

Working Capital Financing and Firm Performance: A Machine Mislearning Approach - A comment on Mahmood, Ahmed, Hussain & Ben-Zaied (2025)

Gwenael Piaser

gwenael.piaser@ipag.fr

IPAG Business School

Research Article

Keywords: Working capital financing, Random forest, Principal component analysis, Model comparison, Research methodology, Replication

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10121836/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Working Capital Financing and Firm Performance: A Machine Mislearning Approach

A comment on Mahmood, Ahmed, Hussain & Ben-Zaied (2025)

June 23, 2026

Abstract. Mahmood et al. (2025) compare a fixed-effects panel model with two tree-based “machine learning” estimators (a regression tree and a random forest) and report, for each, a table of signed coefficients, standard errors, significance levels, an R^2 , and an F -statistic, concluding that the machine-learning models “explain the variance better.” This comment raises three concerns, all verifiable from the article’s own text and tables. (i) Principal component analysis, listed as a competing estimator, is an unsupervised dimension-reduction technique that estimates no relationship and serves only as upstream feature construction. (ii) The quantities reported in the tree and forest tables (signed coefficients, parametric standard errors, significance stars, an F -statistic) are not produced by tree-based methods; they are linear-regression output, as the manually added quadratic and interaction terms confirm. (iii) The estimates are mutually inconsistent: Table 13 combines fixed-effects and random-forest values within one column, and the full-sample random-forest coefficient on WCF takes three incompatible values (0.598, 0.403, 0.323). The in-sample R^2 comparison underlying the central claim is, moreover, uninformative about generalization.

Keywords: Working capital financing; Random forest; Principal component analysis; Model comparison; Research methodology; Replication.

JEL classification: C18; C38; C52; G32.

1. Introduction

Mahmood et al. (2025; hereafter MAHB) study the relationship between short-term working capital finance (WCF) and firm profitability for a balanced panel of Chinese manufacturing firms over 2000–2018 ($\approx 15,238$ firm-year observations). The hypothesized relationship is nonlinear and concave (an inverted U), moderated by ownership status, firm size, and leverage. Methodologically, the article proceeds in two layers. It first uses PCA to compress, for each construct, several literature proxies into a single “master proxy” (the first principal component). It then contrasts a fixed-effects (FE) panel model with a regression tree and a random forest, presented as “machine learning” competitors, and concludes that the tree-based models “provide a more accurate and reliable explanation” of the data because they attain higher R^2 and F values.

The economic question is legitimate and well grounded in the literature. The purpose of this comment is narrower: to assess the *internal consistency* among (i) the methods the article names, (ii) the statistical objects it actually reports, and (iii) the numerical values it prints. Every observation below is verifiable from MAHB’s own text, equations, and tables; no external data or re-estimation is required. Section 2 recalls what regression trees and random forests produce. Section 3 explains why PCA is not an estimator comparable to a regression. Section 4 shows that the reported coefficient tables cannot be the output of tree-based methods. Section 5 documents internal numerical inconsistencies. Section 6 argues that the central model comparison is invalid even granting the authors’ framing. Section 7 lists further inconsistencies, and Section 8 concludes.

2. What regression trees and random forests produce

A regression tree partitions the predictor space into disjoint regions $R_1, \dots, R_M$ and predicts a constant within each region; the fitted function is piecewise constant,

$$\hat{f}(x) = \sum_{m=1}^M c_m \mathbf{1}\{x \in R_m\}, \quad c_m = \text{mean}\{y_i : x_i \in R_m\},$$

where the regions are obtained by recursive binary splitting (Breiman et al. 1984; Hastie et al. 2009). A random forest averages B such trees, each grown on a bootstrap resample of the data and with a random subset of candidate predictors considered at every split, giving $\hat{f}_{\text{RF}}(x) = B^{-1} \sum_{b=1}^B \hat{f}_b(x)$ (Breiman 2001).

The decisive point for what follows is that neither estimator has the form $y = x'\beta + \varepsilon$. There is no parameter vector with one entry per regressor: a single covariate typically enters at many split points, with different thresholds, and its influence on the prediction is local and need not be monotone. Consequently, a fitted tree or forest does *not* yield

- a signed coefficient attached to each variable;
- a parametric standard error, and hence no t -statistic, p -value, or significance “star”;
- an F -statistic, which is defined through the analysis-of-variance decomposition of a *linear* model (Greene 2018; Wooldridge 2010).

