**Valentin T.Z. Berger, Marc Helbling and Santiago Álvarez Tobar**

Mannheim Centre for European Social Research (MZES), University of Mannheim, Mannheim, Germany

June 2026

# Report of the Revision of the IMPIC dataset update (2011–2018)

## 1. Purpose and Scope

This report assesses the scope, distribution, and substantive relevance of revisions made to the IMPIC dataset for the period 2011 to 2018 (Helbling et al., 2024). The IMPIC dataset measures the restrictiveness of immigration policies across most OECD member states based on the coding of legal documents with the assistance of country experts (Helbling et al., 2017). The dataset currently encompasses 32 OECD member states, with data spanning from 1980 to 2018. Coding captures the level of restrictiveness of specific policies in each country and year. Subsequently, items are aggregated at the country level into five main sub-policy areas as well as an overall score of immigration policy.

The revisions examined in this report were implemented as part of a data quality review to ensure the highest data quality for the data covering the updated period from 2011 to 2018. In particular, the review focused on verifying whether the individual items in the final dataset reflect the raw data provided by the country experts in the best possible way. After the data collection, the data contained in the raw dataset was recoded and scored in order to create the final items and indices, which are made available for use by researchers and the public alike.

The objective of the review was to ensure that the aggregation of individual items into the indicator scores accurately reflected the underlying coding logic. The analysis presented evaluates the scope of the resulting corrections, whether the corrections display systematic patterns across policy sub-fields, countries, or years, and underlines the high data quality of the IMPIC dataset.

The unit of analysis is the country-year observation. Comparisons between the original and revised version of the dataset are conducted both at the cell level (country-year-item) and at the policy sub-field level, allowing for detailed diagnostics of where and how revisions occurred across the dataset.

The structure and coding of the indicators are described in detail in the IMPIC Technical Report (Berger et al., 2024), which provides definitions of all variables and explains the scoring logic used in the dataset. For transparency and documentation purposes, the dataset itself, together with the revised version analyzed in this report, is stored in the project repository (http://www.impic-project.eu/data/), ensuring that both the datasets and the associated documentation are permanently accessible for verification and future updates.

## 2. Data Harmonization and Comparison Setup

To ensure a valid comparison, the original and revised version of the dataset were first aligned to guarantee a one-to-one correspondence across country-year observations. Both

datasets were sorted by country and year, and consistency checks were performed to verify that both datasets contain the same set of country-year units and that the ordering of observations is identical. This ensures that differences identified in subsequent analyses reflect genuine data revisions rather than differences in the dataset structure.

All variables shared by both datasets (excluding identifiers such as country and year) were compared directly. Numeric values were standardized by harmonizing decimal notation. Because decimal numbers are represented with limited precision, extremely small numerical differences can arise from rounding during computation. To avoid counting such artifacts as corrections, a small numerical tolerance of $1 \times 10^{-6}$ was applied so that only substantively meaningful differences were recorded. For each indicator, cell-level differences were computed, and a binary variable was created to identify whether a correction occurred.

## 3. Overall Magnitude of Revisions

Across the full set of 17.664 country-year-items (which corresponds to 69 unique items in 32 countries in eight years) a total of 703 scores were corrected and aligned to the raw data. This means that 3,98 % of all cells exhibit any change between both versions. It can be argued that the revisions were not extensive and did involve only targeted corrections.

At the country-year level, only minor corrections are exhibited for the 69 unique items:

- The average number of corrections is 2,75
- The median number of corrections is 3
- The maximum number of corrections is 7, which occurs for 9 out of 256 country-year observations

No country-year exhibits a high concentration of changes, ruling out the possibility of localized or systemic data problems at the unit level.

## 4. Largest Changes and Three Examples of the Revision

The most substantial potential change is a transition from 0 (most liberal) to 1 (most restrictive score), or vice versa. In these cases, the difference between the two datasets equals ±1, which represents the maximum possible change for a binary variable at the country–year level. This case applies only to Ireland and two specific items (policy area family reunification: language skills and residence permit). The 20 observations with the largest correction across all items can be found in Table 1 in the appendix. Importantly, these changes do not reflect systematic errors in the aggregation of indicators used to construct the component averages. Instead, these discrepancies are most likely attributable to differing interpretations of the coding logic employed for specific variables.

