ADB

Asian Development Bank

March 2014

Development Effectiveness Review Rating System

For Inquiries: A. Andrysiak, Strategy and Policy Department
                   (Ext. 6969)

ABBREVIATIONS

ADB   –   Asian Development Bank
DEfR   –   Development Effectiveness Review
MDG   –   Millennium Development Goal

NOTE

In this report, "$" refers to US dollars.

CONTENTS

Page

I. INTRODUCTION

1.   The performance scorecard in the Development Effectiveness Review (DEfR) is designed to provide useful summary information to the Board, Management, and stakeholders. It uses signals to reflect performance relative to targets (colors) and annual change (arrows). The 2013–2016 corporate results framework introduced several new scoring methods to determine these signals.[1]

2.   This technical note describes preexisting and new scoring methods for results indicators (section II) and the composite indicators in the summary scorecard (section III).[2] The scoring process has been automated based on these predetermined and transparent rules to enhance efficiency, speed, and reliability.

II. RESULTS INDICATORS

A.   Colors to Represent Performance

3.   The following colors represent indicator performance in the scorecard: green check (*at or above target*), green (*on track*), amber (*on track but watch*), and red (*off track*). The 2013–2016 corporate results framework utilizes four preexisting and eight new scoring methods to determine these colors (Table 1).[3]

Table 1: Scoring Methods

Category	Scoring method	Notes
Regional performance	1. Millennium Development Goal 2. Progress	Indicator performance in Level 1 (Development Progress in Asia and the Pacific) is assessed relative to Millennium Development Goal targets for 2015 or baseline values.
Core operational results	3. Achievement rate	The ratio of realized and planned outputs and outcomes for Core Operational Results (Level 2) is compared to a satisfactory achievement rate.
2016 target	4. Point target 5. Range	Most indicators are assigned 2016 targets. For such indicators gradual progress from the baseline value is foreseen. As it may be unrealistic to meet a challenging 2016 target in the earlier years of the new results framework, performance is assessed against implicit interim targets that are derived from the baseline and the 2016 target.
Annual target	6. Performance standard 7. Maintain 8. Above baseline	Indicators with annual targets have minimum and/or maximum threshold values that need to be satisfied every year up to 2016. Failure to meet an annual target is cause for immediate concern and flagged in the corporate scorecard as *off track*. Annual targets may be based on baseline values if performance is already close to satisfactory levels.

---

[1] ADB. 2012. *Review of the ADB Results Framework*. Manila.
[2] This guidance is for internal use by the staff of the Asian Development Bank (ADB) and will be updated as necessary to reflect good practices and lessons learned. It builds on earlier documentation on scoring methods: ADB. 2011. *Development Effectiveness Review 2010 Report*. Manila (Supplementary Appendix A).
[3] An overview of scoring methods and related results indicators is in Annex 1.

Category	Scoring method	Notes
Accumulation	9. Cumulative indicator 10. Cumulative ratio	Some indicators measure cumulative performance during the results framework period. Contrary to other indicators, the data period that is covered differs for each year.
Other	11. Increase 12. Monitor	The signal for the final scoring method, increase, is based on improvements relative to the baseline and annual change. No colors are assigned to indicators that are monitored.

Source: ADB Strategy and Policy Department.

B.   Scoring Methods

4.   The following overview describes these scoring methods in more detail.

1.   Millennium Development Goal

5.   The Millennium Development Goal (MDG) scoring method continues to be applied to indicators with MDG targets in Level 1. Attainability assessments are determined by the Asia-Pacific regional MDG report method:[4]

At or above target	Latest MDG data better than cut-off value
On track	2015 projection better than cut-off value
Off track	2015 projection worse than 2015 cut-off value

6.   Some minor refinements were made in line with these general rules. First, since the data sources used for population living on less than $1.25 (PPP) per day and under-5 child mortality do not have a projection for 2015, their classification was drawn from the Asia-Pacific regional MDG report. Second, land area covered by forests does not have a numerical target, but is scored against the MDG objective to "increase the proportion/share of land area covered by forest."[5] Third, the revised results framework for 2013–2016 applies a cut-off value of 1.00 for the ratio of boys and girls in secondary and tertiary education (footnote 1).