What a forest does produce is a measure of *variable importance* (most commonly permutation importance or the mean decrease in node impurity), together with an out-of-bag estimate of prediction error (Breiman 2001). Importance ranks the predictive contribution of a variable; it is unsigned and is not a marginal effect, and its naive use is known to be misleading (Strobl et al. 2007). Notably, MAHB themselves cite a cautionary treatment of this point (Wallace et al. 2023). To recover the direction and shape of a predictor’s influence, one uses partial dependence or individual conditional expectation plots (Friedman 2001) or Shapley attributions (Lundberg and Lee 2017); to obtain valid confidence intervals for targeted parameters one uses purpose-built procedures such as generalized random forests (Wager and Athey 2018; Athey et al. 2019). None of these objects appears in MAHB. Instead, every numerical column, whether labeled “fixed effect,” “decision tree,” or “random forest,” reports a coefficient, a standard error in parentheses, significance stars, an R^2 , and an F . That template is the signature of ordinary least squares (or its within/fixed-effects variant), not of a tree-based estimator.

3. PCA is not an estimator comparable to a regression

In its estimation section, MAHB list “three techniques, principal component analysis, decision tree regression, and random forest regression,” set against the conventional panel estimators (FE, random effects, OLS, GMM) in order “to find the better estimation approach.” This places PCA in the role of a competing estimator, which is a category error.

PCA is an *unsupervised* method: it is the eigendecomposition of the covariance (or correlation) matrix of a set of variables, returning orthogonal linear combinations of maximal variance (Pearson 1901; Hotelling 1933; Jolliffe 2002; Jolliffe and Cadima 2016). It has no dependent variable, estimates no relationship between variables, predicts no target, and delivers no inference. A regression (OLS or fixed effects) is a *supervised* method that estimates the conditional relationship between an outcome and a set of regressors and provides inference about it (Athey

and Imbens 2019). Asking whether PCA or the fixed-effects model “estimates better” the WCF–profitability relationship is therefore not well posed: PCA estimates no relationship.

In MAHB, PCA is in fact used *upstream*, as feature construction: the master proxy for each construct is the first principal component of that construct’s proxy set, and the regression, tree, or forest is then run on these constructed variables. PCA is thus a preprocessing step, not an alternative to the estimator. This reading is confirmed by the comparison tables themselves (Tables 7, 8, and 13), which contrast only FE, tree, and forest and never contain a PCA column.

Using first principal components as the model’s variables also has two specific consequences that the article does not address. First, principal components are identified only up to sign: an eigenvector and its negative are equally valid solutions. Because both the dependent variable (profitability) and the key regressor (WCF) are first-component scores, the *sign* of the estimated WCF coefficient, which is the article’s central economic finding (an initially positive relationship), depends on an unstated orientation convention. MAHB report neither the loadings nor a rule fixing component orientation. Second, the statement that the first component of profitability “captures 70% variance from ten proxies, which means that PC1 of profitability is equal to the seven proxies” conflates a proportion of explained variance with a count of variables. Saying that PC₁ captures 70% of the variance means that, of the total information (the total dispersion) carried by the ten proxies, the first component summarizes 70%; it is an indicator of how well the single component represents the set. MAHB instead apply an invalid cross-multiplication, $0.70 \times 10 \text{ variables} = 7 \text{ variables}$, as though the component had “kept” seven of the proxies and “discarded” the other three. A proportion of variance and a number of variables are not commensurable, and the resulting claim has no defined meaning.

4. The reported coefficient tables cannot be tree-based output

4.1 The general diagnosis

Tables 7, 8, and 13 each report, for every variable, a coefficient with a standard error in parentheses and significance stars, followed by an R^2 and an F -statistic. As we noted in Section 2, these are not quantities that a regression tree or a random forest produces. The numerical content of these tables is therefore the output of a linear regression, irrespective of the heading placed atop each column.

