Contents

1
Introduction

Lane Kenworthy and Alexander Hicks

Macrocomparative researchers use a variety of methodological approaches. This book features analyses of a single substantive topic using several of the most common. The topic is comparative employment performance in affluent countries. The chief methodological approaches are pooled cross-section time-series regression, qualitative comparative analysis (QCA), and small-N analysis.

The aim of the volume is to illustrate in a practical fashion the advantages and drawbacks of these analytical strategies. Instruction and advice is available in numerous monographs, articles, and edited volumes (for example, Greene, 2003; Ragin, 1987, 2000; King, Keohane, and Verba, 1994; Mahoney and Rueschemeyer, 2003; Brady and Collier, 2004; George and Bennett, 2004). But often that advice is provided at a general level. Commonly, substantive illustrations are offered with reference to a single methodological approach. A key question for researchers is when and why to use one methodological approach rather than, or in addition to, another, and what are the payoffs and sacrifices entailed by a particular choice. Although general advice is helpful, the best way to understand the tradeoffs involved is via practical application.

This book was conceived partly as a follow-up to the 1994 volume *The Comparative Political Economy of the Welfare State*, edited by Thomas Janoski and Alexander Hicks. That book was aimed at comparativists interested in the welfare state and in comparative political economy more generally. It included methodological and substantive chapters covering time-series regression, pooled cross-section time-series regression, event history analysis, and qualitative comparative analysis ("Boolean analysis"). This volume differs in three main respects. First, it covers a different set of methodological approaches, focusing exclusively on those that involve macro-comparison – that is, comparison

across countries (or regions). The techniques explored in this book are pooled regression, qualitative comparative analysis, and small-N analysis. Secondly, this book includes no strictly methodological chapters; the methodological techniques are described and discussed in substantive chapters. Thirdly, the substantive chapters in the Janoski–Hicks book examined disparate issues: economic growth, wage trends, active labor market policy, and pension systems. The analyses here for the most part address the same substantive question: What are the determinants of variation in employment performance across affluent countries? Our hope is that the common substantive focus helps to reveal as clearly as possible the advantages and drawbacks of the methodological strategies.

The substantive issue: comparative employment performance

Macrocomparativists engage a wide array of substantive issues. We chose comparative employment performance as the outcome for the analyses in this book. A country's employment rate is measured as the number of people with paying jobs divided by the population age 15 to 64 (the working-age population).

Employment is a useful barometer of labor market performance in a country. For most of the past half-century, unemployment rates have been considered the main indicator of labor market outcomes, but the fact that unemployment can be hidden in various ways – low labor market participation, active labor market programs, and so on – has encouraged a shift toward employment rates.

Employment has intrinsic merit (Jahoda, 1982; Wilson, 1996; Phelps, 1997). With heightened geographical mobility, later marriage, and increased divorce, neighborhood and family ties have dissipated somewhat. As a result, work is an increasingly important site of social interaction. Employment imposes regularity and discipline on people's lives. It can be a source of mental stimulation. It helps to fulfill the widespread desire to contribute to, and be integrated with, the larger society. For many individuals, work is inextricably bound up with identity and self-esteem.

In addition, an increasingly common view is that high employment is critical to maintenance of low or moderate levels of income inequality (Esping-Andersen, 1999; Ferrera et al., 2000; Scharpf and Schmidt, 2000; Esping-Andersen et al., 2002; Kok et al., 2003; Kenworthy, 2004, 2008; OECD, 2005, 2006). Meeting pension and health care commitments for an ageing population will require greater government funds
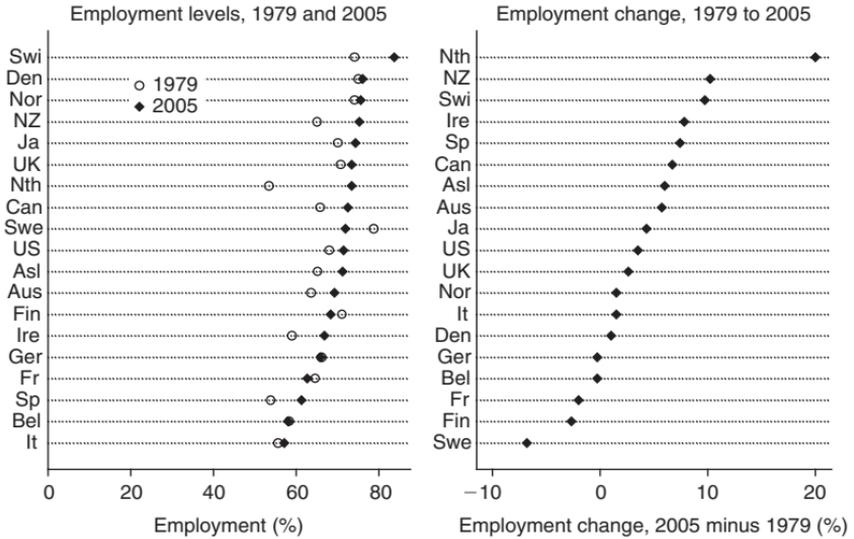
Figure 1.1 Employment levels and employment change, 1979 to 2005
Note: Employment = employed persons as a share of the population age 15 to 64. Portugal is omitted due to lack of employment data for 1979.

in coming decades. Yet governments increasingly find it difficult to raise tax rates, due to capital mobility. This makes it difficult to maintain generous transfers for the working-age population and their children. A rising employment rate helps to increase tax revenues without raising tax rates. And by bringing former benefit recipients into the workforce, it reduces government transfer payments.

