



Economic Diversification and Sustainable Development in Bahrain, 2015–2024: The Moderating Role of Islamic Banking Performance

Amna Mohamed J Alfardan^{a*}, Ashurov Sharofiddin^b

^a *University College of Bahrain, Manama, Kingdom of Bahrain*

^b *IIUM Institute of Islamic Banking and Finance, International Islamic University Malaysia, Kuala Lumpur, Malaysia*

*Corresponding author: amnaalfardan3@gmail.com

Abstract

The objective of this study is to evaluate the impact of economic diversification on sustainable development in Bahrain over the 2015–2024 period and to examine the moderating role of Islamic banking performance in this relationship. Ten annual observations from official sources were disaggregated into quarterly series ($N = 40$) to allow for dynamic specification. Recognising that interpolation inherently adds no new information, all results are reported at both annual and quarterly frequencies. Furthermore, the findings are subjected to a rigorous eight-part robustness protocol comprising effective-sample inference, a year-block bootstrap, annual estimation, four disaggregation methods, structural break testing, twelve lag orders, four index constructions, and fully modified and dynamic ordinary least squares (OLS) estimations. The results indicate that economic diversification has no substantial direct influence on sustainable development across any dimension. This finding holds consistently, reflecting a persistent disparity between productive diversification, with non-oil output accounting for 82.5% of GDP, and fiscal diversification, with non-oil revenue accounting for only 28.0%. Conversely, Islamic banking performance is positively associated with sustainable development at both data frequencies and under all index constructions, although the specific dimension driving this association remains undetermined. While the moderation coefficient is negative in the reported baseline specification ($\delta = -0.044$, $p = 0.006$), it is highly unstable across robustness checks, being significantly negative in six, significantly positive in eight, and insignificant in fifteen of twenty-nine specifications. Notably, all negative estimates rely on interpolation smoothing and a lag order rejected by both information criteria. This study introduces the first three-variable moderated model applied to Bahrain's macroeconomic context. Its primary contribution is methodological: it demonstrates that the observed negative moderation effect is largely an artefact of specific modelling choices, while better-specified estimators suggest an amplification effect rather than a buffering effect.

Keywords: Economic diversification, Sustainable development, Islamic banking performance, Temporal disaggregation, Small-sample inference

© IIUM Press

1. Introduction

Rentier economies rely on a single natural resource for a substantial share of their public revenues, rendering them vulnerable to price shocks, fiscal volatility, and limited long-term growth. Bahrain is a notable outlier within the Gulf Cooperation Council (GCC) region, characterised by limited reserves, an early peak in onshore production, and government finances that remain exposed to oil price fluctuations despite years of diversification policies (Nakibullah, 2018; Espinoza & Senhadji, 2011). Although oil and gas continue to constitute a substantial share of government revenue, non-oil sectors now account for approximately four-fifths of real GDP. This persistent divergence motivates the central question of this study: has the diversification of production contributed to sustainable development, and if not, under what conditions might it do so? Theoretical frameworks provide grounds for questioning any automatic relationship between diversification and development. Structural change theory posits that diversification promotes development only when

© IIUM Press

Article history

Received: 12 July 2026

Accepted: 24 August 2026

resources are reallocated towards more productive activities with strong export potential (Hvidt, 2013). Furthermore, the resource curse literature suggests that even as the oil sector's share of output declines, hydrocarbons can continue to dominate the macroeconomic cycle through fiscal channels (Gylfason et al., 1999). Financial intermediation theory offers an additional link by emphasising that banks influence not only the pace of economic growth but also its quality, as they determine which activities receive long-term financing (Levine, 2005). This perspective is particularly pertinent to Islamic banks, whose financing modalities—such as musharakah and mudarabah—tie credit to real assets and involve risk sharing with the financed enterprise (Chapra et al., 2008). Additionally, their broader social-finance instruments may promote financial inclusion among groups facing barriers in conventional credit markets.

Bahrain hosts one of the most established Islamic banking sectors in the region, with fifteen banks operating under a dedicated Central Bank of Bahrain regulatory framework. Previous empirical work has generally examined the relevant relationships in isolation. For instance, studies on diversification and development, such as Al-Doski and Al-Arkoub (2023), largely omit the financial sector, whereas research focusing on Islamic banking and development (Alghafes et al., 2024; Asaad & Al-Saad, 2025) overlooks the banking sector's role in the broader economic diversification process. To the best of the authors' knowledge, no prior study has estimated a three-variable model in which Islamic banking performance moderates the effect of economic diversification on sustainable development in Bahrain.

A critical data limitation inherent to this study must be acknowledged: the empirical analysis is based on only ten annual observations, which necessarily limits the statistical power of the empirical analysis and the strength of the conclusions that can be drawn. The quarterly series utilised for dynamic estimation are temporally disaggregated from these ten data points and, consequently, introduce no new underlying information. To address this limitation, all results are reported at both annual and quarterly frequencies and are subjected to an extensive robustness protocol. This protocol is explicitly designed to distinguish findings supported by the underlying data from those that may be sensitive to modelling choices. Therefore, the empirical results should be interpreted as exploratory, and caution is warranted against drawing overly broad generalisations.

Three principal findings emerge. First, economic diversification exerts no significant direct effect on sustainable development across any of the dimensions examined. Second, Islamic banking performance is positively associated with sustainable development at both data frequencies, although the exact transmission channel cannot be definitively established. Third, the interaction effect is highly sensitive to model specification: it is negative under the reported baseline specification but becomes positive under the lag orders preferred by information criteria and under estimators designed to account for cointegrating relationships. Taken together, these findings suggest that the relationship between diversification and sustainable development in Bahrain depends not only on the extent of diversification itself but also on the structural and financial conditions under which it is funded. This interpretation is offered as a proposition for future investigation rather than as a definitive causal mechanism established by the present analysis.