Two features of the article corroborate this directly. First, the specification adds the squared term WCF^2 “to find the non-linear WCF–Profitability relationship” (Eq. 1, p. 75) and the interaction terms $WCF \times \text{dummy}$ and $WCF^2 \times \text{dummy}$ to capture moderation (Eqs. 2–3, p. 76). In a linear model these constructed terms are exactly how one permits curvature and interaction. In a tree or a forest they are redundant: the estimator approximates arbitrary measurable functions, including curvature and interactions, without any explicit construction (Hastie et al. 2009). Their presence is meaningful only under a parametric specification, and the article interprets them accordingly (a positive linear term with a negative quadratic term, read as an inverted U). Second, the econometric baseline is labeled the “fixed effect model” in Table 7 (p. 90) but the “ordinary least square model” in Table 8 (p. 91) for the same role, indicating that the estimator labels are used interchangeably.

4.2 The decision tree: the article’s primary estimator

The diagnosis applies with particular force to the regression tree, because the tree, not the forest, is the article’s primary estimator. The industry-level results (Table 9, p. 92) and the moderation analyses for firm size and leverage (Tables 10–12, pp. 95–99) are all labeled “Decision Tree

Regression,” whereas the random forest enters only as a robustness check (Section 6.8). A single regression tree predicts a piecewise-constant step function (Breiman et al. 1984): like the forest, it has no coefficient vector, no parametric standard errors, no significance levels, and no F -statistic. Yet Tables 9–12 report exactly these objects for every variable, in every industry. They are, on their face, linear-regression output presented under a “decision tree” heading.

One operation in the size and leverage analysis makes this unambiguous. To obtain a subgroup slope, the article *adds* a level coefficient and a dummy-interaction coefficient: the WCF slope for large firms is reported as $0.689 = 0.468 + 0.221$, and for high-leverage firms as $0.413 = 0.809 + (-0.396)$ (Sections 6.6, p. 94, and 6.7, p. 97). Summing a main effect and an interaction term to recover a conditional slope is precisely the algebra of a linear model with interaction dummies; it has no counterpart in a regression tree, whose prediction does not decompose into an additive main effect plus an interaction. The manually constructed $WCF \times$ dummy and $WCF^2 \times$ dummy terms serve the same parametric purpose and would be redundant in a tree, which already represents such interactions through its splits.

The decision-tree estimates also fail to reconcile across the paper. The full-sample WCF coefficient is reported as 0.335 in Table 7 (p. 90) but is referred to as 0.589 in the size and leverage discussions (Sections 6.6, p. 94, and 6.7, p. 96), and the component values used there (0.468, 0.221, 0.809, -0.396) appear in none of the tables: Tables 10–12 are disaggregated by industry and contain no pooled full-sample regression with a single size or leverage dummy. The narrative thus relies on decision-tree coefficients that the tables do not contain. The companion full-sample value quoted for the random forest, 0.598 (Section 6.8, p. 100), differs from the decision-tree 0.589 by a transposition of digits and matches neither the corresponding table value (0.323, Table 7) nor that of Table 13 (0.403). Finally, the table references in these sections do not resolve: Section 6.6 locates the coefficients of Eq. (2) in “Table 6” (the year-wise sample distribution) and then in “Table 7” (which contains no size interaction), although the size-dummy decision-tree estimates are in Table 10.

5. The reported estimates are mutually inconsistent

5.1 Table 13 combines outputs from two different models

Table 13 (p. 101) is titled “Random forest regression algorithm results.” Comparing its first (full-sample) column with the two corresponding columns of Table 7 (p. 90) shows that it is assembled from two sources. Table 1 below reproduces them.

In the column that is supposed to report the random forest, the WCF and WCF^2 entries reproduce, to the reported precision, the *fixed-effects* values (not the random-forest values), whereas the seven control coefficients, the R^2 , and the F are the *random-forest* values.

