factors are concentrated among the poor, not because ethnic status itself reduces turnout.

In principle, the solution to this problem of spurious correlation is simple: include all relevant variables in the analysis, for advanced statistical techniques can, to an extent, control for such problems. In practice, however, not all relevant variables will be known and spurious correlation is a continuing danger. Note, though, that even a spurious correlation may be a practical (if risky) basis for prediction.

A spurious correlation is one which arises because both factors depend on a third variable. For instance, the relationship between the proportion of immigrants in an area and its crime rate may be spurious because poverty is the real cause of crime, and both immigration and crime are high in poor areas.

The second issue in interpreting statistical results is that, even if a relationship is genuine, the direction of causation remains to be established. Take an example. Suppose we find that liberal democracies secure higher rates of economic growth than authoritarian regimes. We still face a problem of interpretation. Does the correlation arise because democracy facilitates economic growth, or because a high rate of growth fosters a stable democracy? A case can be made either way, or both; by itself a statistical correlation will not provide the answer. In itself, a correlation does not show the direction of causation.

Worthwhile quantitative comparisons can be made even when the variables take the form of categories (e.g. yes/no), rather than numerical scores. For example, are federations less likely than unitary states to develop welfare states? Is proportional representation linked to coalition government? Are non-Muslim countries more likely to be democratic than Islamic countries? Here, we are dealing with categories, rather than numerical scales: a country is either a federation or not, a government is either a coalition or not. In these circumstances, a straightforward cross-tabulation is the qualitative equivalent of the scatterplot in Figure 19.1. Correlation-like statistics can nonetheless be calculated for such tables (Pennings et al., 2006).

Historical analysis

Most studies in politics – and in comparative politics, especially – focus on the contemporary world; in the main, we leave history to the historians. But this division of labour is, of course, arbitrary; today’s present is tomorrow’s past. Political science can, and perhaps should, make more use of the past as a treasure trove of additional cases, whether of rare events such as genocide and revolution or of particular episodes that exemplify, challenge, or refine existing theories. History can enlarge our database, enabling us in particular to employ the most different design to examine the robustness of findings across distinct time periods.

Quarrying the past by political scientists in this way is rather different from the approach adopted by traditional historians. According to the German historian Leopold von Ranke (1795–1886), history ‘merely wants to show how essentially things happened’. For von Ranke (1824, p. iv), ‘a strict representation of facts, be it ever so narrow, and unpoeitical, is, beyond doubt, the first law’. From this perspective, the obligation of historians is to present the past in its own terms – as a narrative pieced