The following paragraph presents three illustrative examples that demonstrate scenarios in which revisions were made (the order does not indicate the frequency of these scenarios). Firstly, it is important to note that a value of 0 denotes the least restrictive outcome for a specific policy feature. In contrast, a value of 1 denotes the most restrictive outcome, and may also refer to the absence of the policy itself. In the original dataset, some cases appear to have been coded as if the absence of a particular restrictive feature implied that the policy

itself was absent. However, the existence of the policy is indicated by other indicators within the same component. The corrections therefore reflect adjustments to the interpretation of the binary coding rather than structural problems in the construction of the aggregated indicators. Secondly, variables represent the average across tracks. Consequently, the number of entry routes/tracks specified by the expert constitutes the foundation for the calculation. Due to the closure of some tracks and the opening of others, the number of tracks considered for the calculation was, in some instances, erroneous. This is predominant to the area of labor migration, for which as many as six primary entry routes could be specified. Thirdly, the revision process entailed careful consideration of the written comments provided by the experts, followed by a critical comparison with the coding.

## 5. Policy Sub-Field Diagnostics

To assess whether certain policy-subfields were more prone to correction, indicators were grouped into IMPIC's five main sub-policy areas, and correction counts were normalized by the number of cells per component. This analysis reveals the following structure:

- Family reunification (correction rate: 9,8 %) and labor migration (correction rate: 7,6 %) are the subjects of the most significant changes. These two sub-fields account for the vast majority of overall corrections, with a total of 89% of all changes occurring within these two areas. Despite displaying the largest average correction magnitudes, these remain small in absolute terms.
- The three sub-fields asylum and refugees, co-ethnics, and control exhibit very low correction rates (0,7 %, 1,0%, respectively 0,6 %), with correspondingly negligible changes of average magnitudes.

This pattern suggests that changes due to the revision were highly concentrated in the two sub-fields family reunification and labor migration, which involved more complex and interpretive coding decisions. By contrast, other sub-fields display high baseline stability. Importantly, no sub-field exhibits unexpectedly large or erratic correction behavior.

## 6. Country-Level Clustering

To evaluate whether changes to the data were disproportionately concentrated in specific countries, total corrections were aggregated by country and normalized by the number of country-years. The results show the following:

- Changes are most evident for Australia (10,1%) and Canada (8,7% of items corrected).
- No corrections were made for three countries (Slovakia, Spain, and Switzerland)

There is no evidence that a small subset of countries drives the revision process. Instead, corrections are broadly and evenly distributed, indicating that no country systematically deviates from the original scores, and that all countries maintain a consistently high standard of data quality.

## 7. Temporal Distribution of Revisions

To assess whether corrections cluster in particular years, diagnostics were conducted at the year level.

- The total number of changes remains stable between 84 (for 2011) and 90 changes per year (for 2017). This corresponds to an approximate change rate of 4,0 %. Therefore, no year exhibits a spike in either the prevalence or intensity of corrections.

This temporal stability indicates that revisions were not driven by year-specific issues. Instead, corrections reflect consistent application of review criteria across the full-time span.

## 8. Stability of Aggregated Component Scores

This paragraph examines the stability of aggregated component scores for the five main policy sub-fields. To illustrate the effect of the corrections on aggregated sub-fields, Figure 1 presents the time series for three countries selected to reflect three different correction profiles: Canada (comparatively high number of item-level corrections: 48 changes), the Netherlands (moderate number of corrections: 24 changes), and Switzerland (no corrections). The vertical axis shows the difference between the revised and original component averages (revised – previous score), where a value of zero indicates no change in the aggregated score.

Overall, the results indicate a high degree of stability across sub-fields and countries. Canada, displayed on the left, has experienced comparatively many item-level corrections, showing that the aggregated effects remain small. Family reunification shows a modest positive shift of roughly +.014 points across all years, while labor migration shows a small negative shift of approximately -.028 to -.031. These adjustments are consistent across years and do not alter the overall temporal structure of the sub-field scores. Asylum and refugees, co-ethnics, and control remain unchanged, as indicated by their coincidence with the zero-reference line.