2. Literature Review

2.1 Diversification and sustainable development

Economic diversification involves widening the productive base and sources of income so that output and government revenue are not dependent on a single sector. It encompasses productive, export, and revenue dimensions, and this distinction is important: an economy may diversify its production while its budget remains reliant on resource revenues. This is reflected in the evidence from the GCC. Espinoza and Senhadji (2011) demonstrate that non-oil growth in the Gulf remains linked to oil prices through public spending. Nakibullah (2018) notes that Bahrain non-oil growth is accompanied by fiscal dependence on oil. Al-Doski and Al-Arkoub (2023) estimate an inverse relationship between oil reliance and fiscal sustainability across the GCC over the period 2004–2020. Our measure of diversification is a composite of three indicators: the share of non-oil output, the share of non-oil revenue, and the share of manufacturing. It is therefore a cruder measure than the concentration measures commonly employed in the literature. Theil and Herfindahl-Hirschman indexes of export concentration, as well as the Economic Complexity Index, capture compositional diversification in ways that a share-based composite cannot. The present metric was used because its components are available annually for Bahrain throughout the entire study period, unlike disaggregated export and value-added statistics. It represents a measurement compromise and limits comparability with cross-country studies.

The conventional operationalisation of sustainable development, namely, the ability to meet present needs without compromising the ability of future generations to meet their own needs (World Commission on Environment and Development, 1987), is categorised across economic, social, and environmental dimensions.

Cross-country research indicates that diversified economies outperform resource-concentrated ones, although the literature is divided on the underlying mechanisms. For Hvidt (2013), the binding condition is compositional: diversification matters only when it involves meaningful structural change towards higher-productivity tradable activities. Hence, an increase in the non-oil share resulting from non-tradable expansion should have no effect. For Gylfason et al. (1999), the issue is allocative: resource economies channel funding towards short-term activities irrespective of sectoral composition, so the failure operates through the financial system. These two diagnoses may appear similar in aggregate statistics but imply opposite remedies—industrial policy versus financial reform—and neither the GCC nor the Bahraini literature has clearly distinguished between them. A model that incorporates the financial sector as a conditioning variable can help distinguish a compositional failure from an allocative failure, whereas dominant bivariate designs cannot. This study addresses this gap by incorporating the financial sector into the analysis.

H1. Economic diversification has a statistically significant effect on sustainable development in Bahrain, both in composite form (H1) and across its economic (H1a), social (H1b), and environmental (H1c) dimensions.

2.2 Islamic Banking Performance and Sustainable Development

Conventional profitability ratios are not sufficient to measure the performance of Islamic banks. The *maqasid al-Shariah* paradigm asserts that an Islamic bank is an institution that exists to promote public benefit and prevent harm by protecting wealth through productive rather than speculative finance, assisting those in need to safeguard society, and preserving the environment through responsible stewardship (Chapra et al., 2008; Dusuki & Bouheraoua, 2011). This perspective has led to composite metrics that incorporate both financial ratios and sustainability indicators. However, critical literature cautions that some banks invoke the concept of sustainability without genuinely incorporating it into their financing policies.

Recent empirical evidence reflects these varying degrees of integration. Alghafes et al. (2024) find that the implementation of environmental, social, and governance (ESG) practices improves the financial performance of GCC Islamic banks, adding that the environmental dimension requires deliberate strategic action, whereas the social and governance dimensions are more naturally aligned with Islamic finance principles. Asaad and Al-Saad (2025) find that *zakat*, *waqf*, *qard hasan*, and *musharakah* instruments are utilised to reduce poverty and promote equity. Otman (2022) demonstrates that banks with a developmental orientation contribute more effectively to macroeconomic goals, whereas banks focusing on short-term products have a weaker impact. Avdukic and Asutay (2025) identify a continuing discrepancy between declared moral-economy orientations and measurable development impacts. In the local context, Abdulla et al. (2021) note the early institutional adoption of sustainability in Bahrain, although clear performance requirements remain absent.

Islamic banks may have measurable effects when their instruments are directly linked to specific development outcomes, but their impact may be weaker when the relationship operates solely through aggregate productive investment. A common limitation of these studies is that they treat bank performance strictly as an independent driver of development. None examines whether banking performance alters the developmental returns of structural reform—that is, whether it operates through the diversification–development relationship rather than merely alongside it. Establishing that banking performance contributes to development does not, by itself, indicate whether it makes economic diversification more or less developmentally productive. This is the second critical question that must be addressed in a rentier economy pursuing a diversification strategy.

H2: The performance of Islamic banks has a statistically significant effect on sustainable development in Bahrain, both in its composite form and across its economic (H2a), social (H2b), and environmental (H2c) dimensions.

2.3 The Role of Islamic Banking Performance as a Moderator

It is not only the direction of the moderation effect that remains uncertain. The two theoretical frameworks predict conflicting signs, which makes the empirical test informative. Financial intermediation theory

suggests an amplification effect. If banks constitute the channel through which structural change is transformed into productive investment, then a more developed banking sector would increase the developmental return at any given level of diversification, and the interaction effect should therefore be positive (Levine, 2005). The *maqasid* framework, by contrast, implies a substitution effect. A bank that aims to protect society and the environment can generate welfare through *zakat*, *qard hasan*, and environmental financing, none of which depends on structural transformation. As banking performance increases, the economy may rely less on what the diversification process provides, and the interaction effect should therefore be negative (Dusuki & Bouheraoua, 2011). A functioning banking system may operate through both mechanisms simultaneously. The estimated sign is therefore an empirical question concerning which mechanism dominates, while a null effect would indicate that the two mechanisms offset each other.

H3: The performance of Islamic banks significantly moderates the relationship between economic diversification and sustainable development in Bahrain.