This juxtaposition is logically impossible for a single model. For a linear regression estimated on a fixed sample, the coefficient vector $\hat{\beta}$ and the coefficient of determination R^2 are not independent quantities: the latter is fully determined by the former and by the data. Writing SSR for the residual sum of squares and SST for the total sum of squares,

$$R^2 = 1 - \frac{SSR}{SST}, \quad SSR = \sum_i (y_i - x_i' \hat{\beta})^2.$$

Once the data (X, y) and the vector $\hat{\beta}$ are fixed, the residuals (hence SSR , hence R^2) are mechanically determined. The reported $R^2 = 0.542$ is the one that accompanies, in Table 7, the random-forest coefficient vector, which includes $WCF = 0.323$. Replacing that entry with $WCF = 0.403$ (the distinct fixed-effects value) changes the residuals and therefore the R^2 ; the

Table 1: Full-sample estimates as reported by MAHB (coefficient; standard error in parentheses).

Variable	Fixed effects (Table 7)	Random forest (Table 7)	Table 13 ("random forest")	Apparent source
WCF	0.403 (0.026)	0.323 (0.0138)	0.403 (0.026)	= <i>fixed effects</i>
WCF ²	−0.010 (0.005)	−0.002 (0.0029)	−0.010 (0.005)	= <i>fixed effects</i>
Liquidity	0.050 (0.019)	−0.036 (0.0099)	−0.036 (0.0099)	= <i>random forest</i>
Investment	−0.143 (0.021)	0.044 (0.0111)	0.044 (0.0111)	= <i>random forest</i>
Growth	−0.389 (0.017)	−0.297 (0.0078)	−0.297 (0.0078)	= <i>random forest</i>
Asset tangibility	−2.026 (0.493)	−1.493 (0.2512)	−1.493 (0.2512)	= <i>random forest</i>
R&D intensity	−0.070 (0.038)	−0.107 (0.0193)	−0.107 (0.0193)	= <i>random forest</i>
Dividend payout	−5.784 (0.144)	−7.103 (0.0748)	−7.103 (0.0748)	= <i>random forest</i>
R^2	0.342	0.542	0.542	= <i>random forest</i>
F	161.16	1997	1997	= <i>random forest</i>

pair (0.403, 0.542) cannot come from one fitted model. More fundamentally, the coefficient vector of Table 13 is itself hybrid (WCF and WCF² inherited from the fixed-effects column, the seven controls inherited from the forest column), so it is not the output of any single estimation. Table 13 therefore appears to place outputs from at least two models under one heading.

5.2 Three values for one coefficient

The full-sample random-forest coefficient on WCF is reported three times, with three incompatible values: 0.598 in the text of Section 6.8 (p. 100), 0.403 in Table 13 (p. 101), and 0.323 in the random-forest column of Table 7 (p. 90). The corresponding WCF² coefficient is reported as −0.015, −0.010, and −0.002, respectively. These sets cannot all be correct.

6. The model comparison is not valid on its own terms

The article's central claim, that the machine-learning models "explain the variance better," rests on a comparison of in-sample fit: the regression tree ($R^2 = 0.470$) and the random forest ($R^2 = 0.542$) attain higher R^2 than the fixed-effects model ($R^2 = 0.342$). Even granting that a forest was fit and that 0.542 is genuinely its fit, this comparison is uninformative.

A deep tree, and an ensemble of deep trees, can fit a training sample arbitrarily closely; their in-sample R^2 is mechanically inflated relative to a parsimonious linear model. Comparing in-sample fit across model classes of very different flexibility therefore says little about how the models would generalize (Hastie et al. 2009; James et al. 2013). The appropriate comparison uses held-out or cross-validated error, or the out-of-bag error for the forest. MAHB describe a train/validation/test split (Section 4, pp. 76–78) but never state on which sample the reported R^2 and F are computed; if these are in-sample, the comparison is empty. The point is reinforced by the methodological literature the article itself cites on the econometric use of machine learning (Mullainathan and Spiess 2017). In addition, R^2 and F are not the natural performance metrics for a forest, and an F -statistic for a random forest is not defined.