A similar pattern is observed for the Netherlands, displayed centrally. Family reunification exhibits a small downward adjustment of roughly -.009 across the entire period. Labor migration shows slightly larger variation, shifting from approximately -.024 in the early years to about -.007 around the middle of the period, before settling near -.011. These changes remain modest in magnitude and do not alter the relative ordering or long-term evolution of the sub-field scores. As with Canada, scores for asylum and refugees, co-ethnics, and control remain effectively unchanged.

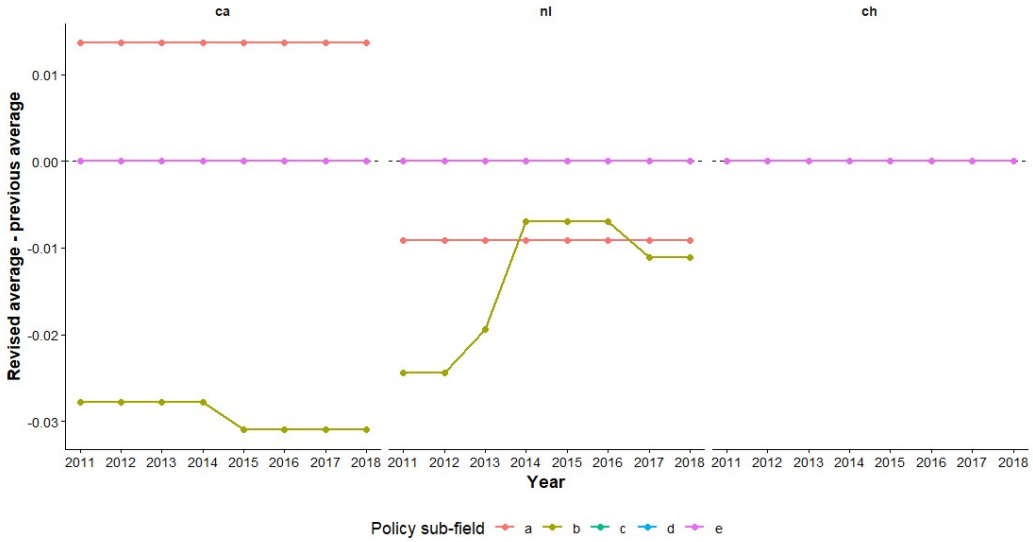
Switzerland, displayed on the right, serves as a low-correction baseline case, the sub-field scores remain identical across the two versions of the dataset. All components coincide with the zero-reference line throughout the period, indicating that the item-level corrections did not affect the aggregated component averages.

Across all three countries and components, the magnitude of changes remains small relative to the 0–1 scale of the index (see Figure 1 in the appendix for all 32 countries). Importantly, the corrections do not produce abrupt discontinuities, trend reversals, or sustained divergence in the sub-field trajectories.

Taken together, these results demonstrate that the item-level corrections improve the accuracy of individual indicators while leaving the higher-level sub-field scores largely

unaffected. The aggregated measures therefore remain stable and comparable across time and countries, indicating that the revisions do not materially alter the substantive interpretation of the dataset.

**Figure 1: Changes in policy sub-fields for three selected countries**




**Note:** The selected countries (Canada -ca- on the left, the Netherlands -nl- in the middle, and Switzerland -ch- on the right) reflect three different correction profiles. Sub-fields without visible lines coincide with the zero-reference line, indicating that the revised and original sub-field average is identical for those cases.

## 9. Occurrence of missing values

The occurrence of missing values (NAs) did not change across the two datasets despite the number of corrections that were conducted. The number of missing values remains stable across both versions, indicating that the revision process did not introduce data loss. Following the exclusion of the sub-field co-ethnics due to its non-applicability in some countries and the resulting numerous missing values, the total number of missing values is 77 or 0.5% of the cases (previous version included 98 missing values). This is attributable to the fact that the majority of NAs in the original database did not stem from issues linked to the aggregation of the respective indicators that constitute each of the averaged values. Instead, these NAs are the result of very clear coding decisions when the experts responded to the questionnaire (specifically, it corresponds to the selection of -8 (not applicable) or -9 (do not know) as a response). The presence of these missing values primarily reflects the construction of a standardized questionnaire applied across 32 OECD member countries and multiple decades, where not all items are applicable or available for every country-year observation.