Figure 1.1 shows employment rates in 1979 and 2005, and changes in employment during that period, for the group of affluent countries examined in this book: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. As of the mid-2000s, employment rates ranged from less than 60 percent of the working-age population in Italy to more than 80 percent in Switzerland. The variation in change between the late 1970s and the mid-2000s was equally large, with the employment rate falling by more than 5 percentage points in Sweden and increasing by 20 points in the Netherlands.

How can countries achieve a high and/or rising employment rate? There are two principal debates around this issue. The first concerns the determinants of overall employment performance and focuses on

the impact of labor market institutions and policies. These include wage levels at the low end of the distribution, employment protection regulations, government benefits, and taxation. If wage levels at the low end of the labor market are high relative to productivity levels – due to a statutory minimum wage, a collectively bargained minimum, or a tight labor market – employer demand for low-end workers may diminish, reducing the employment rate. If it is difficult for employers to fire workers when the economy is bad or the firm's sales are slumping – because they must get the approval of a works council or provide a generous severance package or pay for extensive retraining and job placement – employers may hire fewer workers when times are good. If government benefits – social assistance, unemployment insurance, sickness or disability compensation, pensions, and so on – are fairly generous, easy to access, and of lengthy duration, workers at the low end of the job market may be less likely to search for and accept employment. A high tax burden can reduce the net benefit to a worker from employment and/or increase the cost to an employer of hiring, thereby potentially producing less supply of and demand for labor.

High low-end wages, strict employment protection regulations, generous government benefits, and high taxes are sometimes referred to as "labor market rigidities." The notion that such rigidities impede high and/or rising employment has been around for a long time, but it has been especially prominent since publication of *The OECD Jobs Study* in 1994. The *Jobs Study* was a clear and systematic statement of the rigidities ➜ poor employment performance hypothesis, and it was published at a time when unemployment in a number of western European countries had been high for roughly a decade and showed no signs of imminent decline. Since 1994 dozens of comparative empirical studies have examined the hypothesis. (Some recent studies, which include citations to earlier ones, include Blau and Kahn, 2002; Kenworthy 2004, 2008; Howell, 2005; Nickell, Nunziata, and Ochel, 2005; Bassanini and Duval, 2006; OECD, 2006; Baccaro and Rei, 2007.) Despite this extensive research, there is nothing close to a consensus regarding the merit of the hypothesis.

The second debate is about the impact of so-called family policies (also variously referred to as work–family reconciliation policies and women-friendly policies) on female employment. In countries with employment deficits, the problem consists chiefly of a shortage of women's employment. This can be seen clearly in Figure 1.2, which shows employment rates for men and women as of 2005. A critical task – perhaps *the* critical task – for low-employment countries, therefore, is to identify and
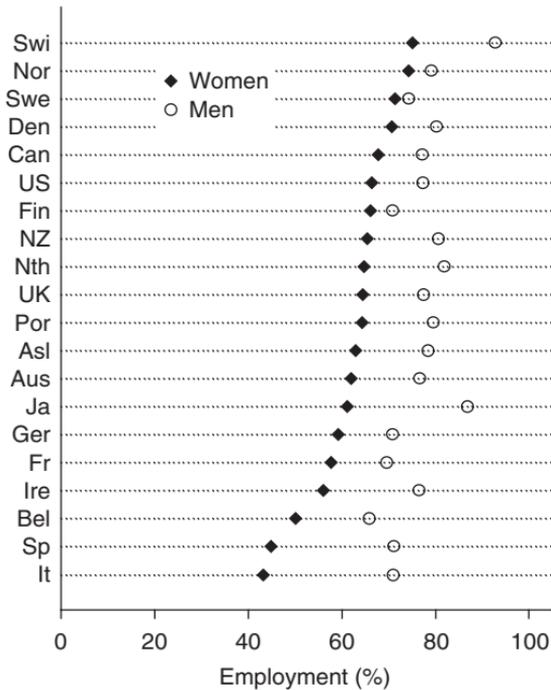
Figure 1.2 Women's and men's employment levels, 2005
Note: Employment = female or male employed persons as a share of the female or male population age 15 to 64.

implement institutional or policy changes that can substantially increase female employment. A number of studies have suggested that the key is generous family policies (Winegarden and Bracy, 1995; Ruhm, 1998; Meyers, Gornick, and Ross, 1999; Plantenga and Hansen, 1999; Rubery, Smith, and Fagan, 1999; Sainsbury, 1999; Daly, 2000; Korpi, 2000; Dingeldey, 2002; OECD, 2001; Stier, Lewin-Epstein, and Braun, 2001; Esping-Andersen et al., 2002, ch. 3; Orloff, 2002; Pettit and Hook, 2002; Ferrarini, 2003; Gornick and Meyers, 2003; Jaumotte, 2003; Morgan and Zippel, 2003; Mandel and Semyonov, 2006; Kenworthy, 2008).

One such policy is public provision or financing of child care. Lack of affordable child care can pose a significant obstacle to employment for women with preschool-age children. A second is paid parental/care leave. The expectation is that if women know they can take a reasonably long break from work without losing their job and without foregoing all of their earnings, more will choose to enter the labor market in the first place and more will return after having a child. A third is government

provision of public sector jobs, which may be more attractive to women than private sector jobs because they are more likely to be available at reduced hours (part-time), to be secure (governments are less likely than private employers to fire employees during economic downturns), and to accommodate family needs such as illness. A fourth is promotion of or support for part-time employment, which women may prefer because the shorter hours facilitate work–family balance. A fifth is the structure of the tax system. Of particular relevance is the degree to which a couple with two earners is penalized relative to a couple with one earner; the greater the tax penalty, the stronger the disincentive for a woman with an employed husband to get a job. A sixth is anti-discrimination and affirmative action laws. To the extent that women's employment is impeded by discriminatory action by employers, such policies are likely to help.

