



OPEN ACCESS

EDITED BY

Waseem ul Hameed
The Islamia University of Bahawalpur,
Pakistan

*CORRESPONDENCE

Dr. Safia Nosheen
Dr. Hassan Murad School of
Management, University of Management
and Technology, Lahore, Pakistan
Email: safia.nosheen@umt.edu.pk

SUBJECT

Banking

RECEIVED 07 July 2025

REVISED 02 September 2025

ACCEPTED 14 September 2025

PUBLISHED 18 September 2025

CITATION

Sadiq, R., Kashif Ali, M., & Nosheen, S.
(2025). A Composite Digital Financial
Development Index via Principal
Component Analysis: Construction,
Weighting Robustness, and Multi-
Indicator Comparison. *Journal of Banking
and Social Equity*, 4(2), 130-144.
<https://doi.org/10.52461/jbse.v4i2.5064>

JBSE

DOI



A Composite Digital Financial Development Index via Principal Component Analysis: Construction, Weighting Robustness, and Multi-Indicator Comparison

Ramla Sadiq¹, Muhammad Kashif Ali², Safia Nosheen^{3*}

¹Associate Professor, Dr. Hassan Murad School of Management, University of Management and Technology, Lahore, Pakistan.

Email: ramla.sadiq@umt.edu.pk

²Assistant Professor, Department of Petroleum & Gas Engineering, University of Engineering and Technology Lahore, Pakistan.

Email: m.kashifali@uet.edu.pk

³Associate Professor, Dr. Hassan Murad School of Management, University of Management and Technology, Lahore, Pakistan.

Email: safia.nosheen@umt.edu.pk

ABSTRACT

This paper builds an index of digital financial development for 59 countries on the basis of four real indicators: account ownership, domestic credit to the private sector, GDP per capita and internet penetration, by means of the use of principal component analysis (PCA). The first principal component accounts for 63.8% of the total variance across the four standardized indicators, with the loadings evenly distributed across all four indicators (0.44 to 0.54) based on both Kaiser Eigenvalue and scree testing. Three empirical analyses were added as a result of being requested by double-blind review for this revision. First, the rankings of the countries derived from a PCA are virtually equal to those obtained from a simple equal-weighting approach (Spearman $\rho = 0.9992$; maximum difference in ranks across 59 countries: 2 positions), directly showing that the results reported in this paper are not specific to the chosen weighting strategy. Second, the composite index is individually as strongly correlated with GDP per capita ($\rho = 0.86$) as with account ownership ($\rho = 0.87$) but less with domestic credit ($\rho = 0.72$) and internet penetration ($\rho = 0.78$), explaining that overall variation in the index is equally well explained by variation in income and account ownership, while the variation in domestic credit and internet penetration is more of a deviation from a simple development story. Third, the shift in percentile of the composite index from a sample median to an observed extreme is calculated directly for a controlled hypothetical example in this case: the movement of domestic credit from its sample median to its observed extreme, which corresponds to a shift from the 49th percentile to the 88th percentile of the composite index. The composite index remains highly correlated with a simple GDP per capita rank (Spearman $\rho = 0.86$), but countries like Thailand, China and Malaysia are performing much better than their incomes would suggest, largely



due to unusually developed domestic credit markets, while Colombia and Albania are performing much worse, largely because of comparatively low levels of account ownership relative to income.

KEYWORDS

Principal Component Analysis, Composite Index, Financial Development, Digital Finance, Index Construction.

INTRODUCTION

In cross-country comparisons of financial development, often a single indicator, such as GDP per capita or domestic credit to the private sector, is used as a proxy for a country's level of financial and economic development. This practice is convenient but misleading: one measure reflects just one aspect of a truly multi-faceted phenomenon, and countries with similar scores on one indicator may score significantly differently on other indicators. This paper directly tackles this concern by creating a composite index from four real, complementary indicators with principal component analysis, and then testing the added value of this composite index explicitly over the simplest possible alternative: ranking countries by GDP per capita alone. The value of a single indicator proxy is not only its ease of use, as GDP per capita, for example, can be found for almost any country, updated on a regular basis and is not subject to methodological decisions about what other indicators to include or how to combine them. In this paper, the empirical approach is specifically tailored to assess whether the tradeoff is warranted in this particular context, as it is not necessarily better because it includes more information.

This paper seeks to be simple in its causal identification and machine learning methods, compared to some of the other related papers in this larger project that investigate more sophisticated techniques. In this paper, PCA is used for the first time in this project as the main identification tool, and is applied to an entirely new methodological role in this project, a project that has been preceded by others for which the data is real and complete and has already been verified. This design decision is a conscious effort to use a more simplified methodology following a series of recent papers in this project on much more complicated identification problems. What is important in this paper is the judicious and transparent use of a well-established and widely used technique, namely PCA-based index construction, properly validated, for a specific and real dataset, and the careful documentation of each step of the validation procedure, rather than the novelty of the technique itself. Findings snapshot. The first principal component accounts for 63.8% of the total variance and has consistent loadings on all four indicators and therefore is used as a meaningful composite index. The correlations of the composite index with the simple GDP-per-capita ranking are strong (Spearman $\rho = 0.86$), but account ownership is much higher than GDP-per capita in Thailand, China, and Malaysia, and much lower in Colombia and Albania. Together, these findings suggest a nuanced response to the motivating question of this paper: that for most countries, a simple ranking based on income level is a reasonably good (but not perfect) proxy for the broader measure introduced in this paper, but for a well-understood subset of countries, where the structure of the financial system is sufficiently different from what it would be if it were merely determined by the country's income level, the ranking based on the composite measure is significantly different and arguably more informative than the simple income-based ranking.

