Corrigendum to "Characteristics are Covariances: A Unified Model of Risk and Return"

Bryan T. Kelly, Seth Pruitt, and Yinan Su

September 2024

We list label sections based on the date the typo or error was discussed in this document, which has been publicly available since June 2020. Sections are ordered in reverse chronological order.

1 September 2024—static versus dynamic loadings

1.1 Changes to the paper

In the published version of "Characteristics are Covariances: A Unified Model of Risk and Return" (*Journal of Financial Economics*, December 2019, 134(3): 237–254) in section 4.10 we discuss the decomposition of characteristics into their historical mean and deviations from that mean, which we denote $\bar{c}_{i,t}$ and $(c_{i,t} - \bar{c}_{i,t})$, respectively (see page 521). In the code `IPCA_empirical_datamaker.m` we made an error in how historical means were calculated, at line 120. We thank Zehong Li for bringing this, and the following related typo, to our attention.[1] An updated version of the code now resides in the Dropbox folder whose link is available at https://sethpruitt.net/downloads/, a folder where all our code has been publicly available since before the paper's publication. This error affects only KPS Table 14 for columns other than those labeled "c". (we use "KPS" to refer to tables in the publication, as opposed to this corrigendum)

Table 1 provides the corrected results.[2] The columns for c in KPS Table 14 are unchanged, and so we do not report them in Table 1. The corrected table largely echoes what we saw in

[1] Furthermore, to avoid collinearity in the first month of data, we changed line 141 to replace the (rank-demeaned) characteristics with deviations from the historical mean only in months 2 and onwards.

[2] Also there is a typo in Table 14's notes: it should read "different coefficients".

Table 1: Corrected KPS Table 14, static versus dynamic loadings

This table reports percentage R^2 from IPCA specifications based on average characteristic levels ("$\bar{c}$"), time series deviations from average levels ("$c-\bar{c}$"), or our baseline specification in which levels and deviations are allowed to enter with different coefficients. It corrects the mistaken numbers reported in KPS Table 14.

	Total R^2			Predictive R^2		
K	$\bar{c}$	$c-\bar{c}$	$\bar{c}$, $c-\bar{c}$	$\bar{c}$	$c-\bar{c}$	$\bar{c}$, $c-\bar{c}$
1	14.1	13.2	14.9	0.29	0.39	0.36
2	15.7	14.4	16.5	0.29	0.40	0.35
3	16.5	15.1	17.6	0.32	0.70	0.44
4	16.8	15.6	18.2	0.33	0.70	0.44
5	17.2	15.9	18.8	0.33	0.69	0.74
6	17.4	16.2	19.1	0.35	0.70	0.73

KPS. There is little gained from breaking apart characteristics into $(\bar{c}, c-\bar{c})$ rather than the characteristic itself c. In the corrected results, the deviations continue to be most important for conditional expectations, as they provide at least twice the predictive R^2 as historical means. This continues to support characteristic dynamics are important for understanding firm-level factor exposures.

A related typo occurs because our data creation code was designed to use the exact same firm-month observations for our main analysis as for the analysis of section 4.10. But in section 4.10 we consider a dataset with both characteristic means and deviations as separate firm-level characteristics, resulting in $L = 73$ as page 521 says. For this reason, we put in a firm-count filter on our data creation: only months with 100 (a round number greater than 73) firm observations are retained. We first noted this data filter in the June 2020 version of this corrigendum, where we verified that our main results are unchanged by reduced the filter number.[3]

The filter results in our using $T = 599$ monthly observations, just as page 513 states. This means that the months in our sample are July 1964 through May 2014: the first two years of the raw data are not used because the cross section of firms is small. Therefore we have an error in what is said (not the results reported) on page 515. We do use every month $t \geq 120$, but this means the parenthetical remark should state "July 1974" instead of "July 1972".

[3]When we do this, we should note that the factor-estimation regression can become undefined for $Z_t'Z_t$ with reduced rank, since then Γ_β could pick off the null space of the matrix. It is a matter of data configuration (i.e. it is random) whether this is the case.

2 May 2024–Typo

There is a typo in the third unnumbered mathline on page 511, the one above "4. Empirical findings". It should read

$$\tilde{x}_t^b = Z_t' Z_t \tilde{\Gamma}_\beta \hat{f}_t + \tilde{d}_t^b.$$

That is, it requires the premultiplication by Z_t' analogous to equation 13 on page 509.

We thank Valentin Luz for alerting us to this.