The book's chapters focus on these two substantive questions: What has been the impact of labor market institutions and policies on overall employment performance? What has been the effect of family policies on women's employment?

Methodological approaches

In attempting to answer a question such as what determines employment performance, various analytical strategies can be pursued. One is to examine individual behavior. Another is to consider patterns across firms or industries within countries. A third is to look at developments over time in a single country. A fourth approach is macrocomparative. Countries are the unit of analysis. The causal factors of interest (policies and institutions) and the outcomes are measured at the level of the nation. Analytical leverage is gained at least in part, and often primarily, by comparison across countries.

Macrocomparative analysis can be conducted using a variety of techniques. In this book the focus is on three: regression, qualitative comparative analysis, and small-N analysis. In this introductory chapter we offer a brief outline of the main distinguishing features, advantages, and disadvantages of these three techniques. They are summarized in Table 1.1.

A key point we wish to stress at the outset is that these approaches are best viewed as complementary rather than competing and overlapping rather than mutually exclusive (Lieberman, 2005; Ragin, 2005). Each is capable of contributing to macro-level analysis in different ways.[1] Overlaps between the three approaches are possible. For example, QCA is not

Table 1.1   Analytical strengths of the three methodological approaches

Pooled Regression
   Designed to assess tendential relationships, manifested in patterns of
      co-variation among variables
   Useful for assessing the net effect of a variable on the outcome
   Enables variation across units (countries) and over time to be analyzed
      together
   Allows formal estimation of the magnitude of impact of a cause

Qualitative Comparative Analysis
   Designed to assess deterministic causal relationships (logically conceived):
      sufficiency and necessity
   Useful for exploring causal configurations (combinations of causes)
   Useful for examining multiple causal paths to the same outcome
   Allows formal estimation of the magnitude of impact of a cause

Small-N Analysis
   Useful for assessment of causal mechanisms via process tracing
   Useful for elimination of hypothesized ("always") sufficient or
      necessary conditions
   Ordinal cross-country comparison can be used to assess hypothesized
      tendential or quasi-deterministic relationships, but generalization
      beyond the studied cases is problematic
   Possibility of considering variables that cannot be included in a large-N
      analysis because data are available for only a few countries
   Possibility of better measurement of variables due to case knowledge
   Possibility of more nuanced attention to interaction among causal factors
      than is possible with regression

uncommonly used for small Ns of 10 or fewer, and an analysis that treats
a few cross-sectionally differentiated units in the "small-N" style might
include a time-series regression of years encompassed by a cross-sectional
slice or QCA analysis of subunits (e.g., of states).

Pooled regression

Regression is the most commonly used analytical technique in macro-
comparative analysis. It is a correlational technique, although causal
interpretation of regression slope estimates backed by special statistical
care (for example, fastidiously specified lag structures) as well as the-
oretical argument sometimes accompany regression analyses. The aim
is to identify statistical associations between hypothesized causes and
outcomes. Such associations are based on co-variation. In regression, an
independent variable is associated with a dependent variable when levels
of the two variables correspond to one another.

The relationships assessed in regression are "tendential" rather than "deterministic." A particular level of a hypothesized cause is thought to make it more likely that the outcome will have a certain level, but it is assumed that there may be exceptions. Regression predictions are made with allowance for error (for example, residuals, confidence intervals).

For analysts interested in understanding differences across the rich countries, a fundamental analytical obstacle is the small number of cases. Depending upon one's definition of "affluent," there are approximately 15 to 25 nations to study. This inhibits estimation of regressions with more than a few independent variables. Another limitation of standard regression in macrocomparative analysis is that it frequently is confined to cross-country variation, ignoring variation over time within countries. Time-series regression does the opposite.

A pooled cross-section time-series regression combines information about variation across countries with information about variation over time within countries. The unit of analysis is the country-year or country-period, rather than either the country or the time period. This not only combines the two kinds of variation; it also substantially increases the number of observations, thereby helping to address the small-N problem. For these reasons, pooled regression has, as Janoski and Hicks (1994) envisioned, become the dominant technique for large-N analysis in macrocomparative research.

Yet pooled regression has important limitations, to which practitioners do not always pay sufficient heed. One is that the determinants of cross-country variation may not correspond to those of over-time variation within countries (Griffin et al., 1986; Kittel, 1999; Kenworthy, 2006, 2007; Shalev, 2007). Over a very long period, we would expect such correspondence, but most analysts do not have lengthy time-series data that are comparable across more than a few countries. One way to partially address this is to pool periods of years rather than individual years (Hicks and Kenworthy, 1998, 2003, this volume; Barro, 2000).

A second limitation concerns time lags in causal effects. Many hypotheses about determinants of change in macrocomparative analysis either implicitly or explicitly refer to relatively long-term effects. Yet most pooled regression analyses use the country-year as the unit. This is likely a function of researchers' desire to significantly boost the number of observations and thereby facilitate estimation of models with a large number of regressors. Sometimes analyses with annual data can pick up medium-term or long-term effects, but that hinges on getting the lag structure correct. More often than not, using annual data to examine

hypothesized medium-run or long-run associations will obscure rather than clarify.

Finally, successful estimation of pooled regressions requires meeting, or addressing, various technical requirements. Bernhard Kittel discusses these issues extensively in chapter 2, so we will do no more than mention them here (see also Kittel, 1999; Ebbinghaus, 2005; Wilson and Butler, 2007). They include independence of observations, non-trending (stationarity), and assessment of the cross-sectional or longitudinal dominance of the data array (the proportions of the variable variances accounted for by the longitudinal and cross-sectional dimensions).