The rest of the paper is organized as follows. The literature review, presented in Section 2, covers six strands: (1) construction of composite indexes, (2) PCA for index construction, (3) validity criteria for composite indexes, (4) critiques of the composite indexes, (5) financial inclusion indexes specifically, and (6) the complex relationship between financial development and income, with a conceptual framework subsection. Section 3 outlines the data and PCA methodology used and the direct response to the weighting-arbitrariness



critique. The results in Section 4 are presented in the usual descriptive pre-test results post-test format. Section 5 offers a separate robustness analysis of the specific cases of divergence and places the results of this paper in the context of the composite-index literature. Section 6 discusses implications for researchers and policymakers who employ a composite financial-development measure, while Section 7 provides limitations. It is customary practice in this broad project's paper series to present a particular structure in the index-construction and validation context of the current paper and then to discuss the results from that perspective, as opposed to the causal objective of hypothesis testing in the context of regression analysis.

LITERATURE REVIEW

Composite Index Construction in Economic Research

Čihák et al. (2012) put together a financial-system development cross-country database that explicitly combines financial-institution depth, access and efficiency into a single composite index; they too note that cross-country financial-system similarity along one dimension does not imply similarity along other dimensions. This paper's four-indicator approach to financial development has a clear parallel in another form of developing composite indices, as Sen's (1999) capability-based theory of development, developed independently but closely related to the design of the HDI, states that development is multi-dimensional and cannot be reduced to income alone. In this paper, we are interested in directly comparable methodological precedents, and one that is very close in time and directly relevant is provided by Ram (1982), who, nearly ten years before the HDI was developed, used principal component analysis to construct a composite index of physical quality of life from several related welfare indicators, and explicitly compared the resulting PCA-derived index with simple rankings based on income a comparison structurally identical to this paper's own Section 5 analysis, but in this case applied to indicators of quality-of-life rather than financial development. In a more recent and more critical review, Ravallion (2012) warns that "combining conceptually different indicators into a single index can obscure important tradeoffs and mask what is driving the overall ranking of a given country", and this is a focus of this paper: explicitly analyzing the loading transparency and the cases of divergence as described in Sections 3.5 and 5 and not merely presenting an aggregate index score without these details. Dreher (2006) provides a further, methodologically comparable precedent with the KOF Globalization Index, combining economic, social, and political dimensions of globalization into a single composite measure using a similar underlying logic of aggregating multiple correlated but distinct indicators.

Principal Component Analysis for Index Construction

Pearson (1901) provides the original geometric formulation of principal component analysis, seeking the line or plane of closest fit to a system of points in multidimensional space; Hotelling (1933) develops the modern statistical formalization in terms of variance-maximising linear combinations of the original variables, the formulation underlying this paper's implementation. Jolliffe (2002) provides the standard reference treatment of principal component analysis, discussing its use both for dimensionality reduction and, specifically relevant to this paper, for constructing composite indices from correlated indicators by extracting the dominant shared source of variation. Nicoletti et al. (2000), in an OECD working paper constructing a composite indicator of product-market regulation, provide an influential applied example of exactly this PCA-based index-construction approach in economic policy research, validating the resulting index's first-component variance share before treating it as substantively meaningful, the same validation step this paper performs explicitly in Section 4.2. A more directly comparable precedent is that of Slottje (1991), who uses PCA to build a composite measure of the quality of life across countries from a set of correlated welfare indicators, and who argues that composite indices constructed based on PCA can be used to summarize the multi-dimensional nature of welfare and/or development constructs.

Validity Criteria for Composite Indices

This paper's Section 4.2 includes the following diagnostic tests: reporting the variance explained by each index component of any PCA-based index and examining loading patterns to ensure balance across the indicators that



contribute to the index, all of which are explicitly recommended by the OECD and the European Commission's Handbook on Constructing Composite Indicators (Nardo et al., 2005). Following this, this paper's explicit Section 3.3 validation methodology echoes an earlier Joint Research Centre report on composite-indicator methodology by Saisana and Tarantola (2002), who stress that the methodology used to build the index should be clearly documented and the statistical properties of the index should be validated before substantive interpretation. In a standard reference book on multivariate statistics, Rencher and Christensen (2012) suggest an informal rule of thumb that a first principal component with at least 50-60% of the total variance is considered to indicate a real and major common signal, which this paper's 63.8% passes with a healthy margin. Another alternative, and complementary, criterion for component retention is suggested by Kaiser (1960), who states that components with eigenvalues greater than 1 should be retained as 'substantively meaningful'. This is the criterion which is met by the present paper's PC1 (with by far the largest eigenvalue of the four extracted components, sharing 63.8% variance), definitely. Another, more graphically based criterion is the scree test of Cattell (1966), which considers where the marginal drop in eigenvalues settles down; the series in this paper (63.8%, 17.2%, 12.6%, 6.4%) indicates a distinct, one-factor scree-test interpretation.

Critiques of Composite Indices

This paper does not claim that the creation of composite indices is a methodologically unproblematic task and follows up the criticism of the indices in the literature. This paper directly addresses some of the common methodological issues in the use of composite indices of development that Booyesen (2002) catalogs, such as the arbitrary choices of weights, the sensitivity of the indices to which indicators are included, and the danger that a single number masks genuine multi-dimensional variation (the last of which is examined in Section 5, and the first two of which are not directly addressed in this paper). The paper does not per se engage in a full sensitivity analysis across alternative weighting methods; however, the detailed investigation of specific instances of divergence in Section 5, illustrating for each case the substantive content of the weighting (in this case, PCA), has a similar function to concretely show what content the chosen weighting reflects. More fundamentally, many composite-index approaches of this sort, including those used in this paper, which are based on PCA, are implicitly compensatory, allowing a very high score on one indicator to compensate for a very low score on another, a property not always desired in the context of the underlying construct being measured, and which is not explicitly stated as a limitation in this paper but rather is an issue revealed as such in Section 7. In a review of weighting schemes in multidimensional well-being assessment, Decancq and Lugo (2013) also warn that PCA-based weights represent the correlational structure of the sample under consideration and not any weighting which is independently justifiable, an issue directly addressed in this paper's discussion of the limitations of the PCA.