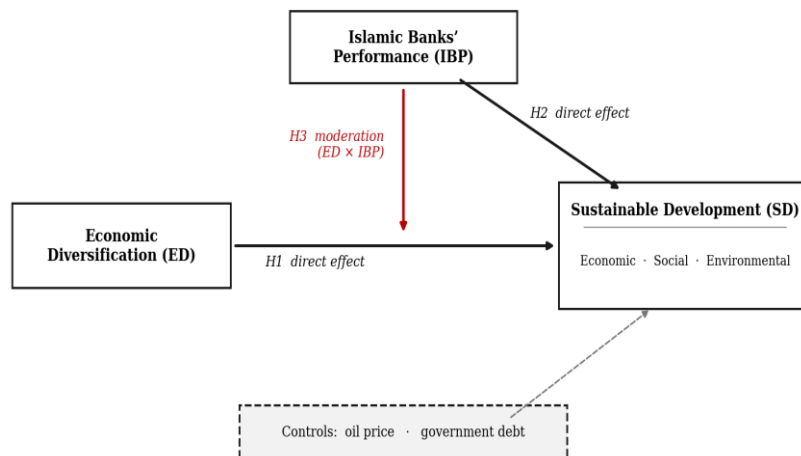


Fig. 1. Conceptual model

3. Research Methodology

The unit of analysis is the Bahraini economy during the period 2015–2024. This decade encompasses the 2015–2016 oil price collapse, the COVID-19 shock, the subsequent recovery, and the acceleration of Economic Vision 2030. Annual data ($N = 10$) were gathered from the Ministry of Finance and National Economy, Central Bank of Bahrain Financial Stability Reports, the Information and eGovernment Authority, the UNDP, the World Bank World Development Indicators, and the IMF World Economic Outlook. Banking metrics were aggregated at the sector level using annual reports of licensed Islamic banks.

The sample has ten independent observations, and no procedure applied here alters that. The annual data were disaggregated into quarterly series ($N = 40$) following Chow and Lin (1971) and the Denton–Cholette method, subject to the restriction that the four quarters in each year reproduce the corresponding annual value (Sax & Steiner, 2013). This was done for one reason: specifications with lag structures in a dynamic context cannot be estimated using only ten observations. Disaggregation produces an intra-year path consistent with annual totals; it does not generate additional information, and the effective information content of the quarterly dataset remains the ten annual values. Two implications follow and are treated as first-order issues. First, the interpolation structure induces serial correlation; therefore, the Newey–West adjustment is used to reduce the potential anti-conservativeness of the standard errors for the $N = 39$ observations. Inference based solely on the quarterly frequency may reflect imposed smoothness rather than underlying economic activity. Second, an interpretive problem arises because the Chow–Lin estimator requires a quarterly indicator series associated with the target variable, and no such series was available for any variable. When no indicator is used, the estimator effectively smooths the series subject to the annual constraints and is not substantially different from the Denton–Cholette method. This explains why, as shown in Section 4.5, the two methods produce similar results while differing from methods that impose less smoothing.

There are three composite indices. The components are standardised to have a mean of zero and a variance of one across the ten annual observations, and each composite is calculated as an equally weighted average of

its standardised components. Equal weighting is intentional: without a theoretically motivated weighting scheme or a sufficiently large sample to estimate weights, alternative weighting would impose a structure that the data cannot reliably sustain. Sensitivity to alternative weighting schemes is discussed in Section 4.5. Economic diversification (ED) is the sum of the non-oil portion of real GDP, the non-oil share of government income, and the contribution of manufacturing to GDP. Sustainable development (SD) is a composite index combining three standardised sub-indices reflecting economic, social, and environmental dimensions. Islamic banking performance (IBP) is a sector-level aggregate of asset-quality and financial-soundness metrics for Islamic banks. Documentation at the component level is provided in the Appendix.

A limitation of the index documentation belongs here, rather than in the limitations section, because it bears on how the results should be interpreted. The component-level series underlying the three composites are no longer available; only the composite indices and a subset of their components could be recovered. The Appendix therefore documents the construction as specified rather than as independently verified. Evidence of reliability can be reported for the sustainable development index, for which the three-dimension series remain available, but not for the diversification or banking-performance indices. Accordingly, all three composites should be regarded as undocumented at the component level. Each model includes two control variables: international oil price (OIL), reflecting the direct effect of the oil cycle, and government debt-to-GDP (DEBT), reflecting the debt-financed component of non-oil growth. DEBT has a substantial association with diversification ($r = 0.878$) and also contributes to the collinearity reported in Section 4.3.

3.1 Specification and estimation

There are three hypotheses and three equations:

$$SD_t = \alpha_0 + \alpha_1 ED_t + \alpha_2 SD_{t-1} + \alpha_3 OIL_t + \alpha_4 DEBT_t + \varepsilon_t \quad (1)$$

$$SD_t = \gamma_0 + \gamma_1 ED_c + \gamma_2 IBP_c + \gamma_3 SD_{t-1} + \gamma_4 OIL_t + \gamma_5 DEBT_t + v_t \quad (2)$$

$$SD_t = \delta_0 + \delta_1 ED_c + \delta_2 IBP_c + \delta_3 (ED_c \times IBP_c) + \delta_4 SD_{t-1} + \delta_5 OIL_t + \delta_6 DEBT_t + v_t \quad (3)$$

To reduce collinearity between the interaction term and its components, ED and IBP were mean-centred before multiplication. The test of H3 is performed using the coefficient δ_3 . Equations 1 and 2 are re-estimated with each dimension of SD as the dependent variable. The lagged dependent variable accounts for the usual persistence of development indicators and results in the loss of one observation; therefore, quarterly regressions are estimated using $N = 39$ and annual regressions using $N = 9$. The specification is a dynamic regression in levels rather than an error correction model. Equations 1 to 3 do not include an error correction term and do not generate an adjustment speed; therefore, the framework should not be categorised as ARDL/ECM. However, the error correction representation is estimated separately in Section 4.5 because the result is directly relevant to the bounds tests.

Integration order was tested using the Augmented Dickey–Fuller and Phillips–Perron tests. Bounds testing was performed using the method proposed by [Pesaran, Shin, and Smith \(2001\)](#). The diagnostic tests indicated autocorrelation and heteroskedasticity; therefore, all standard errors are estimated using Newey–West heteroskedasticity and autocorrelation consistent (HAC) standard errors with four lags. Since robust standard errors are not sufficient to address the inferential problem arising from disaggregation, an eight-part robustness protocol is adopted: effective-sample-size inference based on ten rather than thirty-nine observations; a year-block bootstrap resampling whole years (5,000 replications); direct annual estimation; four alternative disaggregation methods; Quandt–Andrews structural break testing with COVID variants ([Andrews, 1993](#)); twelve ARDL(p, q) lag orders selected using AIC and SIC; four alternative index constructions; and fully modified OLS ([Phillips & Hansen, 1990](#)) with dynamic OLS. A result is deemed supported only if it passes this protocol.