2.   Progress

7.   The existing "progress" scoring method is applied to the remaining indicators in Level 1, with progress relative to the baseline being scored as *on track*. Scoring for performance at or below the baseline has been refined:

On track	Region's performance (rounded) progressed against baseline
On track but watch	Region's performance (rounded) at or below baseline for 1 or 2 consecutive years
Off track	Region's performance (rounded) at or below baseline for 3 or more consecutive years

---

[4] See ADB, the United Nations Economic and Social Commission for Asia and the Pacific, and the United Nations Development Programme. 2013. *Asia-Pacific Aspirations: Perspectives for a Post-2015 Development Agenda.* http://www.adb.org/publications/asia-pacific-aspirations-perspectives-post-2015-development-agenda, p. 7.
[5] The estimates for 1990 were 22.2% for ADB developing member countries and 24.5% for Asian Development Fund countries.

3. Achievement Rate

8. The "achievement rate" scoring method continues to be used to assess the ratio of realized versus planned outputs and outcomes for Core Operational Results in Level 2:

|	
At or above target	Achievement rate (rounded) equals or is better than the satisfactory achievement rate
Off track	Achievement rate (rounded) is below the satisfactory achievement rate

4. Point Target

9. The "point target" scoring method has been used since the 2010 DEfR to reflect the likelihood of attaining long-term targets.[6] The setting of a baseline and 2016 target for results indicators allows for determining implicit intermediate targets that need to be achieved to make attaining the long-term target more likely.[7] Comparing results with these implicit intermediate targets enables the Asian Development Bank (ADB) to determine progress of the results indicator toward the 2016 target and assign a corresponding progress color.

10. The progress color depends on the year and the results indicator value. Amber is assigned as long as the results indicator progresses toward the target within a predetermined band. This band allows for sufficient granularity and avoids frequent and unclear signal changes for results indicators that are progressing toward the target with some minor fluctuations. Within the amber band, a results indicator is deemed to be *on track* to attain the 2016 target, but requires close monitoring as it could easily move *off track* by merely remaining stable.

11. If the results indicator performs worse than the lower border of the band, it is considered to be *off track* and red is assigned. If it performs better than the upper border of the band, the results indicator is *on track* and green is assigned. Finally, green with a checkmark indicates that the rounded-off results indicator is *at or above target*. Note that colors are only used to assess progress toward the target year. In 2016, the target has been either achieved or not. The rules are visually summarized in Figure 1.

Figure 1: Point Target

---
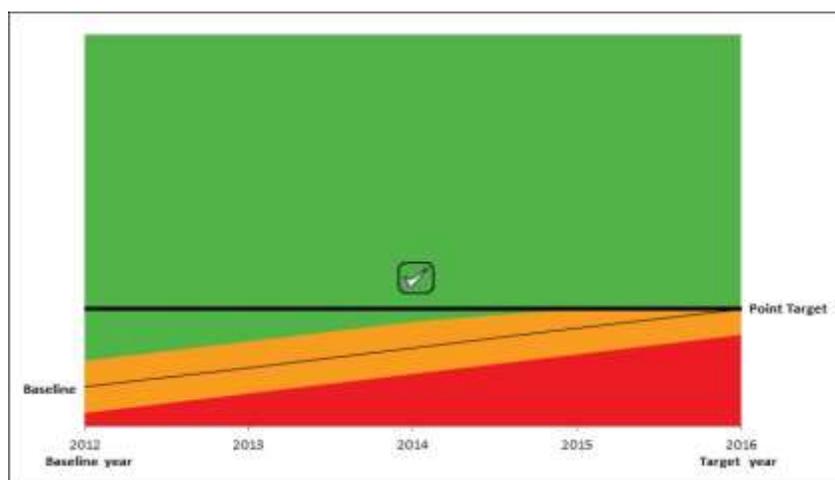
[6] ADB. 2011. *Development Effectiveness Review 2010 Report*. Manila (Supplementary Appendix A).
[7] For simplicity, a constant slope is assumed.