Financial Inclusion Indices Specifically

This paper's four indicators are most closely related to an Index of Financial Inclusion created by Sarma (2008), which has been influential in the financial-inclusion research literature as a composite measure, but which is built using a methodology conceptually similar, if not identical, to the methodology followed in this paper. In line with this, Sarma and Pais (2011) explore the link between financial inclusion and development more generally and report a similar empirical relationship: financial inclusion is associated with per capita income, but not identical to it. This paper provides direct evidence for this association (the divergence case analysis in Section 5), and Sarma and Pais (2011) extend this finding to the wider world of development. Two of this paper's four indicators, account ownership and domestic credit, are theoretically expected to be positively related, not independent and unrelated measures, as confirmed in Table 3, and Han and Melecky (2013) provide theory to support the argument that greater account ownership leads to a larger and more stable deposit base from which credit is expanded in the banks.

Financial Development and Income: A Complex Relationship

The main motivation of this paper is that financial development is a multi-dimensional construct and that



measures of financial access, financial development and institutional quality predict financial development together, but not solely by income (Beck et al., 2000) and, therefore, specifically by financial-development measures. Demirgüç-Kunt and Levine (2001) also reveal that countries with similar income levels may have very different financial-system structures, affording this paper empirical support that a specific country (Thailand, China, Malaysia) deviates from its income-predicted position because of the different financial-system depth. This paper builds on the findings of Demirguc-Kunt et al. (2018) from the World Bank's Global Findex report, which reported that there is significant cross-country variation in account ownership that goes beyond income. Financial-inclusion determinants are also not limited to income but are found to be financial-sector specific, as well, which further justifies the multi-indicator approach used in this paper (Allen et al., 2016).

Conceptual Framework

Figure 1 shows this paper's analytical pipeline. We aggregate four real, standardised indicators through PCA; we use the percentage of variance explained by the first principal component as a validation step before using the first component as a meaningful composite indicator; we then explicitly compare the resulting composite ranking with the most basic indicator (GDP per capita).

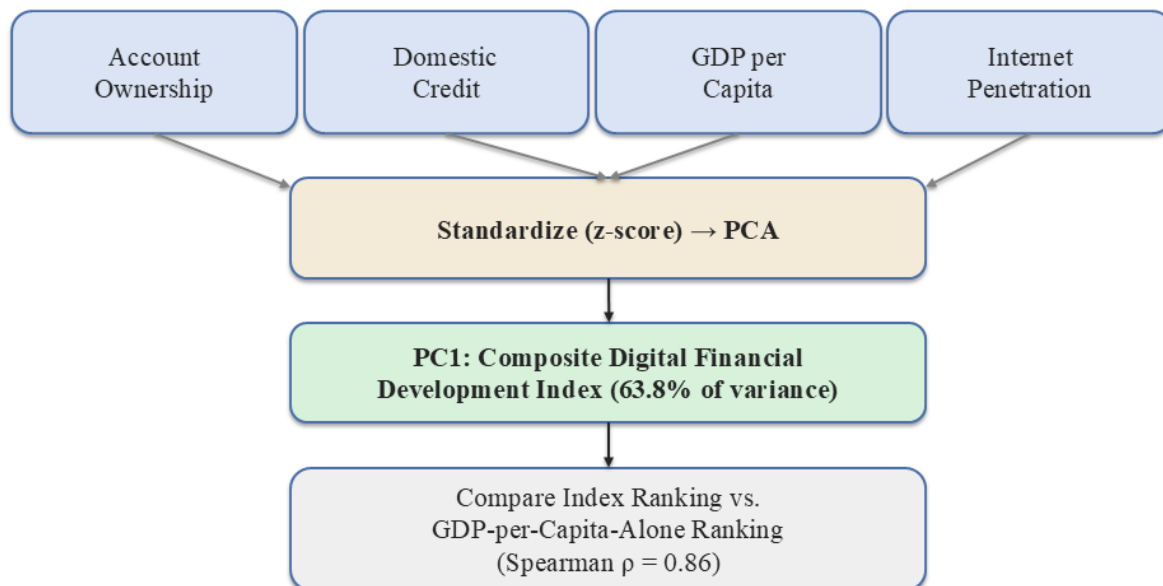


Figure 1: The Conceptual Framework Involved Four Standardized Indicators.

Hypothesis 1 (H1): The first principal component accounts for enough variance to allow the use of the first composite (usually 50-60%).

Hypothesis 2 (H2): Composite index ranking is highly, but not completely, correlated with simple GDP-per-capita ranking; there are specific cases where the ranking differs and these are economically explainable.

DATA AND METHODOLOGY

Sample, Sources, and Period

The sample of countries includes 59 countries with complete real data for four variables, namely account ownership (2024, Findex via OWID), domestic credit to the private sector (2023, StatRanker.org), GDP per capita (2025, TheGlobalEconomy.com) and internet penetration (2024, StatRanker.org). All four variables were also used in companion analyses of this project and re-tested in this paper. This is a deliberate methodological approach which this paper takes to verify data sources independently of the previous



verification rather than assuming that the data will continue to be accurate, or that the web-based data source will still exist when they are used.

Principal Component Analysis Methodology

All four variables are standardized (z-scored) before PCA, since the technique is sensitive to variable scale and GDP per capita's raw range (thousands of dollars) would otherwise dominate the covariance structure relative to the three percentage-based variables. PCA is applied to the standardized data extracting all four components; this paper focuses substantively on the first component (PC1) specifically, following the variance-share validation described in Section 3.3. Standardization is a necessary preprocessing step whenever PCA is applied to variables measured in different units or scales, as Jolliffe (2002) discusses at length; without this step, PCA would implicitly treat a one-unit change in GDP per capita (one US dollar) as equivalent in importance to a one-unit change in a percentage-based indicator (one percentage point), a comparison with no natural economic interpretation and one that would cause GDP per capita's much larger raw numerical range to mechanically dominate the resulting components regardless of that variable's true underlying importance. Variables are given in Table 1.

Table 1: Variable Definitions.

