Visualizing differences between a ternary composition and its center

Jonas Schoeley

December 18, 2017

Von Eynatten etal. (2002) proposed a method to center data in a ternary diagram in order to better visualize the internal structure of very unbalanced compositions. Here I show that data visualized in such a way can easily be interpreted in terms of the percent-point difference of an observation to the compositions average (center).

Figure 1 shows the labor force composition in the year 2016 for 318 regions of the European Union.

The data is quite unbalanced with most regions featuring a high share of employment in the ternary and a low share in the primary sector, leaving most of the ternary diagram empty. In order to better visualize the variance within such unbalanced compositions Von Eynatten etal. (2002) propose to transform the data in a way that moves it towards the center of the diagram, specifically, to pertube the data and the gridlines by the inverse of the compositional mean of the data.

The mean of a composition, also called the *center*, is itself a composition and therefore can be annotated in a ternary diagram. In Figure 1 the center of the data is marked by the intersection of three red-lines: In 2016, the average EU region had a labor-force composition of 4% primary, 24% secondary and 72% tertiary sector.

The grid lines of Figure 1 are repeated in 10 percent-point steps from the red center lines. Such a grid can easily be transformed to show percentage point difference of a composition from its center, one simply has to substract the compositional mean from the grid line labels. Figure 2 shows the labor-force composition of EU regions in terms of their percent-point differences from the average EU region in 2016. The three coordinates defining a point on such a transformed ternary diagram are either negative or positive and always add up to 0. The coordinate (0, 0, 0) denotes the center of the data. Note that the position of the data did not change from Figure 1, only the labeling of the grid.

Performing the centering transformation described by Von Eynatten etal. (2002) on the data and the grid lines while keeping the percent-point difference labels results in a percent-point difference ternary diagram with the compositional mean at its center.

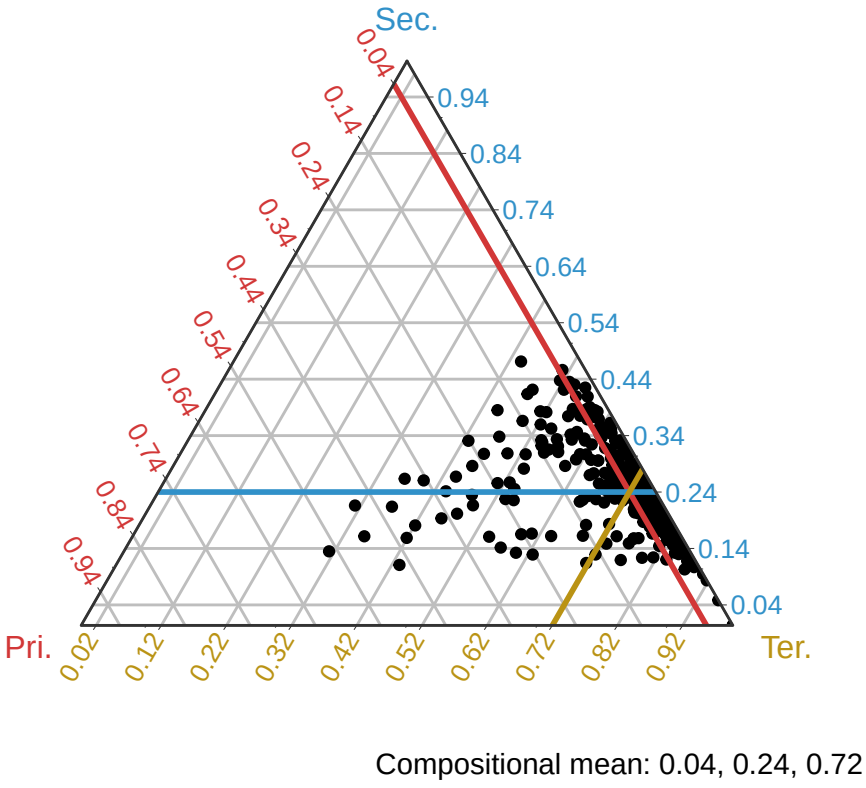


Figure 1: 2016 EU labor force composition by NUTS-2 region. Red lines mark the labor force composition of the average region.

