γ (in particular, $|\delta| \gg 0$, $|\gamma| \gg 0$, $|\delta| \approx |\gamma|$), or when the causal effect is weak ($|\beta| \approx 0$), then

$$|t_{E[\hat{\beta}|X,Z],\text{over-controlled}}| > |t_{E[\hat{\beta}|X,Z],\text{correct}}| \tag{66}$$

This completes the proof.

REFERENCES

Akey, P., A. Robertson, and M. Simutin (2023): "Noisy Factors." Working paper. Available at https://ssrn.com/abstract_id=3930228

Anderson, R. (2011): "Time-Varying Risk Premia." *Journal of Mathematical Economics*, Vol. 47, No. 3, pp. 253–259.

Bailey, D., J. Borwein, M. López de Prado, and J. Zhu (2014): "Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-Of-Sample Performance." *Notices of the American Mathematical Society*, Vol. 61, No. 5, pp. 458-471

Belkin, M., D. Hsu, S. Ma, and S. Mandal (2019): "Reconciling modern machine learning practice and the bias-variance trade-off." Working paper, available at https://arxiv.org/abs/1812.11118

Berkson, J. (1946): "Limitations of the Application of Fourfold Table Analysis to Hospital Data." *Biometrics Bulletin*, Vol. 2, No. 3, pp. 47–53.

BlackRock (2017): "What Is Factor Investing?" Online publication. Available at www.blackrock.com/us/individual/investment-ideas/what-is-factor-investing

Bloomberg (2021): "Bloomberg GSAM US Equity Multi Factor Index." *Bloomberg Professional Services – Indices*. Available through the Bloomberg Terminal. https://assets.bbhub.io/professional/sites/10/Bloomberg-GSAMUS-Equity-Multi-Factor-Index-Fact-Sheet.pdf

Chen, B. and J. Pearl (2013): "Regression and Causation: A Critical Examination of Six Econometrics Textbooks." *Real-World Economics Review*, No. 65, pp. 2–20. http://www.paecon.net/PAEReview/issue65/ChenPearl65.pdf.

Cochrane, J. (2011): "Presidential Address: Discount Rates." *Journal of Finance*, Vol. 66, No. 4, pp. 1047–1108.

Eaton, M. (1983): *Multivariate Statistics: A Vector Space Approach*. Wiley, 1st ed.

Evans, M. (1994): "Expected Returns, Time-Varying Risk, and Risk Premia." *Journal of Finance*, Vol. 49, No. 2, pp. 655–679.

Fama, E. and K. French (1993): "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, Vol. 33, No. 1, pp. 3-56.

Fama, E. and K. French (2015): "A five-factor asset pricing model." *Journal of Financial Economics*, Vol. 116, No. 1, pp. 1-22.

Fama, E. and J. MacBeth (1973): "Risk Return and Equilibrium: Empirical Tests." *Journal of Political Economy*, Vol. 81, No. 3, pp. 607–636.

Ferson, W. (2019): *Empirical Asset Pricing: Models and Methods*. MIT Press, 1st ed.

Giglio, S. and D. Xiu (2021): "Asset Pricing with Omitted Factors." *Journal of Political Economy*, Nol. 129, No. 7, pp. 1947-1990.

Glymour, C., K. Zhang, and P. Spirtes (2019): "Review of Causal Discovery Methods Based on Graphical Models." *Frontiers in Genetics*, Vol. 10, No. 524, pp. 1–15, www.frontiersin.org/articles/10.3389/fgene.2019.00524/full.

Gospodinov, N., R. Kan, and C. Robotti (2014): Misspecification-Robust Inference in Linear Asset-Pricing Models with Irrelevant Risk Factors. *Review of Financial Studies*, Vol. 27, No. 7, pp. 2139-2170.

Greene, W. (2012): *Econometric Analysis*. Pearson Education, 7th ed

Gu, L., H. Zhang, A. Heinz, J. Liu, T. Yao, M. AlRemeithi, Z. Luo, D. Ruppert (2023): "Re-examination of Fama-French factor investing with causal inference methods." Working paper. Available at https://ssrn.com/abstract=4677537

Harvey, C., Y. Liu, and H. Zhu (2016): "... and the Cross-Section of Expected Returns." *Review of Financial Studies*, Vol. 29, No. 1, pp. 5–68.

Harvey, C. and Y. Liu (2020): "False (and Missed) Discoveries in Financial Economics." *Journal of Finance*, Vol. 75, No. 5, pp. 2503-2553.

Jensen, M., F. Black, and M. Scholes (1972): "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theory of Capital Markets*, edited by M. Jensen. Westport, CT: Praeger.

Leamer, E. (1983): "Let's Take the Con Out of Econometrics." *American Economic Review*, Vol. 73, No. 1, pp. 31–43.

López de Prado, M. (2020): *Machine Learning for Asset Managers.* Cambridge University Press. https://doi.org/10.1017/9781108883658

López de Prado, M. (2023): *Causal Factor Investing: Can Factor Investing Become Scientific?* Cambridge University Press. https://doi.org/10.1017/9781009397315

López de Prado, M. and D. Bailey (2021): "The False Strategy Theorem: A Financial Application of Experimental Mathematics." *American Mathematical Monthly*, Vol. 128, No. 9, pp. 825-831.

López de Prado, M. and D. Leinweber (2012): "Advances in cointegration and subset correlation hedging methods." *Journal of Investment Strategies*, Vol. 1, No. 2, pp. 67-115.

Neal, B. (2020): Introduction to Causal Inference: From a Machine Learning Perspective. Course Lecture Notes (December 17, 2020). www.bradyneal.com/causal-inference-course.

Neuberger Berman (2019): “Inside the Quant Investing Trend.” *Quarterly Views*. www.nb.com/documents/public/en-us/Messinger_Client_Letter_2Q19.pdf

Pearl, J. (2009): *Causality: Models, Reasoning and Inference*. Cambridge, 2nd ed.

Pearl, J. (2014): “Understanding Simpson's Paradox.” *The American Statistician*. Vol. 68, No. 1, pp. 8–13.

Pearl, J., M. Glymour, and N. Jewell (2016): *Causal Inference in Statistics: A Primer*. Wiley, 1st ed.

Runge, J., P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic (2019): “Detecting and quantifying causal associations in large nonlinear time series datasets.” *Science Advances*, Vol. 5, No. 11. Available at https://www.science.org/doi/10.1126/sciadv.aau4996

Sadeghi, A., A. Gopal, and M. Fesanghary (2024): “Causal discovery in financial markets: A framework for nonstationary time-series data.” Working paper. Available at https://arxiv.org/abs/2312.17375

Schipper, K. and R. Thompson (1981): “Common Stocks as Hedges against Shifts in the Consumption or Investment Opportunity Set.” *Journal of Business*, Vol. 1, pp. 305–328.

Shen, X., S. Ma, P. Vemuri, M. Castro, P. Caraballo, G. Simon (2021): “A novel method for causal structure discovery from EHR data and its application to type-2 diabetes mellitus.” *Scientific Reports*, Vol. 11: 21025. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8546093/

Stigler, S. (1981): “Gauss and the Invention of Least Squares.” *The Annals of Statistics*, Vol. 9, No. 3, pp. 465-474