

APPENDIX

REGDATA IN BRIEF

RegData uses text analysis and machine-learning algorithms to produce two novel data series. The first series counts the number of restrictions (words such as *must*, *shall*, etc.) in each part of the *Code of Federal Regulations* (CFR), and the second measures the relevance of those CFR parts to the hundreds of industries defined in the North American Industrial Classification System (NAICS). These two metrics have been combined into a single index that measures, at the national level, the degree to which each sector (two-digit NAICS code) and each industry (three- or four-digit NAICS code) are regulated in a particular year.¹ RegData has been applied in numerous research contexts, many of which are catalogued on the website RegData.org. Because RegData is a free and publicly available database, other interested parties are encouraged to download, experiment with, and apply the data in different contexts.

CALCULATION OF THE FRASE INDEX

Among the many applications of RegData, the federal regulation and state enterprise (FRASE) index considers the importance of industries in a particular state to calculate the impact of federal regulation on that state. The nature of this construction means that a state in which the largest industries are heavily regulated will tend to have a high FRASE index score.

We have calculated the FRASE index using the latest version of RegData (2.2).² The FRASE index is the ratio of the impact of federal regulations on a state's private sector to the

impact of federal regulations on the nation's private sector in a given year. A value of 1 would indicate that the state's private sector is affected by federal regulations to precisely the same degree as is the national private sector.

Calculating the FRASE index requires a few steps. First, we calculate the importance of each industry to the private sector in a particular state. To do this, we divide the value added to the state's gross domestic product (GDP) from each private-sector industry i in year t by the entire state's private-sector production in year t .³ We abbreviate contributions to the state's GDP from private-sector production as PSP (private-sector product). Because all calculations described here occur in year t , we omit time subscripts. Thus, the importance of industry i to state s is simply the fraction of the state's PSP produced by industry i :

$$(y_{s,i}/y_s) = \text{industry } i\text{'s fraction of state } s\text{'s PSP,} \quad (1)$$

where $y_{s,i}$ is the value added to state s 's PSP from industry i (observed, from Bureau of Economic Analysis) and y_s is state s 's PSP = $\sum_{i=1}^I y_{s,i}$.

Second, we calculate the importance of each industry i to the national economy. This step involves calculating the fraction of the country's PSP produced by industry i :

$$(Y_i/Y) = \text{industry } i\text{'s fraction of national PSP,} \quad (2)$$

where Y_i = the national value added to PSP from industry i = $\sum_{s=1}^S y_{s,i}$, and Y = the sum of national value added to PSP from all industries = national PSP, or $\sum_{i=1}^I Y_i$.

Third, we combine these two fractions to calculate the importance of industry i to state s relative to the industry's importance to the national economy. This relative importance of industry i to state s serves as a weighting term in later steps:

$$\frac{(y_{s,i}/y_s)}{(Y_i/Y)} = w_{s,i} = \text{importance of industry } i \text{ to state } s \text{ relative to} \\ \text{the industry's importance to the national economy} = \\ \text{weighting term.} \quad (3)$$

Next, we multiply the level of federal regulation of each industry by the weighting term for state s :

$$w_{s,i} r_i = \text{national regulation of industry } i \text{ weighted by its importance to state } s, \quad (4)$$

where r_i = regulation of industry i (observed from RegData). We then sum across all industries in the private sector in state s :

$$\sum_{i=1}^I w_{s,i} r_i = \text{industry-weighted regulation index for state } s. \quad (5)$$

Finally, we scale the industry-weighted regulation index for each state by the total level of regulation summed across all industries in the private sector in the nation:

$$FRASE_s = \frac{\sum_{i=1}^I w_{s,i} r_i}{\sum_{i=1}^I r_i}. \quad (6)$$

To account for changes in the level of national regulation, we have also produced a 1997-basis FRASE index by dividing the industry-weighted regulation index for a state in the current year by the industry-weighted regulation index for the United States overall in 1997. This index is referred to as the constant-basis FRASE index and is further explained in the introduction.

NOTES

1. For a full explanation of RegData 2.2, see Omar Ahmed Al-Ubaydli and Patrick A. McLaughlin, “RegData: A Numerical Database on Industry-Specific Regulations for all United States Industries and Federal Regulations, 1997–2012,” *Regulation and Governance*, 2015, doi:10.1111/regg.12107. See also Patrick A. McLaughlin and Oliver Sherouse, “Industry-Specific Classification of Legal Text,” working paper, forthcoming.
2. Specifically, RegData 2.2 unfiltered.
3. By examining only private-sector industries, we excluded only the industry called *government*.