Qualitative comparative analysis (QCA)

Qualitative comparative analysis is a technique for systematically exploring relations between explanatory factors and outcomes (Ragin, 1987, 2000; Ragin and Rihoux, 2004). QCA was conceived originally with the aim of formalizing the analytical process often pursued by small-N qualitative researchers, to enable the process to be applied more systematically and to a larger number of cases. This process is articulated in terms of the logical language of set theory.

There are two variants of QCA. One, crisp-set QCA (Boolean), uses dichotomous codings of causal conditions and outcomes. The other, fuzzy-set QCA, uses pseudo-continuous codings that vary between zero and one. In both versions, the aim is to identify hypothesized causal factors, or combinations of those factors, that are related to the outcome in a pattern consistent with that of a sufficient or necessary condition – the two main types of deterministic causal relationship. When plotted in a scattergram, a relationship of sufficiency is suggested if the data points all fall above and to the left of a 45-degree line running from the lower-left corner to the upper-right corner. When the hypothesized cause is absent or low, the outcome may be absent/low or present/high; but when the cause is present/high, the outcome is always present/high. A relationship of necessity is indicated by the data points being located below and to the right of the 45-degree line. When the cause is present/high, the outcome may be absent/low or present/high. But the outcome is never present/high unless the cause is present/high. These two patterns are illustrated in Figure 1.3.

"Cause" is the preferred term for sufficient and necessary "conditions" in the QCA literature. Although some have cautioned against identification of the logical conditions of formal languages such as QCA's Boolean logic with real world causes (Ayer, 1956, pp. 170–5; Passmore, 1967, pp. 355–60; Manicas, 2006), we adopt the QCA usage here.
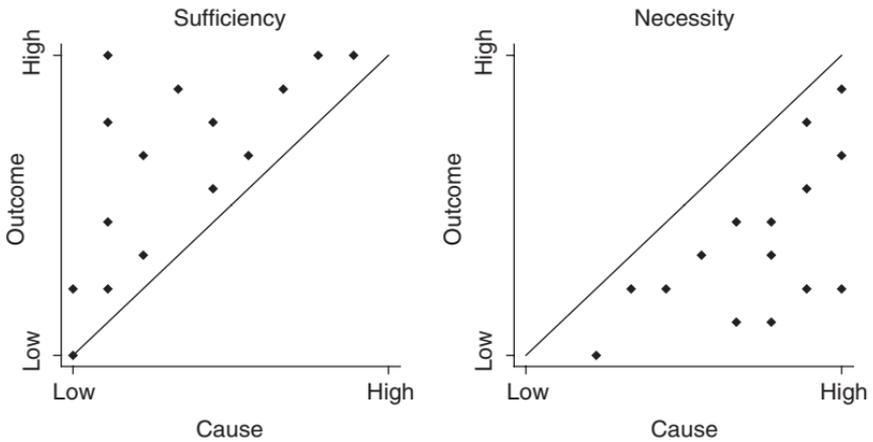
Figure 1.3   Illustration of causal sufficiency and necessity
Note: The lines in the charts are 45-degree lines, not regression lines.

Charles Ragin and others have suggested that a deterministic relationship may have empirical exceptions, due to the randomness of social processes, measurement error, or other reasons. A cause need not be *always* sufficient or necessary; it may be "nearly always" or "usually" sufficient or necessary (Ragin, 1987, ch. 7; Ragin, 2000, pp. 107–16; see also Goertz and Starr, 2003). For some, the idea of introducing a tendential element into an otherwise deterministic notion of causality is oxymoronic. Why not instead simply refer to the relationship as tendential? Consider the pattern in Figure 1.4. Except for one data point, it is consistent with a hypothesis of sufficiency: in cases where the cause is present/high, the outcome is also present/high. One could conceptualize the relationship as tendential; the two variables correlate at .54. But "almost always sufficient" seems a more accurate description of the empirical pattern. Which interpretation is more sensible depends heavily on what the substantive issue is and whether a tendential or deterministic understanding is more compelling on theoretical grounds.

For macrocomparative analysis, QCA has several potential advantages relative to regression. One is its focus on deterministic relationships. As noted earlier, regression is designed to assess tendential causal (or, more simply, explanatory) relationships. When a cause is hypothesized to be sufficient or necessary for an outcome, QCA may thus be a more appropriate method.

Secondly, QCA is adept at exploring causal configurations – situations in which variables have an impact only in combination with a high or
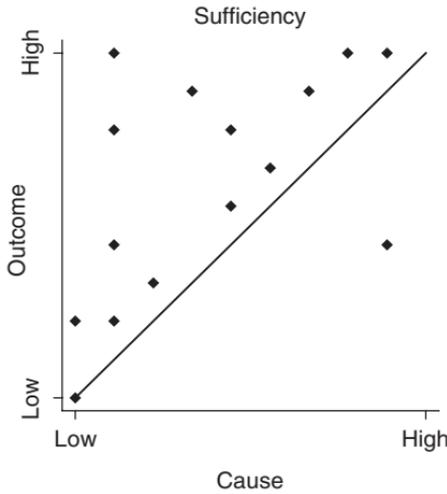
Figure 1.4   Illustration of "nearly always sufficient"
Note: The line in the chart is a 45-degree line.

low degree of one or more other factors. In regression analysis, causal configurations are assessed via interaction terms. But with the small or moderate number of cases that is common to macrocomparative analysis of affluent countries, the number of interactions terms that can be included in a regression model tends to be limited. Pooling the cross-country and over-time variation can alleviate this problem, but the collinearity produced by interaction terms that involve more than two or three variables and the difficulty in interpreting the results makes modeling complex interactions problematic. Moreover, while assessing interactions in regression requires that variables demonstrate product terms, QCA treats any aspects of cases that appear together systematically – in any quantity – as potentially interdependent.