A separate concern affects fit in every model. The data are winsorized twice (Section 5, p. 79). The raw proxies are first winsorized "up to 5%" at each tail as the sample is assembled, and then, after the principal components have been formed, the master proxies are winsorized again "at 10%" in what the article describes as "the final step of data preparation." Winsorizing

replaces extreme values with percentile bounds rather than deleting them, so the procedure compresses the dispersion of the data twice; under the natural reading of “10%” as 10% per tail, roughly one-fifth of the observations are moved to the bounds at the second step alone, on top of the first pass. Two issues follow. Conceptually, winsorizing the principal components is winsorizing linear combinations of the proxies, which distorts the very structure PCA was used to construct. Empirically, because both the dependent variable and the regressors are treated this way, the most extreme and hardest-to-fit observations are pulled toward the center; this mechanically lowers the residual sum of squares (raising R^2), attenuates slope estimates, and narrows standard errors. The treatment of influential observations is known to affect inference materially in this literature (Leone et al. 2019), yet no justification or sensitivity analysis is provided for either pass.

7. Further inconsistencies

The following additional discrepancies, while individually minor, bear on the reliability of the reported results.

- The final sample is stated as 15,238 observations and, two sentences later on the same page, as 15,239 (Section 5, p. 79).
- The textile sub-sample contains 167 observations in the year-wise distribution (Table 6, p. 89) but 679 in the industry regressions (Tables 9–12, pp. 92–99).
- In Table 13, only the first F -statistic carries significance stars (1997***); the remaining three (731.93, 698.78, 723.24) are unstarred, without explanation.
- In-text table references repeatedly fail to match their content: beyond the cases noted in Section 4, Section 6.8 refers to “Table 10” for random-forest results that should appear in Table 13.
- Several reference-list entries are inaccurate. The CART monograph (Breiman, Friedman, Olshen, and Stone, 1984) is listed three times, as 1984a, 1984b, and 1999, and the 1999 reprint is the source cited for the random forest, rather than Breiman (2001), the random-forests paper, which is also listed. The entry for the machine-learning study of director selection is garbled, with a truncated title and an incorrect journal, year, volume, and pages, given as “Erel, Stern, Tan, and Weisbach (2018), Selecting directors using. Mach. Learn. 2003:1–57”; the published article is Erel, Stern, Tan, and Weisbach (2021), “Selecting directors using machine learning,” *Review of Financial Studies* 34(7), 3226–3264.

8. Conclusion

The issues documented above are of three kinds. PCA is presented as a competing estimator, although it estimates no relationship and functions in the paper only as upstream feature construction. The coefficient tables labeled “decision tree” and “random forest” report quantities that tree-based methods do not produce and are, on their face, linear-regression output; the manually added quadratic and interaction terms confirm a parametric specification. And the reported numbers are internally inconsistent, most clearly in Table 13, which combines fixed-effects and random-forest values within a single column, and in the three conflicting values of the full-sample random-forest coefficient on WCF . Finally, even on the authors’ own framing, the in-sample R^2 comparison cannot support the claim that the machine-learning models explain the variance better.

A practical consequence concerns reproducibility. The study draws on the proprietary CSMAR (China Stock Market and Accounting Research) database, distributed under institutional license by GTA Information Technology and not freely available, and it provides no replication dataset, no estimation code, and no list of the retained firm identifiers. A reader therefore cannot reconstruct the reported estimates from the methods as described, nor verify them against the data; resolving the inconsistencies documented here rests with the authors.

References

- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685–725.
- Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *Annals of Statistics*, 47(2), 1148–1178.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.
- Greene, W. H. (2018). *Econometric analysis* (8th ed.). Pearson.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6/7), 417–441, 498–520.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202.
- Leone, A. J., Minutti-Meza, M., & Wasley, C. E. (2019). Influential observations and inference in accounting research. *The Accounting Review*, 94(6), 337–364.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- Mahmood, F., Ahmed, Z., Hussain, N., & Ben-Zaied, Y. (2025). Working capital financing and firm performance: A machine learning approach. *Review of Quantitative Finance and Accounting*, 65(1), 71–106.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(11), 559–572.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8, 25.
- Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228–1242.
- Wallace, M. L., Mentch, L., Wheeler, B. J., Tapia, A. L., Richards, M., Zhou, S., Yi, L., Redline, S., & Buysse, D. J. (2023). Use and misuse of random forest variable importance metrics in medicine: Demonstrations through incident stroke prediction. *BMC Medical Research Methodology*, 23, 144.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.