12.   The need for prudence and meaningful signals are important considerations. They require careful balancing when determining bandwidths. There are significant differences in volatility, so assigning the same fixed bandwidth for each results indicator would be inappropriate and could result in many bandwidths being too narrow or too wide to give appropriate signals to management. ADB developed a nonparametric approach that classifies results indicators into three distinct bandwidth groups based on a parametric analysis of indicator volatility:

(i)   **Volatility-based bandwidth.** To have meaningful signals, bandwidth should relate to indicator volatility. For results indicators with low volatility, a red alert signal should be given sooner when the distance between the result and the implicit intermediate target increases. As a first step, volatility-based bandwidths are calculated. This provides useful information on indicator volatility, although the bandwidths are often so wide that they would be imprudent for management use. For several results indicators, these initial calculations result in total bandwidths that exceed 100% of the average baseline and target. Applying bandwidths that are too wide leads to low-quality management information: red signals would be significantly delayed if a results indicator starts to deteriorate, including when the results indicator has fallen far below both baseline and target.

(ii)   **Bandwidth classifications.** Based on the calculated volatility-based bandwidths, results indicators are classified in three distinct categories. The maximum bandwidth is set at 7.5% of the average of baseline and target.[8] Narrower bandwidths reduce the chance of delayed red signals, in particular for results indicators with very high calculated volatility or limited data availability. Using three bandwidth classifications also allows simplicity while ensuring prudence and objectivity. Results indicators are assigned a corresponding prudent bandwidth based on their category:

(a) Low volatility   2.5% of average of the baseline and target
(b) Medium volatility[9]   5.0% of average of the baseline and target
(c) High volatility   7.5% of average of the baseline and target

(iii)   **Management adjustment.** The point target method continues to include the possibility of a management adjustment. A single adjustment was made to generate a more prudent signal in Level 4.[10] For some indicators, the 2012 baseline (which was updated from the 2011 baseline in the R-paper) already exceeds the 2016 targets (footnote 1). In such cases, scoring will be as follows: green check when performance equals or is better than this target, and red when performance falls short of this target.

5.   Range

13.   The "range" scoring method is a variation of the point target scoring method and is introduced for the two budget indicators that are assigned a target range rather than a specific target for 2016: (i) internal administrative expenses per $1 million disbursement, and (ii) share of operational expenses for portfolio management.

---

[8] Bandwidths are one-sided and a percentage of the average baseline and target.
[9] This category also includes results indicators for which the volatility-based bandwidth cannot be determined.
[10] In the absence of sufficient data to calculate volatility, sovereign operations administered with substantial resident mission involvement (%) was expected to be of *low* volatility for Asian Development Fund operations.

14.   Scoring is similar to point targets, but if performance overshoots the specified range, the signal can become *off track* again. An *on track but watch* band of 2.5%, 5%, or 7.5% of the average baseline and target is added to the outer range to allow for minor overshooting in intermediate years. Similar to the point target method, colors are only used to assess progress toward the target year (Figure 2). In 2016, the target range has been either achieved or not. Although the range method includes the possibility of a management adjustment to the volatility-based bandwidth, no such adjustments were made.

**Figure 2: Range**
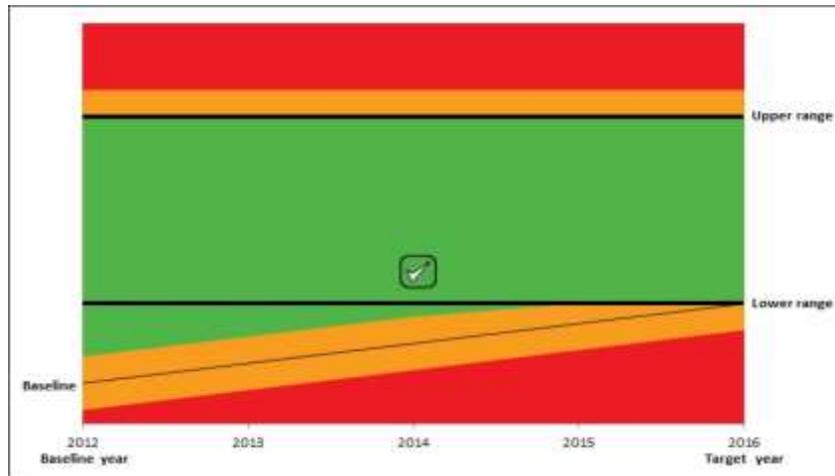
6.   Performance Standard

15.   The "performance standard" scoring method is used for indicators with minimum and/or maximum threshold values that need to be satisfied throughout the results framework period. Failure to meet annual targets is cause for immediate concern and flagged as *off track* (Figure 3).