Thirdly, QCA facilitates identification of multiple pathways to an outcome. Many social phenomena have causes that are relevant to only a fraction of the cases. With a correlational technique such as regression, when the dependent variable is high/present but the independent variable is low/absent, this weakens the estimate of the effect of the independent variable (or interacted combination of independent variables). QCA, by contrast, is designed to reveal patterns of association that differ across subsets of cases. It thereby enables discovery of more complex causal patterns than are generally recognizable via regression.

In the view of some, an additional advantage of QCA is that the deterministic relationships it identifies are more likely to be causal than the

tendential associations identified by regression (Mahoney and Goertz, 2006). This view, however, has been contested and criticized – by logical positivists such as Ayer (1956, pp. 170–5) as well as by realists like Manicas (2006) and interpretivists like Wittgenstein (Passmore, 1967, pp. 355–60) – for a conflation of logical and causal relations (but see Pruss, 2006).

Like any technique, qualitative comparative analysis has limitations. First, if a causal factor is suspected to have a tendential relationship with the outcomes, rather than a deterministic one, QCA is of little use.

Secondly, a QCA analysis is, in a key respect, bivariate. Multiple causal factors are considered, and the "solution set" of causal configurations yielded by a QCA analysis will vary with the set of causal factors entered into the analysis. Specification of additional explanatory conditions, like entry of new control variables in regression, will often modify analytical results. Yet QCA examines the relationship between the outcome and each single hypothesized cause or combination of causes without controlling for – that is, without "holding constant" – any other causes. As Aaron Katz, Matthias vom Hau, and James Mahoney (2005, p. 568; see also Seawright, 2005) point out:

> Fuzzy-set analysis is only a multivariate method in the sense that the technique can explore if combinations of variables represent sufficient causes. However, since each combination is reduced to a single value, each combination is, in effect, treated as a single cause.

Jason Seawright (2002, p. 181) has argued that given the assumption that the causal relationship is a deterministic one, this is appropriate:

> . . . claims of necessary and/or sufficient causation are fundamentally bivariate in nature. The hypothesis entailed in this idea of causation is that no other variable or combination of variables can overcome the effects of the necessary and/or sufficient cause. Therefore, controlling for other variables cannot alter the conclusion of the bivariate analysis, and a bivariate focus is fully appropriate.

However, positing that a causal relationship is deterministic and finding a bivariate pattern consistent with that hypothesis does not rule out the possibility of spuriousness (omitted variable bias). Although regression is scarcely immune to this concern, it is designed to estimate the "net" effect of each variable. Charles Ragin (2005, p. 35) sums up this point effectively: "Regression analysis is a preeminent tool for estimating

net effects; QCA's primary analytic focus is on the different ways causal conditions combine."

Thirdly, to this point there is no QCA counterpart to the combining of cross-country and over-time variation that is possible with pooled regression.

Regression has several tools for assessing the magnitude of a variable's causal effect. One is the variable's coefficient, which estimates the change in the dependent variable given a one-unit increase in an independent variable (net of other independent variables). Another is the R^2 (coefficient of determination), which measures the precision or "goodness of fit" of the coefficient estimates. Others include multiple partial coefficients of determination for subsets of regressors, standardized coefficients, and various techniques of exogeneity and Granger causality assessment. In QCA the principal tools for assessing the strength of relationships are "consistency" and "coverage." Consistency refers to the degree to which the empirical pattern corresponds to that of sufficiency or necessity. An "always" sufficient relationship can be considered stronger than a "nearly always" or "usually" sufficient one. Coverage refers to the share of cases having a particular outcome that feature a particular causal factor or causal configuration. Consistency and coverage are discussed in greater detail in chapter 3 of this volume and in Ragin (2006).

Small-N analysis

By "small-N analysis" we refer to macrocomparative analyses in which the number of countries (cases) studied is ten or fewer. This is, of course, an arbitrary cutoff; there is no number of nations that objectively demarcates "small" from "large." The most common number of nations studied in small-N analyses is one, two, or three.[2]

One of the most important contributions of small-N analysis is descriptive. Studying a small number of countries allows the researcher to learn, and convey to the reader, a level of detailed knowledge that is beyond the reach of an analyst committed to comparing a large number of countries. The small-N researcher typically examines variables, events, actors, and other aspects of the national context in extensive detail. This type of information is inherently interesting to comparativists. It also helps large-N researchers to check their coding of variables, to consider additional causal factors, to think about interactions among causal conditions, and to judge the general plausibility of their causal hypotheses.

What can small-N analyses contribute in terms of theory testing (hypothesis testing)? One contribution is *process tracing* (Mahoney, 2000, 2003; Hall, 2003; George and Bennett, 2004). Process tracing consists of examining theoretically-specified causal pathways (causal mechanisms) in the context of developments in a single country. This is particularly useful where a large-N analysis has suggested support for a particular causal story but where, often due to lack of data, the large-N analysis is unable to examine the causal mechanisms. Of course, any finding of an association or "solution set" in a large-N analysis should be supported by a plausible theoretical story. But investigating the causal paths empirically is equally critical, and small-N analysis is a useful way to do that. When the small-N analysis is of a single country, it is often referred to as a "case study." Small-N analyses of multiple countries sometimes are actually multiple case studies, rather than cross-country comparisons.

A good example of a case study of employment performance is Jelle Visser and Anton Hemerijck's (1997) book-length analysis of developments in the Netherlands, which was revisited in an article by Visser (2002). These two works consider a number of possible determinants of the Dutch "employment miracle" since the early 1980s (see Figure 1.1 above). Visser and Hemerijck carefully trace developments in public policy, economic institutions, and employment patterns in the 1980s and 1990s. They conclude that wage restraint and increased female educational attainment played critical roles, but also that a variety of conjunctural factors, such a growing preference for part-time workers among public sector employers and reactions by employers' associations to an early-1980s agreement to reduce the standard work week, were important. Family policy and union strategies are found to have played distinctly secondary roles.