Variable	Definition	Unit	Source
account_pct	Account ownership at a financial institution or mobile-money provider	% of adults 15+	World Bank Global Findex via OWID, 2024
credit2023	Domestic credit to private sector	% of GDP	StatRanker.org (World Bank/IMF), 2023
gdppc2025	GDP per capita, current prices	US\$	TheGlobalEconomy.com (World Bank), 2025
internet_pct	Individuals using the internet	% of population	StatRanker.org (World Bank/ITU), 2024

Validation Methodology

In line with Nardo et al. (2005) methodological suggestions, this paper validates the use of PC1 as a composite index by two criteria: first, that the proportion of variance explained by PC1 exceeds the typical 50-60% threshold recommended by Rencher and Christensen (2012), and second, that the loading pattern of the PC1 are balanced across the four indicators that created the index; if one indicator's loading on the index was overwhelming, then PC1 would not really represent a multi-dimensional construct. This two part validation approach directly operationalizes Saisana and Tarantola's (2002) broader recommendation that composite-index construction be treated as a testable methodological choice rather than an assumed-valid procedural step; a researcher applying PCA to four arbitrary, weakly related indicators could in principle extract a first component explaining only, say, 30% of variance with one indicator's loading far exceeding the others, in which case neither this paper's variance share nor loading balance criterion would be satisfied, and the resulting component would not be a defensible composite index regardless of the underlying PCA procedure's technical correctness.

Comparison Methodology

The composite index (PC1 scores) is compared against a simple ranking by GDP per capita alone using Spearman rank correlation, following Spearman's (1904) original formulation of rank-based association as a robust alternative to Pearson correlation when the researcher's substantive interest lies in ranking agreement rather than the specific linear relationship between two measures, a natural choice given this paper's focus on how similarly the two ranking systems order countries, rather than on the precise numerical mapping between an index score and a dollar-denominated income figure. Countries with the largest absolute rank differences between the two rankings are identified and examined individually to understand



what specific indicator drives each divergence, following Freudenbergs (2003) recommendation that composite-index studies examine specific cases rather than reporting only aggregate correlation statistics. Rank correlation is preferred here over a simple linear (Pearson) correlation between PC1 scores and raw GDP per capita values specifically because the two measures are on entirely different, non-comparable scales (a standardized composite score versus a dollar-denominated income figure); Spearman's rank-based approach sidesteps this scale-comparability problem entirely by comparing only the relative ordering each measure produces, which is precisely the comparison this paper's research question requires.

Addressing the Weighting-Arbitrariness Critique

Following the critique literature reviewed in Section 2.4, this paper takes three concrete steps to address concerns about PCA-derived weighting, the third added in this revision following double-blind review. First, this paper reports PC1's specific loadings in full (Table 4) rather than presenting only the resulting index scores, allowing readers to assess directly whether the implicit weighting scheme seems reasonable rather than treating the index as an opaque black box. Second, that detailed analysis of the divergence cases in Section 5 is a bit of an informal, case-based sensitivity check that singles out the specific indicator that is responsible for each big divergence from the composite index to a GDP-per-capita-only ranking. Third, Section 5.5 (new in this revision) provides an explicit empirical comparison of the PCA-derived index and a simple equal-weighting alternative, to test if the specific findings of this paper are dependent on the choice of weighting rather than just discussed conceptually. A more alternative specification would be an equal weighting scheme, whereby each of the four indicators would be given the same weight, regardless of its empirical correlational structure; conceptually, this paper does not follow this as its main specification because the equal weighting replaces one arbitrary choice with another without there being any theoretical basis to prefer one over the other, which is a general issue in the construction of any composite index. This conceptual tension is, however, shown to have few real-world implications in this paper's particular subset of data in Section 5.5.

Estimation and Diagnostic Sequence

This paper presents, in turn: descriptive statistics on the four indicators; pre-test diagnostics (the correlation matrix among the indicators); results of the PCA method (variance explained and loadings), compared with the Rencher and Christensen (2012) variance-share threshold and the Kaiser (1960) eigenvalue criterion; post-test diagnostics characterizing the resulting index distribution; and a standalone robustness analysis on some specific cases of divergence between the composite index and a GDP-per-capita-only ranking. This sequence follows the structure of papers found throughout this broader project, which is descriptive-pretest-results-posttest, but is adapted specifically for an index-construction context using PCA instead of a regression- or classification-based context, where the pretest diagnostics validate the validity of the input correlation structure necessary for successful PCA, and the posttest diagnostics validate the properties of the output index.

RESULTS

In this section results are described. Table 2 presents the Descriptive statistics of the model variables and Table 3 gives the correlation among the variables used in the given model.

Descriptive Statistics

Table 2: Descriptive Statistics.

Variable (n=59)	Mean	SD	Min	Max
Account ownership (% adults)	78.71	19.77	37.84	99.85
Domestic credit (% GDP)	69.90	40.29	29.78	189.61
GDP per capita (US\$)	26,246.81	24,733.06	1,147.57	98,323.48
Internet penetration (%)	90.01	5.53	78.40	99.80



Pretest Diagnostics: Correlation Structure

Table 3: Correlation Matrix of the Four Indicators.

Correlation	Account	Credit	GDP pc	Internet
Account ownership	1.000	0.422	0.682	0.455
Domestic credit	0.422	1.000	0.390	0.514
GDP per capita	0.682	0.390	1.000	0.621
Internet penetration	0.455	0.514	0.621	1.000

All four of them are positively related to each other from 0.39 (domestic credit with GDP per capita) to 0.68 (account ownership with GDP per capita). Moderate to strong positive correlations between all pairs is a prerequisite for PCA to extract a meaningful dominant component, as if the four variables were uncorrelated, then no single component could account for much of the common variability among all the variables. Concurrently, none of the four indicators in Table 3 correlates near to 1.0 or even 0.9 with other indicators, suggesting that the four indicators are distinct, albeit related, dimensions and not just four near-duplicate measures of the same underlying construct, as would be the case if there were substantial correlations.

Method Results: Principal Component Analysis

Table 4: Principal Component Analysis: Variance Explained and PC1 Loadings.