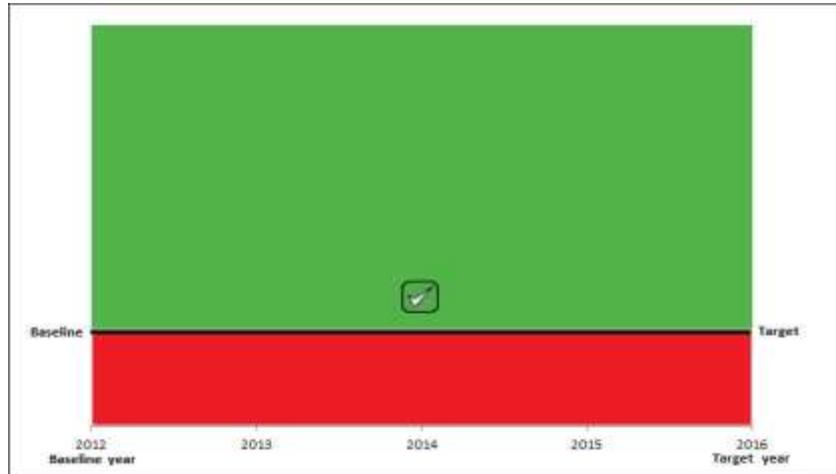|	
At or above target	Performance (rounded) is better than or equal to performance standard
Off track	Performance (rounded) is worse than performance standard

Figure 3: Performance Standard
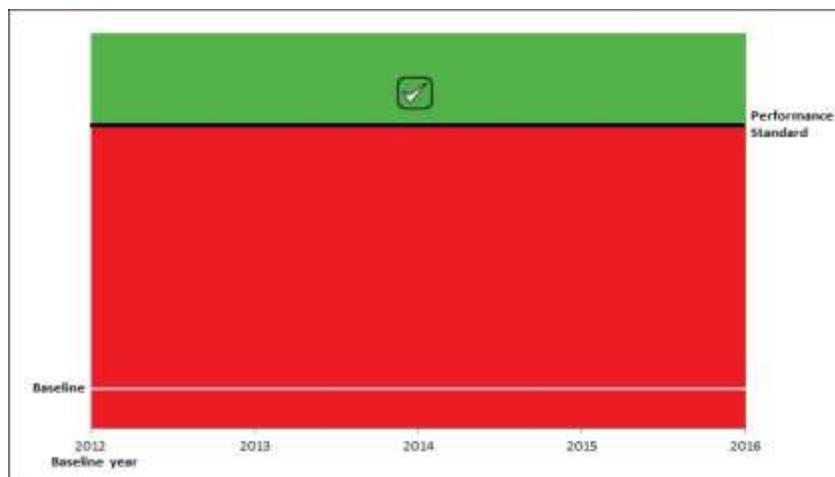
7.   Maintain

16.   The "maintain" scoring method is a variation of the performance standard method. It sets the threshold value that needs to be met in each year at the baseline level (Figure 4).

At or above target   Performance (rounded) is better than or equal to baseline
Off track   Performance (rounded) is worse than baseline

Figure 4: Maintain

8.   Above Baseline

17.   The "above baseline" scoring method is a variation of the performance standard method. It sets the threshold value that needs to be met each year at one measurement unit above the baseline level (Figure 5). Hence, the threshold value for an indicator with a baseline of 24.8 is set at 24.9.

At or above target   Performance (rounded) is better than or equal to baseline plus one
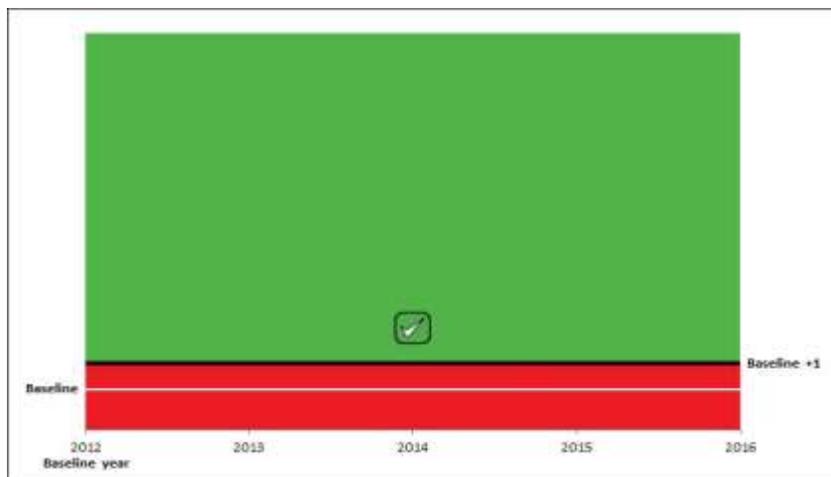*Off track*   Performance (rounded) is worse than baseline plus one