## 10. Interpretation and Assessment

Taken together, the presented diagnostics provide a coherent and reassuring picture of the revision process:

- Corrections are limited in magnitude and affect a small share of cells.
- No country-year, country, policy sub-field, or year exhibits unusually high concentrations of changes.
- Revisions are concentrated in the two policy sub-fields family reunification and labor, where conceptual complexity is highest, and are minimal in the three other sub-fields.
- The distribution of corrections is stable across both countries and years, despite a small number of localized discontinuities and persistent level shifts in specific component trajectories.

## 11. Conclusion

The analysis presented in this report indicates that the revised IMPIC dataset remains highly stable relative to the previous IMPIC version. Corrections are overall small, evenly distributed, and conceptually interpretable, with no evidence of systematic deviations across policy sub-fields, countries, or years. The revision further enhances the already exceptional data quality of the IMPIC dataset.

## 12. References

Berger V, Bjerre L, Breyer, M, Helbling M, Römer F, and Zobel M. (2024) 'The Immigration Policies in Comparison (IMPIC) Dataset: Technical Report V2', *WZB Discussion Paper* SP VI 2016-201r. Berlin: WZB Berlin Social Science Center

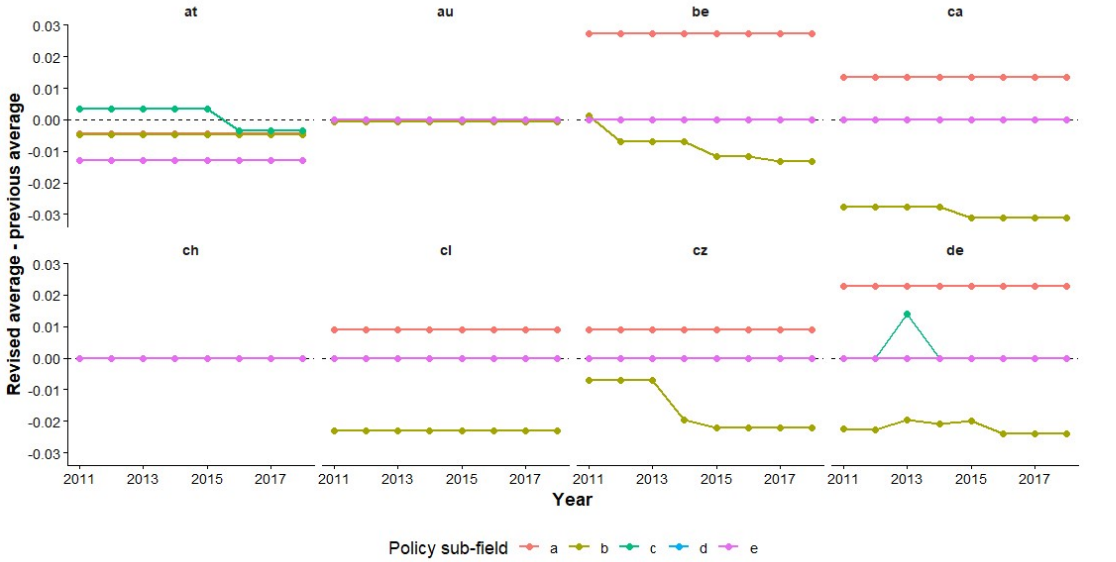
Helbling M., Abou-Chadi T., Berger V., Bjerre L., Breyer M., Römer F., and Zobel M. (2024) 'IMPIC Database v2'. Immigration Policies in Comparison Project (IMPIC).