3 December 2023—arbitrage portfolios

3.1 Changes to the paper

In the paper we report mistaken empirical results for the "arbitrage" portfolios, in KPS Table 7. Our code erroneously over lagged—that is, the weight w_{t-1} used Z_{t-2} (in the code) instead of the proper Z_{t-1} (as correctly written on page 517 in Section 4.5.3). For bringing this to our attention in the code `IPCA_empirical_GBGA_outofsample.m` we thank the research team of Luca Pezzo, Raja Velu, Lei Wang, and Zhaoque (Chosen) Zhou. An updated version of the code now resides in the Dropbox folder whose link is available at https://sethpruitt.net/downloads/, a folder where all our code has been publicly available since before the paper's publication. To reiterate, only KPS Table 7 changes and the error had no effect on any other results.

Table 2: Corrected KPS Table 7, IPCA pure-alpha portfolios

This table reports the out-of-sample annualized Sharpe ratios for a portfolio designed to exploit characteristic-based mispricing estimated from Γ_α in the unrestricted IPCA model. It corrects the mistaken numbers reported in KPS Table 7.

	K					
	1	2	3	4	5	6
Sharpe ratio	1.42	1.82	2.40	3.00	2.91	2.71

The corrected results are in Table 2. Notably these annualized Sharpe ratios are above one and reach as high as 3. Therefore two statements we make in the text should be altered—we discuss them in the order of the page they appear.

On page 504 it is stated:

> However, we find that pure-alpha compensation is far smaller than that earned from harvesting factor risk premia.

This is no longer true because the pure-alpha portfolio Sharpe ratio is not "far smaller" than the tangency portfolio Sharpe ratio reported in KPS Table 6, when $K \geq 3$. Indeed, we find the Sharpe ratios of comparable size.

On page 517 we discuss KPS Table 7 and state:

> Pure-alpha Sharpe ratios range from 0.55 to 1.07, and are substantially smaller than the factor risk premium portfolios reported in Table 6... And, with these Sharpe ratios, they appear far from consistent with the typical "near arbitrage" interpretation of mispricing-based alpha.

Obviously now the numerical range is 1.42 to 3.00 and so these are more consistent to the typical "near arbitrage" interpretation of mispricing-based alpha.

3.2 Support for the qualitative argument in KPS

Our qualitative conclusion in KPS was that an investor trading factor risk does not gain a significant amount by trading pure-alpha. After correcting the above mistakes, we continue to argue this is the case.

To demonstrate this, we calculate two notions of a tangency portfolio. The first is the out-of-sample tangency portfolio attainable for an investor using only the restricted K-factor IPCA model with alphas fixed at zero (these are exactly the results in the second row of Panel A of KPS Table 6). The second uses the K factors from the restricted model, but also gives the investor access to the pure-alpha portfolio implied by the unrestricted model. As Table 3 shows, inclusion of the pure-alpha portfolio has no material benefit for out-of-sample performance for $K > 3$.

These findings support the statement (hidden in the ellipsis above) on KPS page 517, albeit using a novel notion of the tangency portfolio:

> While pure-alpha portfolios have the virtue of factor neutrality, they do not appear particularly attractive in terms of mean-variance efficiency.

Out-of-sample performance degrades when we add the pure-alpha portfolio to the factor space (of a sufficiently rich dimension), implying it adds nothing useful to investors' mean-variance objective.

Table 3: Tangency portfolios, restricted factors and pure-alpha portfolio

This table reports the out-of-sample annualized Sharpe ratios for two notions of a tangency portfolio. The first is the out-of-sample tangency portfolio attainable for an investor using only the restricted K-factor IPCA model with alphas fixed at zero. The second uses the K factors from the restricted model, but also gives the investor access to the pure-alpha portfolio implied by the unrestricted model.

	K					
	1	2	3	4	5	6
Factors	0.62	0.62	2.51	3.09	3.89	4.05
Factors + pure-alpha	1.57	1.90	2.72	3.13	3.10	3.59

4 June 2020—Typos

Table 3, Panel C, row SMB, column 1 has a typo: it should read "29.7" instead of "2.97"

Footnote 26 should read "at least 12 non-missing months" instead of 60. If we instead require 60 months, the results are identical.

Given the root-N consistency of the factor estimates, we filtered the data to look only at months with at least 100 cross-sectional observations. If we reduce this filter to 6 observations (so that $Z_t'Z_t$ has rank ≥ 6 every month) we pick up 12 more months of data. Our results are unchanged.

We thank Philip Bold for alerting us to these.