Figure 5: Above Baseline

9. Increase

18. The "increase" scoring method is used for web-distributed knowledge solutions. Its signal is based on two criteria: (i) the average annual improvement since baseline year is larger than or equal to 3%, and (ii) the annual improvement is larger than or equal to 3%.

On track — Both are met
On track but watch — Only one is met
Off track — Neither is met

10. Cumulative Indicator

19. The scoring method for "cumulative indicators" is used for project development transactions for public–private partnerships (total number from 2013). If the indicator is not expected to ramp up gradually, the scoring method assumes that performance increases in equal increments each year and accumulates to the cumulative target (hence a straight target line). Some deviation of this line is allowed in which the indicator is considered to be *on track but watch*. The standard initial bandwidth is 10% of the average baseline and target, which would result in bandwidths of 7.5% (2013), 5% (2014), and 2.5% (2015).

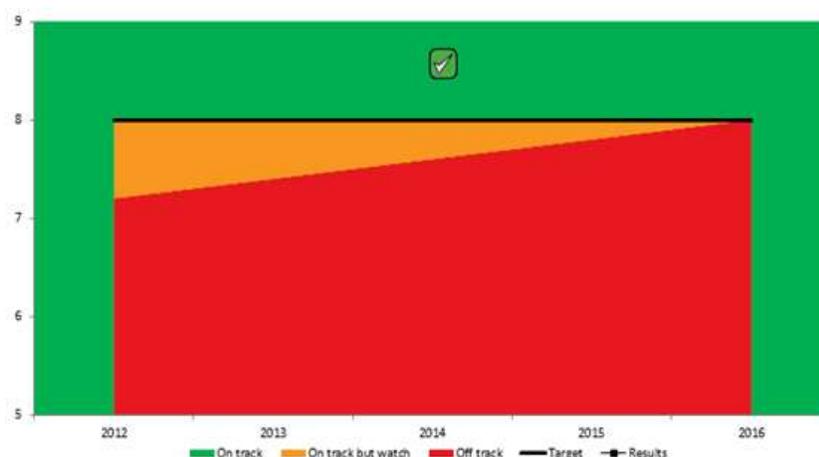**Figure 6: Cumulative Indicator**
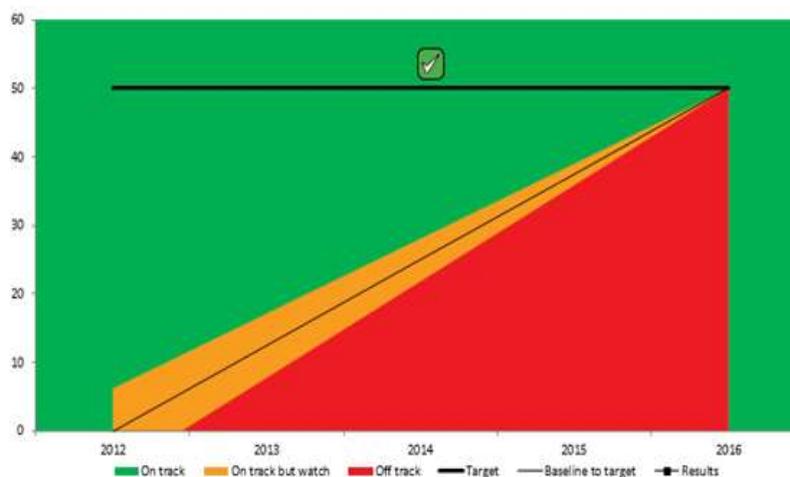
20. Similar to the point target and range methods, the cumulative indicator method includes the possibility of a management adjustment to the standard initial bandwidth for an indicator where a gradual performance ramp up is foreseen. In such cases, performance is still considered *on track but watch* in earlier years although deviating considerably from the straight baseline to target line. For project development transactions for public–private partnerships (total number from 2013), a management adjustment was made as performance is expected to accelerate in later years.