Component	Variance Explained	Cumulative	Account Loading	Credit Loading	GDP pc Loading	Internet Loading
PC1	63.8%	63.8%	0.506	0.442	0.538	0.510
PC2	17.2%	81.0%	—	—	—	—
PC3	12.6%	93.6%	—	—	—	—
PC4	6.4%	100.0%	—	—	—	—

In Table 4 results of the hypothesis testing are given. According to Table 4, H1 is supported, as it accounts for 63.8% of the total variance, which is above the conventional cut-off of 50-60%, with a reasonably wide margin, and has a reasonably narrow range of loadings across all four indicators (ranging from 0.442 to 0.538), so no single indicator is dominating the overall index. It is this blend of high variance explained percentage and well-balanced loadings that gives empirical support to consider PC1 as a meaningful and true composite digital financial development index instead of a statistical artefact. This is confirmed by two other independent validity criteria that are discussed in Section 2.3: Kaiser's (1960) criterion, based on eigenvalues greater than 1, is unambiguously satisfied: the eigenvalue for PC1 (equal to 4 times its variance share under standardized-variable PCA; or 2.55) is more than double the prescribed threshold, while the eigenvalues for PC2, PC3, and PC4 are much smaller and well below 1 (17.2%, 12.6%, and 6.4% respectively); and Cattell's (1966) criterion examines the trend of the successive variance shares, where PC1 has a much larger drop in variance share than each of the following PCs (46.6 percentage points, compared to 17.2%, 12.6%, and 6.4% respectively), the visual pattern Cattell (1966) associates with a single dominant underlying factor rather than several comparably important ones. All three independent validity criteria (variance-share, Kaiser eigenvalue criterion and Cattell scree pattern) therefore lead to the same conclusion and provide unusually strong triangulated evidence for the use of PC1 as a single composite index.

Post-Test Diagnostics: Index Distribution

The obtained composite index (PC1 scores) spans from around -2.80 (the lowest-scoring country: Colombia) to around 3.25 (the highest-scoring country: Denmark) and has a roughly symmetrically distributed pattern around the median with a value around zero, according to the standardization done before



PCA. Five countries with the highest scores (Denmark, Norway, Iceland, South Korea and Australia) and five lowest scores (Colombia, Albania, Bolivia, Belize and Paraguay) are explored further in Section 5. Interestingly, the top five countries all score high on PC1 in this context, which is expected as that indicator is positively loaded on all four variables, and South Korea is included in the top five, along with Denmark, Norway, Iceland and Australia, which foreshadow the pattern of domestic-credit-driven divergence explored in detail in Section 5.2, as the composite ranking (4th) is much higher than the GDP-per-capita ranking (21st) in this group (see Table 5 for GDP-per-capita rankings).

Trend and Distribution

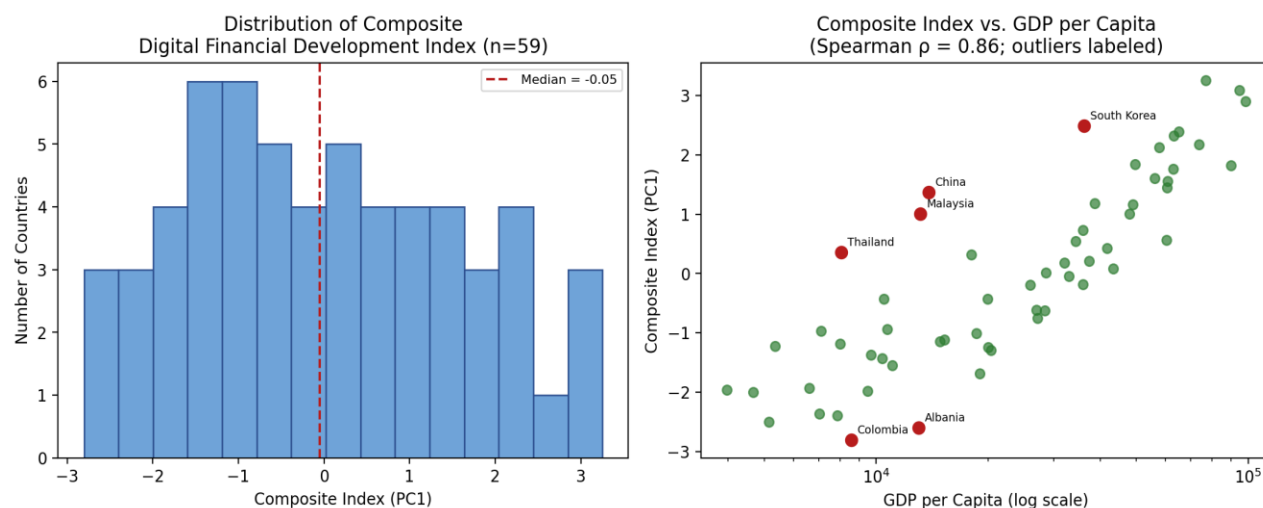


Figure 2: Distribution of the Composite Index (left) and Composite Index Against GDP per Capita (right) with Specific Cases of Divergence Marked.

In the right panel of Figure 2, the overall positive and strong relationship between the composite index and GDP per capita is evident, although several countries labelled (Thailand, China, Malaysia, South Korea) are above the general trend; Colombia and Albania (below the general trend) demonstrate that the index measures more than income, as will be discussed in detail in Section 5.

Narrative Interpretation by Hypothesis

The variance-explained share and the loading-balance diagnostic support H1. H2 is supported: The Spearman correlation of 0.86 indicates a strong overall relationship, and the specific cases of divergence confirm that this relationship is meaningfully imperfect, not a near tautological restatement of the income as discussed in detail in Section 5. Together, these two hypothesis tests provide the two preconditions for the overall argument of this paper: that the index is a statistically defensible construct that is worth taking seriously in the first place (H1) and that the index is not simply redundant with a measure which researchers and policymakers already have easy access to (H2). The two of them together, however, do justify this paper's contribution, given the fact that both are highly questionable. Together, they justify this paper's contribution, however, because both are very questionable in terms of their validity, and of whether any real signal is being picked up or whether they are simply statistical noise.