Small-N analyses of more than one country sometimes are simply multiple case studies. Frequently, however, they are comparative: they attempt to gain analytical leverage from cross-country comparison. How so? In an insightful paper, James Mahoney (2000, pp. 399–406; also Mahoney, 2003) points out that small-N analyses often engage in a simple form of correlational analysis, which he refers to as *ordinal comparison*. The countries are rank-ordered on the outcome and on a hypothesized causal variable, and the analyst draws inferences about causal impact based on the consistency of the rankings for the two variables. Suppose a researcher compares employment performance in Sweden, Italy, and the United States during the 1990s. The analyst might argue that performance was strongest in the United States, followed by Sweden and then Italy, and that non-market institutions and policies were weakest in the

United States and strongest in Italy, with Sweden in between. Among these three countries, the ranked positions on the two variables are consistent, which supports the notion that "rigidities" had an adverse effect on employment outcomes. The analysis is depicted in Figure 1.5.

The strength of such an analysis is likely to lie in the coding of the variables and perhaps the attention to interactions among them. However, this type of implicit correlational analysis is problematic with a small number of cases. It is difficult to take very many factors into account (to control for them). And while an inference based on well-done analysis of this type can certainly be suggestive of a tendential causal relationship, it leaves open the possibility that the countries analyzed are atypical. Of course, large-N correlational analyses are never definitive; they too can do no more than suggest a causal relationship. Still, all else equal, the larger the number of countries analyzed, the less reason there is to worry about representativeness or generality of findings (Geddes, 1990; King, Keohane, and Verba, 1994).

Mahoney (2000, pp. 391–8) argues that small-N comparison offers analytical leverage chiefly via its ability to *eliminate a hypothesized sufficient or necessary condition*. A hypothesized sufficient condition, he suggests, can be tested using John Stuart Mill's "method of difference." Here cases are selected that differ on the outcome; in at least one country the outcome is present and in another it is absent. Since "sufficiency" implies that where the cause is present the outcome will occur, any cause that is present in the country in which the outcome is absent can be ruled out as a sufficient condition. A hypothesized necessary condition can
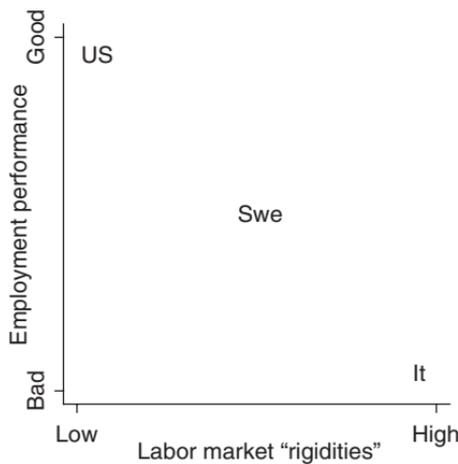

Figure 1.5 Illustration of small-N ordinal comparison

be assessed using Mill's "method of agreement," in which the outcome is present in all countries selected for analysis. Any cause that is absent in any of the countries can be ruled out as a candidate for "necessity." Findings in these types of analyses that are consistent with a hypothesis of sufficiency or necessity can be treated as supportive, but by no means definitive, since there may be other nations – possibly many of them – for which the hypothesis is contradicted. Mahoney's point is illustrated in Figure 1.6.

We find Mahoney's observation illuminating. However, a causal factor can be eliminated as sufficient or necessary based on a single case; there is no need for cross-country comparison. Indeed, if the aim is to eliminate a hypothesized sufficient or necessary condition, comparison across countries is irrelevant. Comparison implies variation, and in testing a hypothesized deterministic relationship there is no need for variation. To test a hypothesized sufficient condition, one should select only cases in which the condition is present/high; countries in which it is absent/low offer no analytical leverage. To test a hypothesized necessary condition, a researcher should select only cases in which the outcome is present/high. For the purpose of eliminating a hypothesized sufficient or necessary condition, then, the utility of studying more than one country is not that it enables comparison across the countries. Rather, the gain is simply that examining multiple countries increases the opportunity for elimination.

Given this, does a small-N comparative analysis have any advantage over a large-N analysis for eliminating a hypothesized sufficient or necessary condition? After all, the larger the number of countries, the greater the opportunity for elimination – and the greater the confidence in the hypothesis if it is not eliminated. There are two potential advantages to keeping the N small. One is that the likelihood of measurement error may be reduced because of the researcher's greater case knowledge. Another is that the researcher may be able to consider variables for which data are not available for a larger set of countries.

A. Eliminating a hypothesized sufficient condition			B. Eliminating a hypothesized necessary condition		
	Cause	Outcome		Cause	Outcome
Country 1	Present	Present	Country 1	Present	Present
Country 2	Present	Absent	Country 2	Absent	Present

Figure 1.6 Mahoney's argument for use of small-N analysis to eliminate a hypothesized sufficient or necessary condition

For a tendential or quasi-deterministic theory, small-N analysis is more limited in its analytical utility. A single nonconforming case, and perhaps even two or three, cannot eliminate a hypothesized tendential, quasi-sufficient, or quasi-necessary cause. Then again, a nonconforming case increase grounds for skepticism. This is particularly true if the case is a "most favorable" one – one that there is reason to suspect will be highly likely to support the theory, or one that proponents of the theory frequently refer to as an illustration (Eckstein, 1975).