4. Results

4.1 Descriptive statistics and correlation

Table 1 presents descriptive statistics at the annual frequency, reflecting the level at which the data were originally observed. Two features emerge prior to the regression analysis. First, the average non-oil share of GDP (82.46%) is substantially higher than the average non-oil share of government revenue (27.99%),

indicating that production has diversified much more than government revenue. Second, a high capital adequacy ratio (20.01%) coexists with a modest return on assets (0.47%), suggesting that some of the sector's capital strength may not be translating into high-return productive investment.

Table 1. Descriptive statistics, 2015–2024 (N = 10)

Variable	Mean	SD	Min	Max
Economic diversification (composite)	0.000	0.702	−0.901	1.255
Share of GDP from non-oil sector (%)	82.46	2.11	79.50	86.00
Share of non-oil revenue (%)	27.99	6.87	18.50	38.50
Manufacturing share of GDP (%)	14.47	0.38	13.60	14.90
Composite sustainable development	0.000	0.484	−0.472	0.914
Performance (composite) of Islamic banks	0.000	0.969	−0.821	1.932
Return on assets (%)	0.470	0.330	0.100	1.100
Capital adequacy ratio (%)	20.01	3.70	16.80	27.70
Human Development Index	0.881	0.014	0.859	0.899
Oil price (USD/barrel)	66.07	18.45	41.80	99.80
Government debt to GDP (%)	106.4	23.08	66.70	134.0

Note: Composite indices are standardised. Descriptive statistics are not provided for the interpolated quarterly series since doing so would misrepresent the sample size.

The correlation analysis identifies three associations of interest. Islamic banking performance is strongly correlated with sustainable development ($r = 0.911$), economic diversification is correlated with sustainable development ($r = 0.703$), and economic diversification is strongly correlated with government debt ($r = 0.878$). This final association is of particular substantive relevance, as it suggests that a portion of Bahrain's non-oil growth may be linked to government debt rather than private-sector-led expansion. Furthermore, these high pairwise correlations caution against interpreting the simple associations causally, as diversification and sustainable development may be jointly influenced by common macroeconomic factors, such as oil prices and government spending.

4.2 Integration, Cointegration, and Causality

All series in the quarterly data are non-stationary in levels and stationery in first differences, except for oil, which is marginally stationary in levels. The results of these tests must be interpreted with caution: mechanical smoothing of the data increases persistence and biases unit root tests towards non-rejection. Consequently, the uniform $I(1)$ classification may be partly an artefact of the interpolation rather than an underlying structural property of the Bahraini economy, while the underlying annual sample of $N = 10$ is too small to support meaningful unit root testing.

The bounds tests help distinguish between the two candidate relationships. The model relating development to banking performance alone produces an F -statistic of 9.227, which exceeds the upper bound at the 1% level. The joint model produces an F -statistic of 5.922, which exceeds the upper bound at the 5% level. In contrast, the model linking development to diversification alone produces an F -statistic of 0.006, which is well below the lower bound. These tests may be oversized when applied to interpolated data, and the effective sample size is well below the range for which the tabulated critical values apply. Therefore, the results are presented as indicative of relative strength rather than as definitive evidence of cointegration. The Granger causality tests reveal no evidence that diversification precedes development at any of the lags tested, whereas banking performance precedes development at two lags ($F = 11.198$, $p < .001$).

4.3 Regression results

The three primary specifications are reported in Table 2. The diversification coefficient in Model 1 is positive but far from significant ($\beta = 0.010$, $p = .625$). Instead, persistence dominates, as evidenced by a lagged development coefficient of 1.026.

This coefficient warrants careful attention. If the coefficient is equal to or greater than unity, the process is non-mean-reverting, in which case a levels specification is inappropriate and the remaining coefficients are estimated in a near-unit-root regression with non-standard inference. Formally, we cannot reject the null

hypothesis of $\phi = 1$ for Model 1 ($t = 0.617, p = .541$) or Model 3 ($t = 0.116, p = .909$). The development composite is therefore consistent with a random-walk process over this window. This means that the level relationships suggested by the bounds tests rest on a dependent variable whose own dynamics cannot be distinguished from a unit root, and the standard errors on the remaining regressors may not be reliable in the conventional sense. A first-difference specification would avoid this issue but would discard the level relationships that fundamentally motivate this investigation. Given the limitation of ten annual observations, this choice cannot be resolved empirically. In Model 2, banking performance enters positively and significantly ($\beta = 0.066, p = .041$). Model 3 incorporates the interaction term, which is negative and significant ($\delta_3 = -0.044, p = .006$), while economic diversification becomes positive and significant for the first time. At face value, this represents the buffering pattern hypothesized in this study. However, this interpretation does not survive the robustness analysis presented in Section 4.5.

All three models are subject to concerns regarding multicollinearity, as indicated by the variance inflation factors (VIF). In Model 3, all variables identified as significant possess VIFs exceeding the commonly used threshold of 10: diversification (24.8), banking performance (11.9), the interaction term (10.2), the lagged dependent variable (29.9), and debt (47.8). This is consistent with the strong correlation observed between economic diversification and government debt ($r = .878$). This severe collinearity likely explains a pattern that would otherwise be difficult to interpret: the debt coefficient is positive and negligible in Model 2 but negative and substantial in Model 3. Sign reversal across nested specifications is a classic indication of collinearity rather than proof of a substantive economic relationship; therefore, the individual coefficients in Model 3 must be interpreted as imprecise, even when their p -values appear favourable.