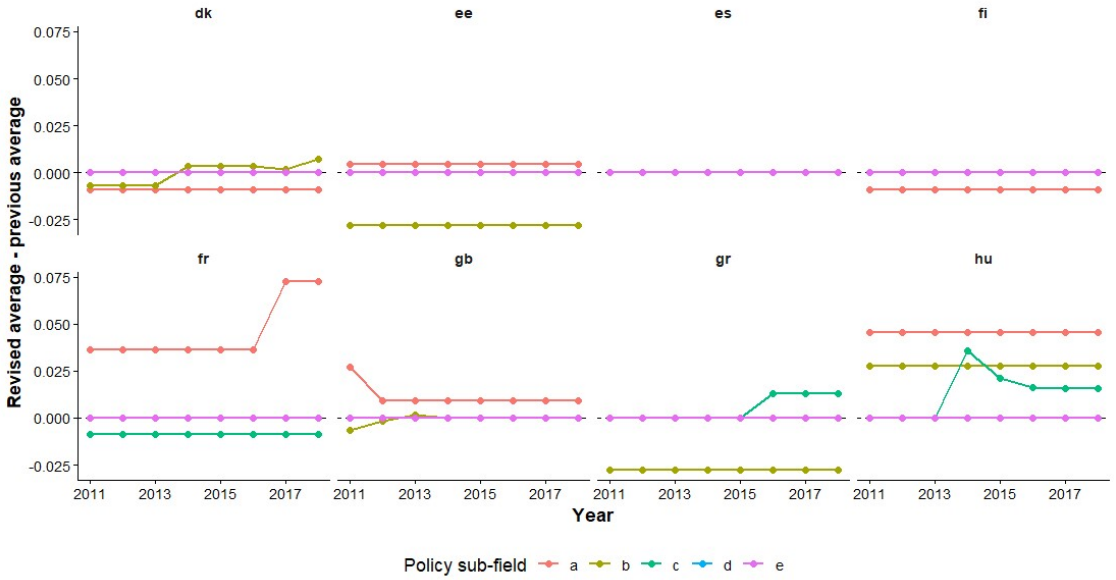
Helbling M., Bjerre L., Römer F., and Zobel M. (2017) 'Measuring Immigration Policies: The IMPIC Database', *European Political Science*, doi: doi.org/10.1057/eps.2016.4