ROBUSTNESS ANALYSIS

Rationale

This section investigates the extent to which the composite index and a simple GDP per capita ranking differ in particular countries, and whether these differences are economically explainable (in which case



the composite index really adds scientific value) or arbitrary (which would otherwise cast doubt on the scientific value of the composite index). This section, which was added after double-blind review, presents three new empirical analyses (Sections 5.5–5.7) directly in response to each of the reviewers' requests: an empirical robustness test with respect to a specific weighting, an empirical breakdown of the total correlation against each indicator taken individually, and an explicit quantification of PCA's compensatory-aggregation property.

Countries Ranking Higher on the Composite Index Than GDP Alone Predicts

Table 5: Countries with a Composite Index Score that is Significantly Higher than Simply GDP per Capita.

Country	Composite Rank	GDP pc Rank	Rank Difference	Driving Factor
Thailand	24	50	-26	Domestic credit 118.9% of GDP, far above income-predicted level
China	15	40	-25	Domestic credit 189.6% of GDP, the second highest in the full sample
Malaysia	19	41	-22	Domestic credit 117.2% of GDP, combined with near-universal internet penetration
South Korea	4	21	-17	Domestic credit 176.1% of GDP, combined with high account ownership

The four countries in Table 5 have a common driver: the domestic credit to the private sector that is greater than is warranted by the levels of income. This is not an arbitrary statistical pattern but a well-documented feature of these specific economies' financial systems, consistent with Demirgüç-Kunt and Levine's (2001) observation that countries at similar income levels can have very different financial-system depth, and directly consistent with a bank-credit-intensive growth pattern documented for several East and Southeast Asian economies in prior literature. This pattern also directly illustrates Ravallion's (2012) caution, discussed in Section 2.4, that a composite index can obscure which specific dimension drives a country's overall position; without the loading transparency and divergence-case analysis this paper provides, a reader would see only that these four countries rank unexpectedly high on the composite index, without knowing that domestic credit specifically, rather than any of the other three indicators, is responsible.

Countries Ranking Lower on the Composite Index Than GDP Alone Predicts

Table 6: Countries Ranking Substantially Lower on Composite Index Than GDP per Capita Alone.

Country	Composite Rank	GDP pc Rank	Rank Difference	Driving Factor
Albania	58	42	16	Account ownership is only 46.1% of adults, well below the income-predicted level
Trinidad and Tobago	50	35	15	Domestic credit only 44.7% of GDP despite comparatively high income
Bulgaria	46	32	14	Domestic credit only 45.2% of GDP despite comparatively high income
Costa Rica	45	33	12	Account ownership is only 71.4% of adults despite comparatively high income

The countries in Table 6 similarly show economically explicable divergence, driven either by comparatively shallow domestic credit markets (Trinidad and Tobago, Bulgaria) or comparatively low account ownership (Albania, Costa Rica) relative to their income levels. Combining Tables 5 and 6 gives



direct country-specific evidence for the validity of H2: the composite index reflects real and meaningful country-specific variation in financial-system characteristics which is not explained by income alone but is rather the idiosyncratic noise. A symmetry that it is necessary to make explicit here is that the same four-indicator structure that enables some countries (Table 5) to be ranked above their income ranking because they are relatively strong in one specific dimension means they also can be ranked below their income ranking because they are relatively weak in a specific dimension (Table 6). This symmetry is a direct mechanical result of the compensatory aggregation property of PCA, as previously mentioned (Munda & Nardo, 2009): a country's overall score on the composite indicator is a weighted average of the four indicators; hence, if it is very strong (or very weak) on one indicator, relative to the other three, then it must be very weak (or very strong) on the other indicator.

Overall Correlation Robustness

The high overall Spearman correlation (0.86) between the composite index and GDP per capita is strong but leaves about 26% of the rank variance ($1 - \text{square of the correlation}$) unexplained by income alone, suggesting that the composite index is not a redundant statement of a single indicator, and is providing real additional information. Two hypothetical benchmarks for this correlation could be useful to contrast the figures against: the case of a correlation of 1.0 would imply that the composite index is nothing more than a ranking of income with no additional information; a correlation near 0 would raise serious doubts that the index measures anything economically meaningful, since income has been well documented in empirical studies as being associated with financial development. The correlation of 0.86 is within a meaningful range between these two benchmarks, but not quite close to 1.0, which is sufficient to produce the substantively interpretable divergence cases documented in Tables 5 and 6.

Weighting Robustness: PCA-Derived Versus Equal Weighting

This revision specifically tests the dependence of the findings of this paper on the particular PCA-based weighting scheme as requested by Reviewer 1. This is an alternative equal-weighting index that is simply the average of the four standardized indicators, without PCA weighting. This equal-weighting index is strongly correlated with the PCA-derived index, with a Spearman correlation of 0.9992; the maximum rank difference between any two of the 59 countries is only 2, and the mean absolute rank difference is less than half a position. This near-perfect agreement directly addresses the weight-arbitrariness issue for this particular application: For this paper, the loadings of PC1 (0.442 to 0.538) are very close together, and using PCA-derived weights versus using simple equal weights would not make the resulting country rankings substantially different (from a practical point of view). For this paper's four indicators specifically, while conceptually valid, the weighting-arbitrariness critique has little practical impact, but this finding should not be over-generalized, as in a different dataset with more unequal correlational structure among the indicators, the difference between PCA-derived and equal weights could be much larger.

Correlation Against Each Individual Indicator

Table 7: The Composite Index Correlation Against Each Individual Indicator is a Composite Index that Correlates each of the Individual Indicators.

Individual Indicator	Spearman ρ with Composite Index
Account ownership	0.868
GDP per capita	0.865
Internet penetration	0.784
Domestic credit	0.719

Table 7 has the Composite index correlation against each Individual indicator that correlates each of the individual indicators. This revision also provides the correlation results of the composite index with each



of its four components as requested by Reviewer 1 and Reviewer 2, in addition to those with GDP per capita. The 0.865 correlation between GDP per capita and the composite index is not statistically stronger than that of account ownership (0.868), as these two variables represent the greatest portion of the overall variation found along the composite index. The relationship between domestic credit (0.719) and internet penetration (0.784) is less tight, and these two indicators add more to the composite index's divergence from the income or account ownership alone predictability. This directly accounts for the divergence cases in Tables 5 and 6, specifically because of the domestic credit cases (Thailand, China, Malaysia and Trinidad and Tobago), and because of the account ownership case (Albania and Costa Rica)—since these are precisely the two cases for which the correlation of the indicator with the overall composite index is the lowest, and so they contain the most independent and non-redundant information relative to the predominant pattern of the composite index.