Table 2. Regression findings, composite sustainable development (quarterly, N = 39)

Variable	Model 1	Model 2	Model 3
Constant	0.038*** (0.008)	0.034*** (0.008)	0.057*** (0.012)
ED (centred in Models 2, 3)	0.010 (0.020)	0.011 (0.018)	0.087*** (0.030)
IBP (centred)	—	0.066** (0.032)	0.065** (0.033)
ED × IBP	—	—	−0.044*** (0.016)
SD (t−1)	1.026*** (0.042)	0.904*** (0.087)	1.011*** (0.096)
OIL	−0.013 (0.010)	−0.026* (0.015)	−0.044*** (0.016)
DEBT	0.020 (0.019)	0.024 (0.018)	−0.054** (0.023)
R ²	0.9913	0.9928	0.9934
F	1481.0	1880.5	2190.5

Note: Newey-West HAC standard errors (4 lags) in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. A high R^2 is a consequence of the lagged dependent variable, which on its own explains much of the variance in a very persistent series, and should not be interpreted as explanatory power.

4.4 Results at the dimension level and the form of moderation

The dimension-level estimations are presented in Table 3. Hypotheses H1a through H1c are rejected because economic diversification does not have a significant effect on any individual dimension. Islamic banking performance does not have a significant impact on the economic dimension but has a significant positive impact on the social ($\beta = 0.113, p = .007$) and environmental ($\beta = 0.385, p < .001$) dimensions. The oil coefficients are substantial: positive and significant for the economic dimension ($\beta = 0.399$) and strongly negative for the environmental dimension ($\beta = -0.286$). These results suggest that the economic dimension of sustainable development in Bahrain remains closely linked to the oil cycle, whereas the environmental dimension exhibits an inverse association with oil prices.

Table 3. Dimension level estimates (quarterly, N = 39)

Variable	SD_ECO	SD_SOC	SD_ENV
ED, Equation 1	0.020 (0.867)	−0.018 (0.823)	−0.043 (0.776)
ED, Equation 2 (centred)	0.030 (0.798)	−0.013 (0.798)	0.004 (0.958)
IBP, Equation 2 (centred)	0.114 (0.303)	0.113*** (0.007)	0.385*** (0.000)
OIL	0.399** (0.019)	−0.140** (0.036)	−0.286*** (0.000)

Variable	SD_ECO	SD_SOC	SD_ENV
R ² (Eq. 2)	0.9357	0.8908	0.9572

Note: *p*-values in brackets. Newey-West HAC standard errors (4 lags). Controls and lagged dependent variables included, coefficients not displayed for space. *** $p < 0.01$, ** $p < 0.05$.

The marginal effect of economic diversification decreases monotonically as banking performance increases, falling from 0.128 at the observed minimum to 0.006 at the maximum. Notably, this effect is no longer statistically distinguishable from zero above the IBP_c threshold of 1.18—a level within the observed range that corresponds to the 2023–2024 period of strong banking recovery. Taking the Model 3 estimates at face value, Figure 2 depicts the simple slopes and marginal effects alongside their confidence intervals. However, this interpretation is conditional: the estimates are subject to the fragility of the underlying interaction, which the robustness analysis in Section 4.5 shows to be decisive.

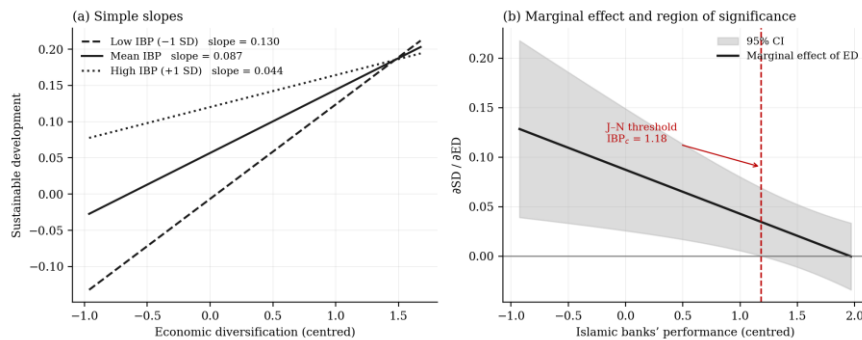


Fig. 2. Simple slopes and the marginal effect of diversification across levels of banking performance

4.5 Robustness

Inference and frequency: When reassessed using degrees of freedom based on the ten annual observations, the interaction retains its sign but becomes statistically insignificant ($p = .071$). Using a year-block bootstrap that resamples complete years, its 95% confidence interval ranges from -0.188 to 0.307 and includes zero. When estimated directly using the annual data, the interaction is negative in all specifications but never significant ($p = .106$ in the complete model and $p = .419$ in the parsimonious model). The annual full specifications also produce lagged dependent-variable coefficients of 2.17 and 2.72, which are not economically interpretable and imply overfitting at $N = 9$. The parsimonious annual model, the only annual specification with sufficient residual degrees of freedom, yields a clear result: economic diversification (0.213, $p = .009$) and banking performance (0.386, $p < .001$) are both positive and significant, while the interaction term is not included.

Dimensional attribution: The dimensional attribution does not survive the annual check. The quarterly estimates attribute the banking influence to the social and environmental dimensions and find no significant effect on the economic dimension. The annual estimates show the opposite pattern: substantial in the economic dimension (0.723, $p = .010$), absent in the social dimension (0.072, $p = .679$), and marginal in the environmental dimension (0.443, $p = .068$). What remains stable across frequencies is that banking performance is favourably associated with development in composite form. This study does not, and cannot, resolve which dimension conveys this association. The reliability evidence below supports this conclusion.

Parameter stability: The Quandt–Andrew’s test indicates a sup-Wald statistic of 414.1 in 2020Q3, confirming the instability indicated by the CUSUM and CUSUMSQ violations. The interaction coefficient changes from -0.190 ($p < .001$) before 2020 to 0.001 ($p = .946$) after 2020, while the banking performance coefficient reverses from -0.802 to $+0.214$. The parameters therefore do not merely change in magnitude; their signs change across regimes. Averaging full-sample estimates across two regimes with opposite signs would therefore obscure the distinct relationships across the regimes.