A recent small-N macrocomparative analysis of employment performance is Ronald Schettkat's (2005) study of Germany and the Netherlands. Schettkat provides qualitative and quantitative information to suggest that these two countries have been similar in their degree of institutional and policy "rigidities." According to the tradeoff hypothesis, therefore, both should have had poor employment performance during the 1980s and 1990s. Yet the Netherlands arguably had very good employment performance. Schettkat provides extensive detail to support this coding decision.

Schettkat does not say whether he considers the hypothesis he is assessing to be deterministic, quasi-deterministic, or tendential – that is, whether labor market "rigidities" are hypothesized to be a sufficient condition for bad employment performance or to increase the likelihood of bad employment performance. If the hypothesis is one of sufficiency, the Dutch case contradicts it. This is shown in Figure 1.7. But note that the German case (like country 1 in Figure 1.6), and therefore the cross-country comparison, is not needed to reach this conclusion. If Schettkat is treating the tradeoff view as a tendential hypothesis, then the cross-country comparison is helpful. But here the fact that the analysis includes only two countries is problematic. After all, the Netherlands could simply be an exception to the general tendency.

Like pooled regression and QCA, then, small-N analysis offers certain advantages but also has important limitations. Its main assets are descriptive detail, care in measurement, ability to consider causal

	Hypothesized sufficient cause: labor market "rigidities"	Outcome: bad employment performance
Germany	Present	Present
Netherlands	Present	Absent

Figure 1.7 Summary of analysis in Schettkat (2005)

variables for which data may not be available for a larger set of countries, process tracing as a means of assessing causal mechanisms, and elimination of hypothesized sufficient or necessary conditions. Its principal drawbacks are limited generalizability and concern about omitted variable bias.

Overview of the chapters

The book's chapters are organized partly by substantive topic and partly by methodological approach. Table 1.2 provides a summary.

Chapters 2 and 3 examine the rigidities ➔ poor employment performance hypothesis. These two chapters use the same data set. The employment data are for private sector employment in low-end services – hotels, restaurants, wholesale and retail trade, and community/social/personal services (ISIC revision 2, sectors 6 and 9). These data are available only through the mid-1990s, but they provide the truest test of the rigidities hypothesis (see Iversen and Wren, 1998; Kenworthy, 2004, ch. 5). The analyses focus on six labor market institutions and policies: earnings inequality (a proxy for low-end wage levels), wage increases, payroll and consumption taxes, employment protection regulations, unemployment benefit generosity, and public employment.

In chapter 2, Bernhard Kittel offers a clear and careful illustration of some of the potential pitfalls of pooled regression. Perhaps most important, he finds that the choice to pool observations adds little information and introduces significant estimation problems. Adding annual observations does not add relevant variation to several key explanatory variables, as they are largely constant over time. And the dependent variable turns out to be nonstationary. This leads Kittel to prefer a small-N cross-sectional design over a moderate-N pooled one. The lesson is not that simple cross-sectional models are always or even usually preferable, but rather that analysts should use pooled regression where doing so makes theoretical and empirical sense, not simply because data availability makes it possible to do so. Kittel's substantive conclusion is that generous unemployment benefits may reduce the employment growth of private consumer-service jobs, at least when employment protection regulations are not stringent and when benefit generosity is measured using gross (pretax) rather than net benefits.

In chapter 3, Jessica Epstein, Daniel Duerr, Lane Kenworthy, and Charles Ragin use fuzzy-set qualitative comparative analysis to explore the impact of the same institutions and policies on growth of employment in private sector consumer-oriented services. They focus on paths

Table 1.2 Summary of the contributions

Chapter	Author(s)	Dependent variable(s)	Unit(s) of analysis	Countries and years	Method
2	Kittel	Level and change in employment in private-sector consumer-oriented services	Country-years and countries	14 countries, 1979 to 1995	Regression: pooled and cross-sectional
3	Epstein, Duerr, Kenworthy, and Ragin	Change in employment in private-sector consumer-oriented services	Countries	14 countries, 1979 to 1995	Fuzzy-set QCA
4	Misra and Jude	Level and change in women's total, full-time, and part-time employment	Countries	France and the Netherlands, 1960s through 1990s	Small-N analysis: within-country process tracing and cross-country ordinal comparison
5	Eliason, Stryker, and Tranby	Level of women's labor force participation	Country-years	14 countries, 1960 to 1999	Fuzzy-set QCA with compliers average causal effects (CACE) analysis
6	Hicks and Kenworthy	Level and change in women's employment	Country-decades	14 countries, 1960s through 1990s	Regression: pooled and cross-sectional
7	Visser and Yerkes	Employment status and employment transitions among women: full-time, long-hour part-time, short-hour part-time, non-employed	Individuals	Germany, the Netherlands, and the United Kingdom, 1992 to 2002	Regression: multinomial logit
8	Ebbinghaus	Early-exit rates and regimes among men and women age 55–59 and 60–64	Early-exit regimes and countries	10 countries, 1970 to 2003	Small-N analysis: cross-regime ordinal comparison and within-regime process tracing
9	Przeworski	Change in labor force participation	Country-years	135 countries, 1950 to 1990	Regression (pooled) with selection bias estimators

to slow employment growth or employment loss ("poor employment performance"). One of the aims of the chapter is to carefully illustrate the mechanics of a fuzzy-set QCA analysis, which is far less commonly used than regression in macrocomparative research. The chapter illustrates the potential usefulness of QCA in situations where researchers want to explore hypotheses of sufficiency and/or necessity and where the interest is primarily cross-sectional. Although QCA is particularly adept at examining multiple causal paths to the same outcome and at considering combinations of causal factors, in this particular analysis it turns out that there are only two causal paths consistent with a sufficiency hypothesis and only one of them involves multiple causal factors. The results center on one simple causal configuration and another singular causal factor: (1) low earnings inequality combined with high payroll and consumption taxes; (2) high unemployment benefit generosity.