11. Cumulative Ratio Indicator

21. This scoring method for "cumulative ratios" is used for cumulative public–private partnership leveraging achieved using ADB financing (ratio of leveraging amount to ADB financing approved). A cumulative ratio becomes more and more certain as it approaches the target year. Some deviation of the targeted ratio is allowed in which the indicator is considered to be *on track but watch* (Figure 7). The standard initial bandwidth is 10% of the average baseline and target, which results in bandwidths of 7.5% (2013), 5% (2014), and 2.5% (2015).

Figure 7: Cumulative Ratio Indicator

22.   Similar to the point target and range scoring methods, the cumulative ratio includes the possibility of management adjustment to the standard initial bandwidth for indicators where a gradual performance ramp-up is foreseen. No such adjustment was made.

12.   Monitor

23.   No color is assigned for indicators that are monitored.

C.   Signals for Biennial or Triennial Data

24.   Signals for results indicators are based on the latest available data within the results framework period. For instance, a 2014 signal for quality at entry of sovereign projects rated satisfactory will be carried over to 2015 although no new data will become available in that year.[11] This ensures that composite indicators provide a more complete picture of performance.

D.   Arrows to Report Annual Change

25.   Arrows represent annual deteriorations or improvements. A downward or upward arrow is shown if the change is large enough to justify a signal (Figure 8). No arrow is shown if annual change remains below a threshold of 3% of the most recent previous value. For results indicators that are monitored, the arrow shows the direction of movement only. No arrows are shown for Core Operational Results in Level 2, as there is no comparison to past baseline achievement rates.

Figure 8: Annual Change of Results Indicators

⇧   significant improvement (≥3%)

remained stable (no arrow)

⇩   significant deterioration (≥3%)

---

[11] No signals are calculated for baseline years, so there are no signals for quality at entry in 2013.

E.   Integrating Progress and Annual Change

26.   Single integrated symbols indicate both target attainability and periodic change (tables 2 and 3). This provides more information at a glance. For instance, a green box with a downward arrow means that performance is *on track* but deteriorated from the previous year. This might prompt management to analyze the indicator in more detail. A red box with an upward arrow may show that measures are already paying off, although progress is still unsatisfactory.

Table 2: Signals in Level 1

Progress	Signal	Annual Change [a]
MDG Indicators		
At or above target—region's performance at or above cutoff value for 2015	✔	
On track—region's performance on track to achieve cutoff value for 2015	⬆	Improved
	⬛	Remained stable
	⬇	Deteriorated
Off track—region's performance fell short of desired progress to attain cutoff value for 2015	⬆	Improved
	⬛	Remained stable
	⬇	Deteriorated
Non-MDG Indicators		
On track—region's performance progressed against baseline	⬆	Improved
	⬛	Remained stable
	⬇	Deteriorated
On track but watch— region's performance at or below baseline for 1 or 2 consecutive years	⬆	Improved
	⬛	Remained stable
	⬇	Deteriorated
Off track—region's performance at or below baseline for 3 or more consecutive years	⬆	Improved
	⬛	Remained stable
	⬇	Deteriorated

MDG = Millennium Development Goal.

[a]   An arrow indicates a significant (≥ 3%) change from previous performance: an arrow pointing up indicates improvement; an arrow pointing down indicates deterioration. There are no arrows for Core Operational Results (level 2), as there is no comparison to past and/or baseline achievement percentages. For results indicators that are monitored, the arrow shows the direction of movement only.

Source: ADB Strategy and Policy Department.

Table 3: Signals in Levels 2–4

Progress	Signal	Annual Change [a]
At or above target—performance at or above target	✅	
On track—performance exceeded desired progress to attain 2016 target	⬆️ 🟩 ⬇️	Improved Remained stable Deteriorated
On track but watch—performance met desired progress to attain 2016 target but requires close monitoring	⬆️ 🟨 ⬇️	Improved Remained stable Deteriorated
Off track—performance fell short of desired progress to attain target	⬆️ 🟥 ⬇️	Improved Remained stable Deteriorated
Monitor—no performance assessment	⬆️ ⬜ ⬇️	Increased Remained stable Decreased

[a] An arrow indicates a significant (≥ 3%) change from previous performance: an arrow pointing up indicates improvement; an arrow pointing down indicates deterioration. There are no arrows for Core Operational Results (level 2), as there is no comparison to past and/or baseline achievement percentages. For results indicators that are monitored, the arrow shows the direction of movement only.