Order of lag: According to the Akaike and Schwarz criteria, the reported ARDL (1,0) is the weakest of the twelve ARDL (p, q) specifications. Both criteria select ARDL (2,2), with an interaction coefficient of $+0.251$

($p = .004$), which is positive, substantial, and opposite in sign to the reported estimate. Among the twelve specifications, one produces a substantial negative interaction—the specification reported in the baseline results—four produce significant positive interactions, and seven produce insignificant interactions.

The conditional error correction model: The cointegration issue is further examined by directly estimating the error correction representation. The adjustment coefficient is -0.020 ($t = -0.25$), well above the lower bound of -2.86 for $k = 3$ reported by Pesaran et al. (2001), and is statistically indistinguishable from zero ($p = .803$). There is therefore no evidence of error correction; that is, there is no detectable tendency for the system to return to equilibrium following a shock. The implied half-life of approximately 35 quarters is an artefact of dividing by an adjustment coefficient that is statistically indistinguishable from zero. The bounds F-statistic for the lagged levels (5.275) exceeds its upper bound, whereas the t-statistic for the adjustment coefficient is well above its own lower bound. This represents the contradictory-signal case discussed by Pesaran et al. (2001), in which the appropriate conclusion is that the evidence for cointegration is inconclusive. Hence, the framework is referred to as a dynamic regression in levels, and no cointegration is asserted.

Reliability of the index: Cronbach's alpha for the three development dimensions is 0.537 annually and 0.475 quarterly, both below the customary 0.70 criterion. The source of this low internal consistency is the economic dimension, which has a correlation of -0.019 with the social dimension and 0.224 with the environmental dimension. Its item-rest correlation is 0.148 , and deleting this dimension increases alpha to 0.808 . Principal component analysis (PCA) is consistent with this finding: the first component explains 62.5% of the variation, but the second eigenvalue is 1.009 , above the Kaiser threshold, indicating that the index is two-dimensional. PC1 loads 0.94 on the social dimension and 0.97 on the environmental dimension but only 0.23 on the economic dimension. This provides a meaningful interpretation: the economic dimension moves with the oil cycle, whereas the other two do not, suggesting that the composite averages series driven by different underlying forces. The component series for the diversification and banking performance indices are not available; therefore, similar statistics could not be calculated for them. Alternative constructions and alternative estimators the interaction remains significant in one of the four alternative constructions of the development index: principal-component weights, min-max normalisation, rank transformation, and an alpha-optimal two-dimension version. Conversely, the banking performance coefficient is positive and significant across all five constructions and is highest where the index is most internally consistent: 0.324 in the alpha-optimal version compared with 0.065 with equal weights. Aggregation choices that improve measurement therefore strengthen support for H2 and reduce support for H3. The estimators designed for cointegrating regressions—fully modified OLS and dynamic OLS—produce positive and significant interaction coefficients in all three specifications: $+0.145$ ($p < .001$), $+0.124$ ($p = .019$), and $+0.263$ ($p = .003$).

A census of the interaction estimates: The robustness exercises generate a wide range of specifications, and it is more informative to count them than to characterise them vaguely. All estimates are summarised in Table 4. Of the twenty-nine specifications, six produce a significant negative coefficient, eight produce a significant positive coefficient, and fifteen produce a coefficient statistically indistinguishable from zero. The count alone provides limited evidence because the specifications are not independent draws. More important is the structure underlying these results. All negative estimates are based on Chow–Lin or Denton–Cholette-smoothed data and are estimated using the restrictive ARDL (1,0) lag order. The COVID-19 and subsample variations are not independent confirmations; rather, they re-estimate the same specification on subsets of the same smoothed series. All positive estimates result from relaxing one of these two choices: using disaggregation without intra-year smoothing, adopting the richer lag structures selected by both information criteria, or using cointegrating estimators. The sign therefore appears to reflect two modelling decisions rather than a stable feature of the Bahraini economy. Neither decision has a strong empirical justification: the smoothing does not add information, and the selected lag order is the one explicitly rejected by the information criteria.

Table 4. Census of estimates of the interaction coefficient (δ_3)

Group	Specifications	Range of δ_3	Common feature
Significant negative	6	−0.190 to −0.025	Chow-Lin or Denton smoothing with ARDL(1,0)
Significant positive	8	+0.111 to +0.296	Smoothing eliminated, lag order relaxed or cointegrating estimator
Not significant	15	−0.106 to +0.094	Annual data, corrected inference, weaker smoothing, index rebuilt

Note: The negative group includes the reported specification, Denton-Cholette disaggregation, the COVID dummy, 2020 omitted, the pre-2020 subsample and min-max normalisation. The positive group includes step disaggregation, ARDL(2,1), (2,2), (3,2) and (4,2), FMOLS and DOLS with one and two leads and lags. Significance at the 5 per cent level.

Table 5. Summary of hypothesis tests

Hypothesis	Coefficient	p	Survives protocol	Decision
H1: ED → SD (including H1a–H1c)	0.010	0.625	n/a (null)	Not supported
H2: IBP → SD	0.066	0.041	Yes	Supported
H2a: IBP → SD_ECO	0.114	0.303	n/a (null)	Not supported
H2b: IBP → SD_SOC	0.113	0.007	Quarterly only	Supported, channel unknown
H2c: IBP → SD_ENV	0.385	0.000	Partially	Supported, channel unknown
H3: ED × IBP → SD	−0.044	0.006	Sign not identified	Not established

5. Discussion

The rejection of **H1** represents a boundary condition rather than a contradiction of the diversification literature. Structural change theory argues that the gains from development depend on the reallocation of resources towards higher-productivity, export-oriented activities. The Bahrain data demonstrate quantitative sectoral expansion without a corresponding fiscal transformation. Non-oil output averages 82% of GDP, whereas non-oil revenue averages only 28%. The diversification index is also strongly correlated with public debt ($r = .878$). Growth in construction and services supported by public borrowing increases the measured non-oil base but does not necessarily decouple the development path from the oil cycle. The positive oil coefficient in the economic-dimension model ($\beta = 0.399$) indicates that the economic dimension of development responds more strongly to oil prices than to measured diversification, consistent with Espinoza and Senhadji (2011) and Al-Doski and Al-Arkoub (2023). This null result is robust across numerous specifications, frequencies, and disaggregation methods and is therefore the least sensitive to the data restrictions of the study.