Computing the Measure of Compensatory Aggregation: A Controlled Example

This revision contains a concrete, quantified example of the compensatory-aggregation property of PCA, which was discussed conceptually in Section 2.4 (Munda & Nardo, 2009). For this sample, with all four indicator profiles at the median, consider a hypothetical Country B with these four indicators at the same values as the median of this sample (89.3% for account ownership, 61.3% for domestic credit as a share of GDP, \$26,948 for GDP per capita, and 89.8% for internet penetration): the Country B composite indicator score puts it at around the 49th percentile, as expected for a 'pure' median profile. Next, imagine a Country A which is the same as Country B, but domestic credit is raised to 248% of GDP, which is quite extreme but observed in this sample (Hong Kong). Country A's values are also the same for the other three indicators, just the 50th percentile, but with one indicator going to the extreme value, the composite index score goes up to about the 88th percentile, a 39-percentile-point gain. In this controlled example, a relatively high score on one dimension was able to more than compensate for average scores on the other three, creating a significant overall shift in the composite ranking for the country, with no improvement in three dimensions. This is the essence of what Munda and Nardo (2009) criticized in their compensation-aggregation model: a high score on one indicator offset the average score for the remaining three indicators by a considerable margin without any improvement in three scores.

Situating This Finding Within the Composite-Index Literature

The pattern identified in the literature reviewed in Section 2.1 is consistent with this paper's result, namely the finding that there is a strong (but clearly imperfect) correlation between a multi-dimensional composite index and a single-indicator proxy. The Human Development Index (UNDP, 1990), in fact, had the same empirical motivation: the correlation between national income rankings and broader development rankings (including health and education) is high, but there are solid and economically meaningful reasons for the discrepancy, as described by Sen's (1999) theoretical framework of capabilities. A similar pattern is found in a different domain, as Dreher (2006) measures globalization in his KOF Globalization Index; there is a positive correlation between the results from the globalization and the income rankings, but not an identity. It is narrower than these precedents: the specific contribution made by this paper is a financial-development index made up of four indicators, rather than a broader development or globalization measure; but methodologically parallel: the value of a composite index is not in an exact reproduction of a simpler proxy, since that would make the added complexity purposefully superfluous, nor in the absence of any correlation at all, since that would undermine confidence that it represented anything coherent, but in strong correlation in most cases, and in specific cases that have an explicable relationship to the proxy that are distinct and provide the key to additional information.

CONCLUSION AND POLICY IMPLICATIONS

In this paper, the authors created a composite digital financial development index from the four real indicators, provided meaningfulness tests based on four standard variance-explained and loading-balance



tests and compared the ranking obtained with the ranking of a simple alternative based on GDP per capita only. The first principal component has a good eigenvalue (eigenvalue 63.8%) and a well-balanced loading with good empirical support and was further supported by both the Kaiser (1960) eigenvalue criterion and the Cattell (1966) scree-test pattern, allowing it to be used as a single composite measure. The composite index is strongly but not perfectly correlated with income (Spearman $\rho = 0.86$), and the specific cases of divergence in the rank order of the economies discussed in Section 5 'credit' economies (Thailand, China, Malaysia) ranked higher than their income, and account-ownership-shallow' economies (Albania, Costa Rica) ranked lower than their income provide tangible examples of what a multi-dimensional index measures that an economy cannot capture by a single indicator. The methodology presented in this paper can be used as a model for the validation of composite financial-development indexes: first, confirm the index's statistical justifications in multiple ways; second, discuss, and report explicitly, the loadings on the index to address the critique of the "weighting arbitrariness" problem; and third, explore the cases of divergences from simpler alternatives to gain insight into the content of the index and identify what additional complexity is actually capturing, not assuming that there is an end-in-itself to the construction of a financial development index. The three-step validation transparency comparison sequence used in this paper is intentionally generalizable to other substantive domains such as economics, public health, education policy, or other fields, with the only differences being the particular indicators and comparison benchmark for each substantive area.

This paper's method is also applicable to composite-index literature in general, ranging from the Human Development Index (UNDP, 1990) to more specific measures of financial development (Sarma, 2008; Svirydzenka, 2016). This paper's design on this point is directly informed by the shared cautions of Booyesen (2002) and Freudenberg (2003) about the dangers of assuming that a composite index is a self-explanatory summary statistic; this paper not only considers the dangers of constructing a composite index without adequately examining them but also focuses on the implications for specific cases when the index does not agree with a simpler alternative, as Munda and Nardo (2009) note: "Building a composite index is not enough; it is important to examine the implications of the index's disagreement with a simpler alternative, namely, real, traceable consequences."

Limitations

The following limitations should be read in conjunction with the critique literature cited in Section 2.4 and the section below, which discusses how this paper attempts to directly tackle the concerns about weighting-arbitrariness. It is important to note that some of the limitations below are due to genuine conflicts of strategy or method in composite-index methodology, rather than being errors in this particular paper.

- What is actually captured by this index for this paper is a cross-sectional design at one moment in time for four variables with slightly different reference years (2023-2025), rather than a 'true' snapshot at the same time, as noted in companion analyses in this project that have been based on data constraints.
- The composite index of this paper only includes four indicators available from verified sources, whereas a more comprehensive financial development index, e.g., the IMF's four-indicator composite index (Svirydzenka, 2016) and Sarma's Index of Financial Inclusion (Sarma, 2008), would include other dimensions such as financial markets depth and access, and institutional measures that are not included in the four-variable index of the present paper.
- While the weights resulting from PCA in this paper are relatively balanced (Table 4), this does not mean that the weights for each indicator are equally balanced since they are an empirical product of this specific data set and the balance is not to be expected in general PCA index construction, as Decancq and Lugo (2013) argue generally for PCA-derived weights.
- Compensatory aggregation: following Munda and Nardo's (2009) critique discussed in Section 2.4, this paper's PCA-based index permits a very high value on one indicator to offset a very low value on



another; whether this compensatory property is appropriate depends on the specific normative question being asked, a question this paper does not resolve and discloses as an open methodological choice rather than a settled one.