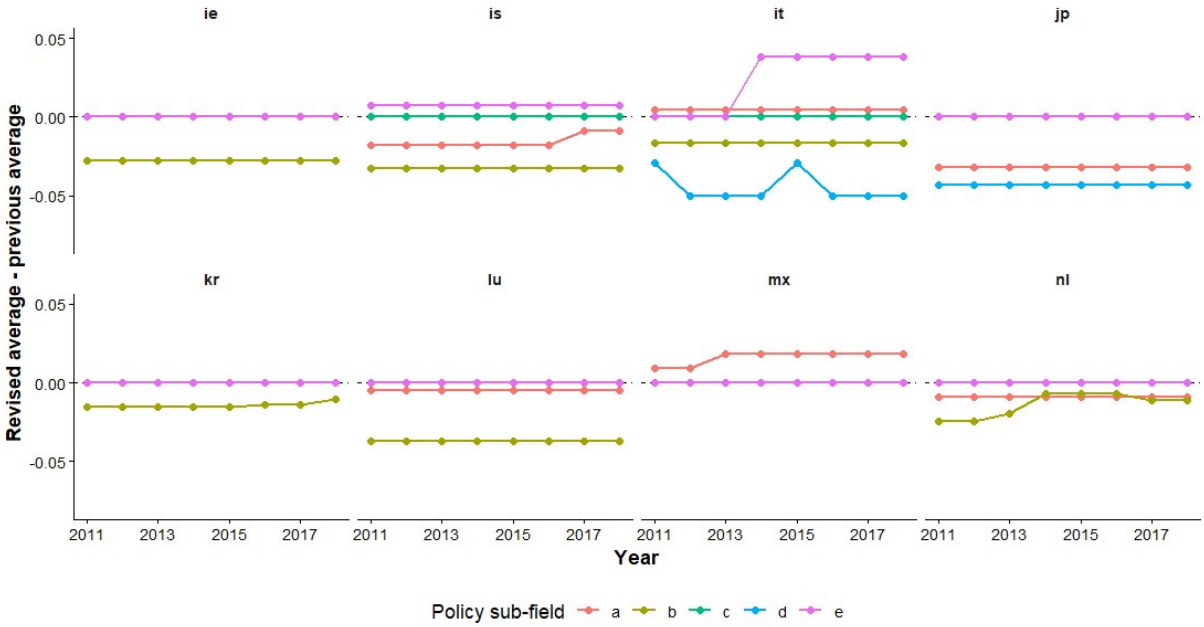
# 13. Appendix

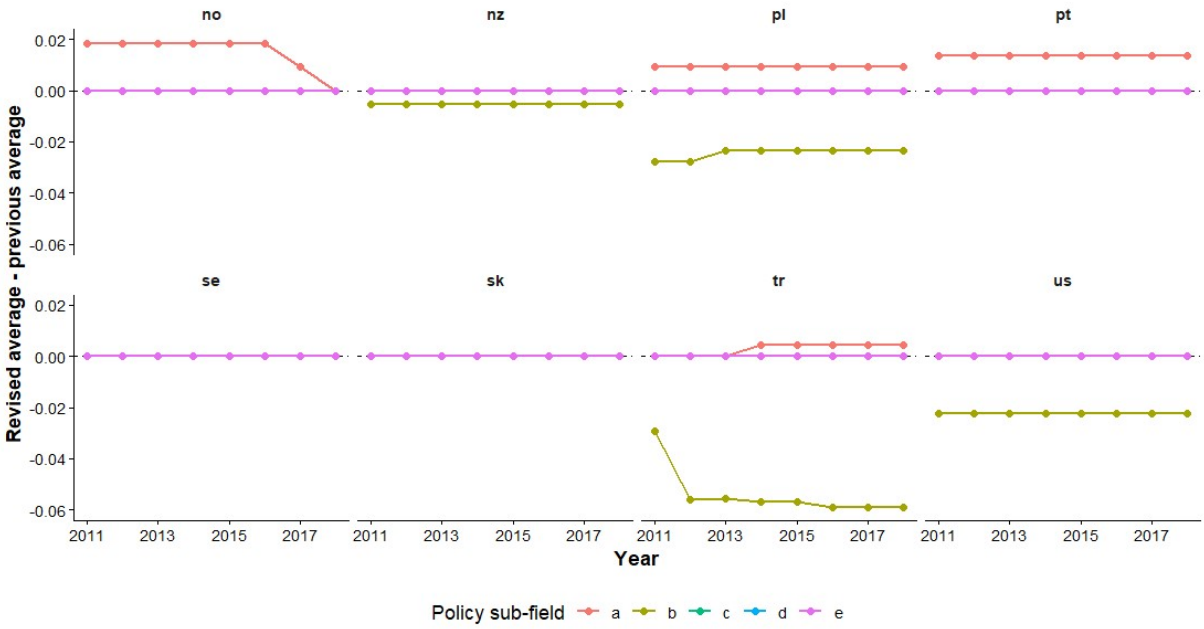
## Appendix 1: Changes in policy sub-fields for all 32 countries

**Note:** Sub-fields without visible lines coincide with the zero-reference line, indicating that the revised and original sub-field average is identical for those cases.




ie
is
it
jp
kr
lu
mx
nl
Revised average - previous average
0.05
0.00
-0.05
2011
2013
2015
2017
Year
Policy sub-field
a
b
c
d
e


no
nz
pl
pt
se
sk
tr
us
Revised average - previous average
0.02
0.00
-0.02
-0.04
-0.06
2011
2013
2015
2017
Year
Policy sub-field
a
b
c
d
e

## Appendix 2: 20 observations with the largest correction across all items

| Country | Year | Indicator | Original value | Revised value | Absolute difference |
|---|---|---|---|---|---|
| Ireland | 2011 | AvgS_a06 | 1 | 0 | 1 |
| Ireland | 2011 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2012 | AvgS_a06 | 1 | 0 | 1 |
| Ireland | 2012 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2013 | AvgS_a06 | 1 | 0 | 1 |
| Ireland | 2013 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2014 | AvgS_a06 | 1 | 0 | 1 |
| Ireland | 2014 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2015 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2016 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2017 | AvgS_a08 | 1 | 0 | 1 |
| Ireland | 2018 | AvgS_a08 | 1 | 0 | 1 |
| France | 2017 | AvgS_a06 | 0 | .8 | .8 |
| France | 2018 | AvgS_a06 | 0 | .8 | .8 |
| Iceland | 2011 | AvgS_b03_1_min | .583 | 0 | .583 |
| Iceland | 2012 | AvgS_b03_1_min | .583 | 0 | .583 |
| Iceland | 2013 | AvgS_b03_1_min | .583 | 0 | .583 |
| Iceland | 2014 | AvgS_b03_1_min | .583 | 0 | .583 |
| Iceland | 2015 | AvgS_b03_1_min | .583 | 0 | .583 |
| Iceland | 2016 | AvgS_b03_1_min | .583 | 0 | .583 |