The **H2** pattern identifies an association between Islamic banking performance and sustainable development, with an important proviso. At the quarterly frequency, the social effect may operate through channels that are less prominent in conventional banking, including *zakat*, *qard hasan*, microfinance, and housing finance for lower-income households (Asaad & Al-Saad, 2025). The environmental effect is the largest coefficient in the study, consistent with the Central Bank of Bahrain's principles of sustainable finance in 2022 and 2023, and with the findings of Alghafes et al. (2024), who argue that the environmental dimension of ESG requires deliberate strategic effort in GCC Islamic banks. The environmental model also shows a strongly negative oil coefficient ($\beta = -0.286$), which makes the interpretation more pronounced: oil abundance and environmental sustainability appear to move in opposing directions in Bahrain, while Islamic banking performance is positively associated with the environmental indicator. However, the annual estimates place the banking effect on the economic dimension instead. Thus, this dimensional interpretation is supported by the quarterly series but is not established consistently across frequencies. Both frequencies, however, show a positive association in composite form, which is consistent with Otman (2022) and with Avdukic and Asutay (2025) regarding the gap between stated developmental orientation and measured outcomes.

The moderation hypothesis (**H3**) deserves separate consideration. Section 2.3 presented competing theoretical predictions, and the data do not provide a definitive basis for distinguishing between them. The reported specification favours substitution, and five variants of it do so as well. All six share two modelling choices that are problematic on their own terms: interpolation smoothing that adds no information to the ten annual observations and a lag order that performs worst among the twelve specifications according to both information criteria. Relaxing either choice reverses the sign. The specifications supporting amplification are therefore more strongly justified by the model-selection and estimation evidence, representing an asymmetry in the quality of the supporting specifications rather than simply in their number.

The policy implications differ substantially depending on which interpretation is correct, which is why the ambiguity matters. Under the buffering interpretation, a robust Islamic banking industry is a partial substitute for diversification, with the two representing alternative channels of development. Under the amplification interpretation, they are complements: diversification generates development returns to the extent that banks allocate resources towards it, and neither strategy acts alone. The second interpretation is closer to the structural picture documented here—a diversification index correlated at .878 with public debt, non-oil growth concentrated in sectors dependent on government spending, and a banking sector with a capital adequacy ratio of 20.01% but a return on assets of only 0.47%. This may describe an economy diversifying with substantial reliance on borrowed funds while its banks maintain capital that is not being translated into high returns. It is an economy in which the two channels may not yet be effectively connected, which is what a positive but unrealised interaction would depict. The data suggest this interpretation rather than confirm it.

The ambiguity is further compounded by the structural break identified in 2020Q3. Whether the sign reversal across this point represents a regime change following the pandemic and the subsequent supervisory focus on sustainable finance, or the instability that may arise when only five annual observations are divided into two regimes, cannot be determined from the present data. This may be the most important question raised by the analysis. One further methodological implication is that the diversification coefficient is insignificant in Models 1 and 2 but becomes significant when the interaction is introduced. This suggests that studies estimating only direct effects in similar settings may report coefficients that are affected by omitted interaction effects rather than relationships that are genuinely absent.

6. Conclusion

Measured economic diversification had no significant direct effect on development in any dimension over the 2015–2024 period. This is the most robust result of the study, and it holds across all specifications, at both frequencies, and with all disaggregation methods. The main empirical contribution is that Islamic banking performance was favourably associated with sustainable development in its composite form, and this association remains supported across the robustness analyses at both frequencies, although the quarterly and annual estimates differ regarding which dimension underlies the association. The reported specification showed a negative interaction, but this result is not stable: across twenty-nine specifications, it is significantly negative in six, significantly positive in eight, and insignificant in fifteen, with the pattern being systematic. The paper consequently makes no assertion that moderation exists in either direction. If the better-supported specifications point in either direction, they tend towards amplification, which is the opposite of what the reported specification implies.

First, the findings point to a fiscal and institutional constraint rather than a sectoral one. A more relevant policy objective is to address the disparity between an 82% non-oil share of output and a 28% non-oil share of revenue. The oil cycle continues to influence the economic dimension of development through public finance. This points to the revenue side of Economic Vision 2030—the pace of non-oil revenue mobilisation and the treatment of debt-financed non-oil activity in fiscal reporting—rather than to further quantitative expansion of non-oil output.

Second, the developmental return of Islamic banking can potentially be improved through changes in the composition of finance and by making sustainability obligations measurable. In terms of composition, this would involve a shift from short-term consumer products towards *musharakah* and *mudarabah* structures serving productive non-oil sectors. The tools available to the Central Bank of Bahrain include differential capital or liquidity treatment and disclosure rules that make financing composition more transparent. In terms of measurement, the 2022 and 2023 sustainable finance principles provide useful guidelines. Their transformation into reporting requirements incorporating three distinct indicators—green *sukuk* issuance,

ijarah-based renewable energy financing volumes, and the proportion of finance subject to environmental screening—would produce the observable series that are absent from the present analysis.

Third, *Shariah* governance needs to be considered as a development variable in its own right, as it may determine whether *maqasid* principles are incorporated into financing policy or remain primarily declaratory. Data on board composition, jurisdiction over financing choices, and audit scope at the bank level could be related to the developmental profile of each portfolio, helping to address the aggregation problem encountered in the present study.

This study is subject to several limitations. First, the sample size is constrained by the availability of annual data, with only ten observations ($N = 10$) over the study period. This is insufficient to support the dynamic specifications typically employed in this literature and substantially limits the statistical power of the analysis. Accordingly, the inferential results should be interpreted with an effective sample size of ten. From a specification perspective, the unit-root behaviour of the lagged dependent variable cannot be rejected, individual coefficients in Model 3 are imprecisely estimated because of multicollinearity, and the confirmed structural break indicates that full-sample estimates may mask differences in parameter relationships across regimes. Second, the study faces limitations in the measurement of the composite indices. The component series underlying the indices are not fully available; consequently, the indices cannot be completely replicated from publicly published data. This limitation constrains the interpretation of the sustainable development index, particularly its environmental dimension, which is subject to greater measurement uncertainty. Future confirmatory research should therefore reconstruct the indices from their underlying source components using fully documented and reproducible procedures. Where the necessary source data are publicly available, such reconstruction should be feasible.