Source: ADB Strategy and Policy Department.

III. COMPOSITE INDICATORS

A. Colors to Report *Off-Track* Indicators

27.   The color of composite indicators is based on the percentage of *off track* (red) results indicators, which are the areas of immediate concern. Consequently, the composite indicator ignores annual change in individual results indicators unless they resulted in an altered *off track* status. Monitored indicators play no role either. These rules stress whether or not underlying results indicators are *on track* toward attaining targets (Table 4). An exceptional rule is used for situations in which half of the results indicators are red and the other half are green. In this case, an amber, rather than red, signal is assigned.

Table 4: Composite Signals

Signal		Explanation
🟢	***good***	Two-thirds or more of results indicators in the group achieved a green or amber signal.
🟡	***mixed***	More than half but less than two-thirds of results indicators in the group achieved a green or amber signal. [a]
🔴	***poor***	Half or less of results indicators in the group achieved a green or amber signal.

[a] Exception: If half of results indicators are *off track* (red) and the other half exceed desired progress or are at or above target (green), a composite rating of amber, rather than red, is given.

Source: ADB Strategy and Policy Department.

B.   Results Indicators without Signals

28.   To provide a complete picture of the entire group of underlying results indicators, composite indicators include biennial or triennial results indicators where data in the previous year or years were unavailable (para. 24). However, if there are results indicators for which no signal color could be determined within the results framework period, composite signals are only calculated if at least half of underlying results indicators have a signal.

ANNEX 1: SCORING METHODS AT A GLANCE

Type	Target	Status	Notes		Applied to [a]
MDG	2015	Existing	• Better than MDG target	✅ (green check)	Level 1, indicators 1, 5–8, 10, 14–17
			• Projection better than MDG target	🟩 (green)	
			• Projection worse than MDG target	🟥 (red)	
Progress (non-MDG)	Annual	Existing	• Better than baseline	🟩 (green)	Level 1, indicators 2–4, 9, 11–13, 18–22
			• Rounded value at or below baseline for 1 or 2 consecutive years	🟨 (yellow)	
			• Rounded value at or below baseline for 3 or more consecutive years	🟥 (red)	
Achievement rate	Annual	Existing	• Better than target	✅ (green check)	Level 2 (Core Operational Results)
			• Worse than target	🟥 (red)	
Point target	2016	Existing	• At or above target	✅ (green check)	Level 2 (Quality at Completion)
			• Within band from baseline to target	🟨 (yellow)	Level 3, indicators 3, 6–9, 11, 17–19, 21, 23–24, 28
			• Above band from baseline to target	🟩 (green)	Level 4, indicators 2, 6, 9
			• Under band from baseline to target	🟥 (red)	
Range	2016	New	• See point target		Level 4, indicators 4, 5
			• Within the target range	✅ (green check)	
			• An *on track but watch* area is added to the top of the range to allow for temporary overshooting	🟨 (yellow)	
Performance standard	Annual	New	• Performance standard or better	✅ (green check)	Level 3, indicators 1, 2, 4, 5, 20, 26, 27
			• Worse than the performance standard	🟥 (red)	Level 4, indicators 1, 3, 7
Maintain	Annual	New	• Baseline or better	✅ (green check)	Level 3, indicator 10
			• Worse than the baseline	🟥 (red)	Level 4, indicator 8
Above baseline	Annual	New	• One measurement unit above the baseline or better	✅ (green check)	Level 3, indicator 22
			• Worse than the baseline plus one measurement unit	🟥 (red)	
Increase	Annual	New	• Criteria used: 1. Average annual improvement since baseline year larger than or equal 3% threshold 2. Annual improvement larger than or equal 3% threshold		Level 3, indicator 25
			• Both	✅ (green check)	
			• Only one	🟨 (yellow)	
			• Neither	🟥 (red)	

14

Type	Target	Status	Notes		Applied to [a]
Cumulative indicator	2016	New	• Performance is expected to increase in equal increments each year and accumulate into the cumulative target		Level 3, indicator 12
			• Above cumulative target	✅ green	
			• Cumulative indicator is assumed to hover along this expectation (*on track but watch*) with less leeway in later years as less time remains to correct performance	🟨 yellow	
			• Below band	🟥 red	
Cumulative ratio indicator	2016	New	• Above cumulative target: green check	✅ green	Level 3, indicator 13
			• Cumulative ratio is assumed to hover around the target (*on track but watch*) but allowed more leeway to fall below the target in initial years	🟨 yellow	
			• Below narrowing band	🟥 red	
Monitor	None	New	• Only arrows are shown • No impact on summary scorecard	⬜ white	Level 3, indicators 14–16

[a] Refers to the results framework indicator numbers used in the 2013–2016 corporate results framework and development effectiveness reviews.
Source: ADB Strategy and Policy Department.