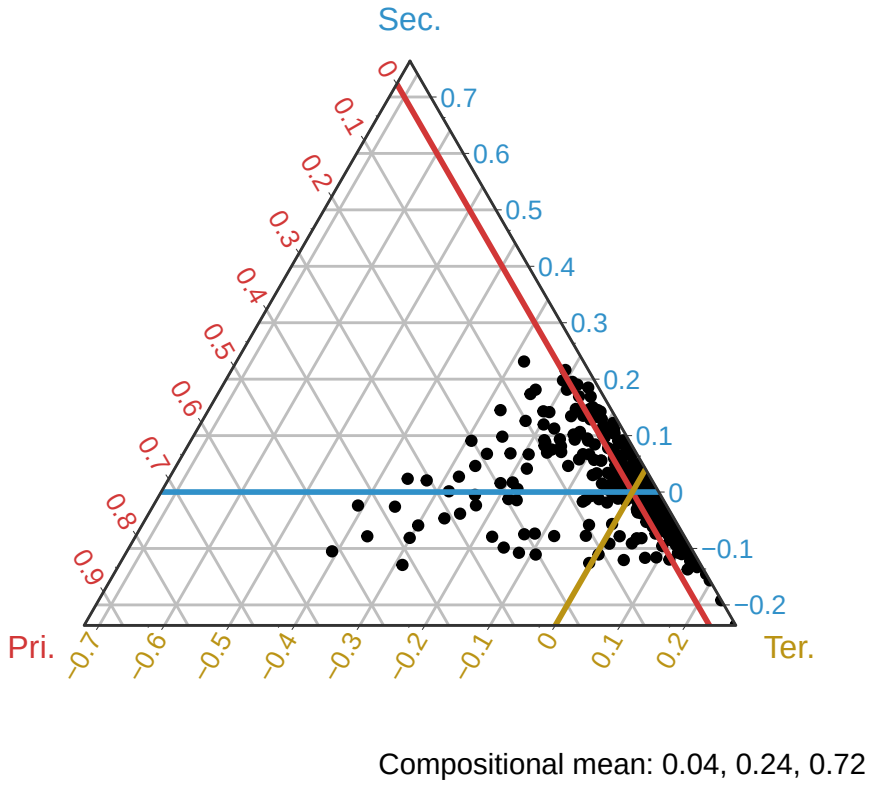
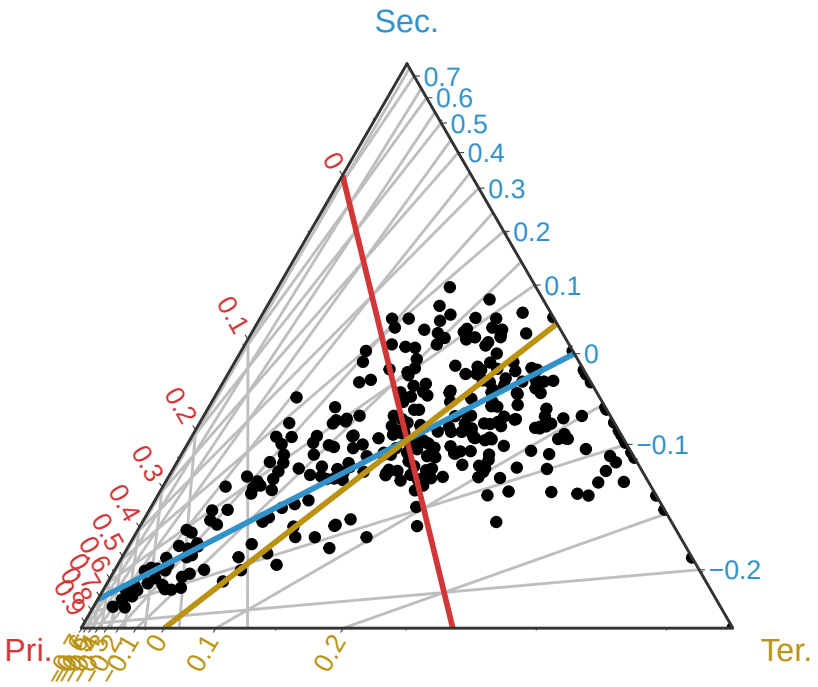