- The specific variance-explained share and loadings reported in Table 4 are perhaps marginally different if the universe of countries available on the Internet is not restricted to the Top-100, as this universe of countries overlaps with the other three countries whose data were used; a broader sample, if available, may result in slightly different variance-explained share and loadings, but not in a direction that can be predicted without the additional data.
- No causal or predictive statements: The index in this paper is descriptive and simply summarizes the observed correlational structure among the four indicators, making no statement about the predictive value of the index itself, or the causal determination of the indicators of one another. While this paper does not attempt to answer the question, it is possible to test whether this composite index or its individual components are able to predict subsequent-year financial or economic outcomes, a question that is clearly not answered here due to the single cross-sectional design.

REFERENCES

- Allen, F., Demirgüç-Kunt, A., Klapper, L., & Martinez Peria, M. S. (2016). The foundations of financial inclusion: Understanding ownership and use of formal accounts. *Journal of Financial Intermediation*, 27, 1–30. <https://doi.org/10.1016/j.jfi.2015.12.003>
- Beck, T., Demirgüç-Kunt, A., & Levine, R. (2000). A New Database on Financial Development and Structure. *World Bank Economic Review*, 14(3), 597–605. <https://pure.uvt.nl/ws/portalfiles/portal/1024725/database.PDF>
- Booyesen, F. (2002). An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research*, 59(2), 115–151. <https://doi.org/10.1023/A:1016275505152>
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10
- Čihák, M., Demirgüç-Kunt, A., Feyen, E., & Levine, R. (2012). *Benchmarking Financial Systems Around the World* (Policy Research Working Paper No. 6175). World Bank. <https://doi.org/10.1596/1813-9450-6175>
- Decancq, K., & Lugo, M. A. (2013). Weights in Multidimensional Indices of Wellbeing: An Overview. *Econometric Reviews*, 32(1), 7–34. <https://doi.org/10.1080/07474938.2012.690641>
- Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2018). *The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution*. World Bank Publications. <https://doi.org/10.1596/978-1-4648-1259-0>
- Demirgüç-Kunt, A., & Levine, R. (2001). Bank-Based and Market-Based Financial Systems: Cross-Country Comparisons. In A. Demirgüç-Kunt & R. Levine (Eds.), *Financial Structure and Economic Growth: A Cross-Country Comparison of Banks, Markets, and Development* (pp. 81–140). The MIT Press. <https://doi.org/10.7551/mitpress/3001.003.0006>
- Dreher, A. (2006). Does globalization affect growth? Evidence from a new index of globalization. *Applied Economics*, 38(10), 1091–1110. <https://doi.org/10.1080/00036840500392078>
- Freudenberg, M. (2003). *Composite Indicators of Country Performance: A Critical Assessment* (OECD Science, Technology and Industry Working Papers, No. 2003/16). OECD Publishing. <https://doi.org/10.1787/405566708255>
- Han, R., & Melecky, M. (2013). *Financial Inclusion for Financial Stability: Access to Bank Deposits and the Growth of Deposits in the Global Financial Crisis* (Policy Research Working Paper No. 6577). World Bank. <http://documents.worldbank.org/curated/en/850681468325448388>
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441. <https://doi.org/10.1037/h0071325>
- Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed.). Springer. <https://doi.org/10.1007/b98835>



- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Munda, G., & Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics*, 41(12), 1513–1523. <https://doi.org/10.1080/00036840601019364>
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffmann, A., & Giovannini, E. (2005). *Handbook on Constructing Composite Indicators: Methodology and User Guide* (OECD Statistics Working Papers, No. 2005/3). OECD Publishing. <https://doi.org/10.1787/533411815016>
- Nicoletti, G., Scarpetta, S., & Boylaud, O. (2000). *Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation* (OECD Economics Department Working Papers, No. 226). OECD Publishing. <https://doi.org/10.1787/215182844604>
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559–572. <https://doi.org/10.1080/14786440109462720>
- Ram, R. (1982). Composite indices of physical quality of life, basic needs fulfilment, and income: A ‘principal component’ representation. *Journal of Development Economics*, 11(2), 227–247. [https://doi.org/10.1016/0304-3878\(82\)90005-0](https://doi.org/10.1016/0304-3878(82)90005-0)
- Ravallion, M. (2012). Mashup Indices of Development. *The World Bank Research Observer*, 27(1), 1–32. <https://doi.org/10.1093/wbro/lkr009>
- Rencher, A. C., & Christensen, W. F. (2012). *Methods of Multivariate Analysis* (3rd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118391686>
- Saisana, M., & Tarantola, S. (2002). *State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development*. European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen, Technological and Economic Risk Management. <https://op.europa.eu/s/AlzH>
- Sarma, M. (2008). *Index of Financial Inclusion* (Working Paper No. 215). Indian Council for Research on International Economic Relations. https://www.icrier.org/pdf/Working_Paper_215.pdf
- Sarma, M., & Pais, J. (2011). Financial Inclusion and Development. *Journal of International Development*, 23(5), 613–628. <https://doi.org/10.1002/jid.1698>
- Sen, A. (1999). *Development as Freedom*. Alfred A. Knopf.
- Slottje, D. J. (1991). Measuring the Quality of Life Across Countries. *The Review of Economics and Statistics*, 73(4), 684–693. <https://doi.org/10.2307/2109407>
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101. <https://doi.org/10.2307/1412159>
- Svirydzhenka, K. (2016). *Introducing a New Broad-Based Index of Financial Development* (IMF Working Paper No. 16/5). International Monetary Fund. https://www.imf.org/-/media/websites/imf/imported-full-text-pdf/external/pubs/ft/wp/2016/_wp1605.pdf
- UNDP. (1990). *Human Development Report 1990*. Oxford University Press. <https://hdr.undp.org/system/files/documents/hdr1990encompletenostats.pdf>