ANNEX 2: DETERMINING BANDWIDTHS

The following three-step approach to determine bandwidths has been in place since 2011. The bandwidth calculation described in this annex is used for the point target and range scoring methods.

A.   Step 1: Volatility-Based bandwidth

Volatility-based bandwidths are calculated as a first crude input for deriving the bandwidth classification of results indicators. The volatility-based bandwidth around the target line Y^T, is determined by assuming that the observed results indicator, Y_i for year i can be expressed by the baseline value, Y_0, the average change needed to reach the target, $d^T = \frac{(\text{Target}- Y_0)}{\text{\# years to Target}}$ and a random error, ϵ_i, such that

$$Y_i = Y_0 + d^T \times i + \epsilon_i = Y_i^T + \epsilon_i,$$

and $\epsilon_i \sim N(0, \sigma_\epsilon^2)$.

ϵ_i is a random variable with zero mean. Volatility of ϵ_i will determine the volatility-based bandwidth, with a highly volatile error term leading to a broader bandwidth and vice versa. Estimating the parameters of ϵ_i is a challenge due to the lack of sufficient data for many results indicators. Where data are available, the standard deviation of ϵ_i can be estimated using the linear regression model. With the estimated variance, the confidence intervals (or bandwidth) around the observed value can be estimated using t-statistic. The volatility-based bandwidth will be $\pm t_{(\alpha/2, n-2)} \times \hat{\sigma}_\epsilon$, where $\hat{\sigma}_\epsilon$ is the estimated standard deviation of ϵ_i. The confidence level is set at 95%.

Since there are results indicators for which the targets are lower values than the baseline (e.g., average sovereign operations processing time and months from fact-finding to effectiveness), a different set of criteria has to be applied depending on the type of results indicator.

For indicators where higher values are targeted:

Criteria	Color
$Y_i > Y_i^T$ + margin of error and $Y_i <$ Target	🟩
Y_i^T − margin of error $\leq Y_i \leq Y_i^T$ + margin of error and $Y_i <$ Target	🟧
$Y_i < Y_i^T$ − margin of error and $Y_i <$ Target	🟥

For indicators where lower values are targeted:

Criteria	Color
$Y_i < Y_i^T$ + margin of error and $Y_i >$ Target	🟩
Y_i^T − margin of error $\leq Y_i \leq Y_i^T$ + margin of error and $Y_i >$ Target	🟧
$Y_i > Y_i^T$ − margin of error and $Y_i >$ Target	🟥

B.   Step 2: Initial Bandwidth Classifications

Based on these volatility-based bandwidths, results indicators are classified in the following three broad categories and assigned a corresponding bandwidth.

Category	Volatility-Based Bandwidth	Assigned Prudent Bandwidth
Low volatility	0.0%–5.0%	2.5%
Medium volatility	5.0%–15.0%	5.0%
High volatility	>15.0%	7.5%

Bandwidths are one-sided and a percentage of the average baseline and target.

This second step makes the approach more prudent from a management perspective. As most bandwidths are narrowed in the process, red signals will be given sooner if the results indicator starts to deteriorate. Note that a classification into narrower bandwidths is also likely to provide fewer "*watch*" signals for results indicators that are considered to be *on track*.

C.   Step 3: Reclassification based on Management Judgment

As limited availability of data affects the reliability of volatility calculations, it needs to be supplemented by management judgment in order to derive more prudent results. Such a reclassification may be particularly relevant in situations where (i) the bandwidth cannot be determined, (ii) data availability is limited, (iii) the target year is close, or (iv) there are already predetermined institutional bandwidths.[12]

---

[12] Results indicators for which no volatility-based bandwidth could be calculated due to limited data are usually considered to be of medium volatility.