Figure 2: 2016 EU labor force composition by NUTS-2 region in terms of the percent-point difference to the average EU region. Red lines mark the labor force composition of the average region.



The red lines divide the ternary diagram in 6 segments (see figure 4).

- A: More than average share of primary sector labor force, less than average share of secondary and tertiary sector labor force,
- B: more than average share of primary and secondary sector labor force, less than average share of tertiary sector labor force,
- C: more than average share of secondary sector labor force, less than average share of primary and tertiary sector labor force,
- D: more than average share of secondary and tertiary sector labor force, less than average share of primary sector labor force,
- E: more than average share of tertiary sector labor force, less than average share of primary and secondary sector labor force, and finally
- F: more than average share of tertiary and primary sector labor force, less than average share of secondary sector labor force.

While the size and shape of the six segments depends on the center of the data, the interpretation of the segments given above is always true.

[1] Von Eynatten, H., Pawlowsky-Glahn, V., & Egozcue, J. J. (2002). Understanding perturbation on the simplex: A simple method to better visualize and interpret compositional data in ternary diagrams. Mathematical Geology, 34(3), 249–257. https://doi.org/10.1023/A:1014826205533

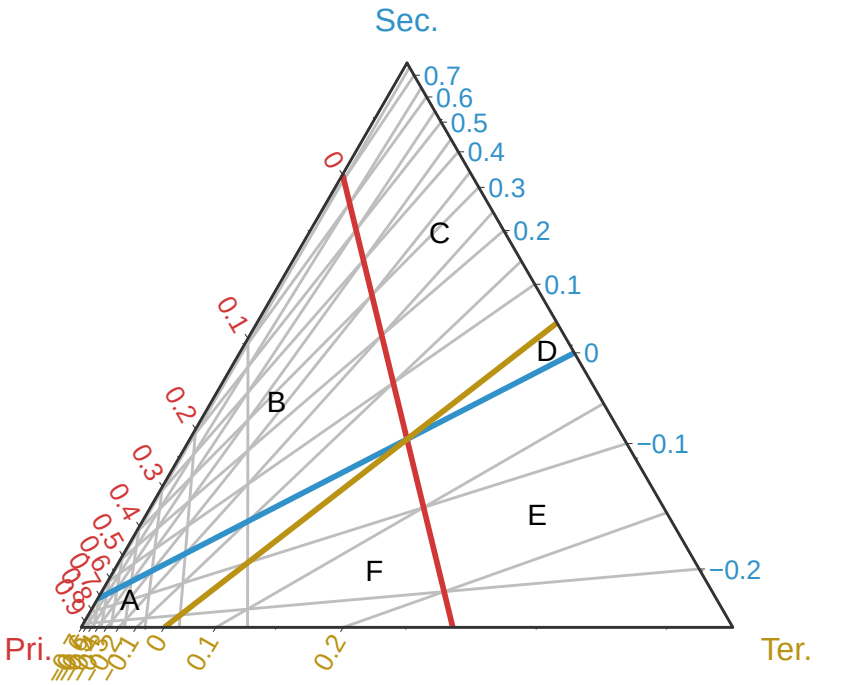


Figure 3: The center lines segment the ternary diagram into 6 regions.