These limitations point to three priorities for future research. First, the composite indices should be reconstructed using fully documented component-level data and transparent aggregation procedures. Second, a GCC panel framework could provide the cross-sectional variation absent from the current design and enable a more rigorous assessment of whether the observed amplification pattern generalises across Gulf economies. Third, bank-level panel data could incorporate *Shariah* governance variables directly into the analysis and provide greater insight into how governance structures shape the developmental composition of Islamic banking portfolios. Collectively, these extensions would provide a stronger empirical basis for evaluating the relationships examined in the present study. The current findings therefore provide a basis for further investigation and highlight the importance of improved data collection and measurement for future confirmatory research.

Acknowledgements

The comments received on an earlier draft prompted the robustness analysis reported in Section 4.5, which significantly changed the results of the work.

References

- Abdulla, Y., Ebrahim, R., & Kumaraswamy, S. (2021). Islamic finance and sustainable development goals in Bahrain. In 2021 International Conference on Sustainable Islamic Business and Finance (pp. 8–13). New York: IEEE.
- Al-Doski, K., & Al-Arkoub, H. (2023). The impact of economic diversification on fiscal sustainability in the Arab Gulf Cooperation Council countries for the period 2004-2020. *Journal of the University of Duhok*, 26(1), 1–20.
- Alghafes, M., Al-Malkawi, H.N., & Pillai, R. (2024). Influence of key ESG factors on Islamic banks' financial performance: Evidence from GCC countries. *Research in International Business and Finance*, 71, 102445.
- Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821–856.
- Asaad, A., & Al-Saad, A. (2025). The role of Islamic finance in achieving sustainable development: Islamic banks as a model. *International Islamic University Malaysia Journal*, 9(2), 1–22.
- Avdukic, A., & Asutay, M. (2025). Testing the development impact of Islamic banking: Islamic moral economy approach to development. *Economic Systems*, 49(2), 101229.
- Chapra, M.U., Khan, S., & Al Shaikh-Ali, A.A. (2008). *The Islamic vision of development in the light of Maqasid al-Shari'ah*. London: The International Institute of Islamic Thought.
- Chow, G.C., & Lin, A. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4), 372–375.

- Dusuki, A.W., & Bouheraoua, S. (2011). The framework of Maqasid al-Shari'ah and its implication for Islamic finance. *ISRA International Journal of Islamic Finance*, 2(2), 316–336.
- Espinoza, R., & Senhadji, A. (2011). How strong are fiscal multipliers in the GCC? (IMF Working Paper 11/61). Washington: International Monetary Fund.
- Gylfason, T., Herbertsson, T., & Zoega, G. (1999). A mixed blessing: Natural resources and economic growth. *Macroeconomic Dynamics*, 3(2), 204–225.
- Hvidt, M. (2013). Economic diversification in GCC countries: Past record and future trends (Kuwait Programme Research Paper 27). London: London School of Economics.
- Levine, R. (2005). Finance and growth: Theory and evidence. In Aghion, P., & Durlauf, S.N. (Eds.). *Handbook of economic growth* (Vol. 1A, pp. 865–934). Amsterdam: Elsevier.
- Nakibullah, A. (2018). Economic diversification in Bahrain. *Applied Economics and Finance*, 5(5), 67–74.
- Otman, A. (2022). Islamic banks and sustainable development goals in the Arab world: A case study of selected countries. Abu Dhabi: Arab Monetary Fund.
- Pesaran, M.H., Shin, Y., & Smith, R.J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3), 289–326.
- Phillips, P.C.B., & Hansen, B.E. (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies*, 57(1), 99–125.
- Sax, C., & Steiner, P. (2013). Temporal disaggregation of time series. *The R Journal*, 5(2), 80–87.
- World Commission on Environment and Development. (1987). *Our common future*. Oxford: Oxford University Press.

Appendix. Construction of the composite indices

Components are standardised to zero mean and unit variance across the ten annual observations before aggregation; each composite is an equally weighted average of its standardised components. Components marked † are specified in the study design but do not appear in the surviving estimation dataset, so the table documents construction as specified rather than as verified.

Table A1. Component-level documentation

Index / component	Definition	Source	Weight
Economic diversification (ED)			
Non-oil GDP share	Non-oil sectors as % of real GDP	MOFNE / IGA	1/3
Non-oil revenue share	Non-oil revenue as % of total government revenue	MOFNE	1/3
Manufacturing share	Manufacturing value added as % of GDP	IGA / World Bank	1/3
Sustainable development (SD)			
Economic dimension †	Non-oil real growth; macroeconomic stability measures	MOFNE / IMF	1/3
Social dimension	Human Development Index; public spending on education and health	UNDP / MOFNE	1/3
Environmental dimension †	Environmental expenditure; energy efficiency indicators	IGA / World Bank	1/3
Islamic banks' performance (IBP)			
Return on assets; return on equity	Sector-aggregate net income over assets and over equity (%)	CBB; bank reports	1/4 each
Capital adequacy ratio	Regulatory capital over risk-weighted assets (%)	CBB	1/4
Non-performing financing	Non-performing over total facilities (%), reverse-coded	CBB	1/4
Developmental and sustainability dimensions †	Industrial and SME financing shares; environmental and social project financing; financial inclusion	CBB; bank reports	not in dataset

Note: MOFNE = Ministry of Finance and National Economy; IGA = Information and eGovernment Authority; CBB = Central Bank of Bahrain. The three development dimensions have standard deviations of 0.81, 0.39 and 0.79 rather than 1.00, the signature of averaging several standardised components rather than a single series, which is consistent with the specification above but does not verify it.