Chapters 4–7 shift the focus to women's labor force participation and employment, with an emphasis on the impact of family policies.

In chapter 4, Joya Misra and Lucian Jude examine the effect of family policy on women's employment in a small-N analysis of France and the Netherlands. Part of their aim is to carefully trace over-time developments in these two countries and thereby explore in a detailed and nuanced fashion the role of family policy, economic conditions, and cultural support. They also are interested in understanding two differences between these countries: (1) higher full-time female employment in France by the 1960s and 1970s; (2) dramatic growth in (mainly part-time) women's employment in the Netherlands beginning in the 1980s versus stagnation in France. Based on their analysis, they argue that a combination of supportive family policy, greater economic need for women's employment, and cultural support explains both the initial higher levels of women's employment in France as well as the dramatic growth of women's employment in the Netherlands. However, cultural and policy differences in respect of caregiving for young children remain, helping explain the much higher levels of part-time employment among Dutch women.

One of the purposes of this volume is to highlight the advantages and disadvantages of alternative methodological techniques. Equally important, however, is to move beyond these discussions to emphasize payoffs resulting from combining multiple methods. In chapter 5, Scott Eliason, Robin Stryker, and Eric Tranby combine fuzzy-set QCA methods with an analysis of compliers average causal effects (CACE) to explore the impact of left government on family policies and of family policies on female labor force participation from the 1960s through the 1990s. The principal

family policies they consider are child care, maternity leave, public employment, and child benefits. They offer a methodological innovation in coupling the QCA analysis with a CACE analysis. The aim of the latter is to assess whether the effect resembles one that would have been observed of compliers had the treatment been randomly assigned, as would be the case in a typical experimental design framework.

They conclude that both demand-side and supply-side factors causally influence female labor force participation. On the demand side, they find that an expanded public sector has a substantial impact on female labor force participation. On the supply side, they find that maternity leave and public day care programs also have non-negligible causal effects on female labor force participation, although in some cases modest in comparison to demand-side mechanisms.

In chapter 6, we (Hicks and Kenworthy) use Eliason, Stryker, and Tranby's family policy data to explore the relationship between family policy generosity and female employment via regression analysis. We argue that even if the effect of family policy generosity on women's employment is conceptualized in a deterministic fashion – as a sufficient condition – there may be reason for concern about omitted variable bias. We examine the possibility that the association between family policy and female employment is spurious – a product of the fact that both are associated with women's educational attainment.

We examine unconstrained pooled models, pooled models with fixed effects for time or country, and a cross-sectional model. The pooled models without unit effects for countries suggest that both family policy and women's educational attainment have tended to boost women's employment rates. In pooled models with country unit effects, however, there is little or no indication of a family policy impact. In cross-sectional models, we once again find support for effects of both family policies and female education. Further exploration of the over-time developments within countries confirms that support for the hypothesis that generous family policies tend to increase female employment rates rests largely on the cross-sectional association. This does not mean family policies do not affect women's employment, but it suggests less confidence than if there were supportive evidence both across countries *and* within countries over time.

In chapter 7, Jelle Visser and Mara Yerkes pursue a very promising analytical strategy in macrocomparative research: the use of individual-level panel data in a small, but deliberately chosen set of countries. They use individual-level data to examine the effects of institutions, policies, and women's preferences regarding employment and working hours in

Germany, the United Kingdom, and the Netherlands. These three countries exhibit the highest incidence of women working part-time in the OECD. They also share a "breadwinner legacy": each traditionally discouraged employment of mothers with young children, though they have in various ways moved away from that position in recent decades. Visser and Yerkes exploit the variation in welfare state and industrial relations institutions and policies across the three countries and the availability of individual panel data to explore how strongly the breadwinner legacies still affect the choice for and nature of part-time work of women.

Visser and Yerkes first estimate the effects of motherhood on the probability of adult women to be full-time employed, part-time employed, or entirely outside the labor force ("inactive"). They focus on differences across birth cohorts, controlling for education and household status. Next they analyze transitions from part-time employment into inactivity or full-time employment, focusing on the impact of motherhood. Comparison with the transitions from full-time jobs into long-hour or short-hour part-time jobs or into inactivity can help answer the question of whether part-time employment encourages particular groups of women to remain in the labor force. Finally, they examine the impact of "choice," bringing into play working-time preferences of women and analyzing whether or not they lead to transitions in the desired direction.

In chapter 8, Bernhard Ebbinghaus analyzes comparative patterns of early exit from the labor market. Affluent countries differ sharply in their employment rates among those aged 55 to 64 and in the degree to which those rates have shifted over the past several decades. Ebbinghaus examines the impact of "early-exit regimes" (see also Ebbinghaus, 2006). These regimes are defined by three factors: social policy orientation (protection systems, which are hypothesized to differ in the degree to which they pull older workers out of the labor force), the organization of production (production systems, which are hypothesized to differ in the degree to which they push older workers out of employment), and the organization of labor relations (partnership traditions, which are hypothesized to differentially mediate pull and push factors). Ebbinghaus identifies five regimes. He selects ten countries (this is at the upper end of what we call "small-N" analysis; Ebbinghaus refers to it as "medium-N") that enable comparison both across and within these regimes. The cross-regime analysis is based on ordinal comparison as described above: regimes are ranked on the degree to which they are expected to promote early exit, and this ranking is correlated with a quantitative measure of early exit (relative exit rate). The within-regime analysis aims to account for the unexplained variation among countries within regimes; it is akin

to attempting to account for the residuals in a large-N quantitative analysis.

Through his combination of within-case process-oriented and cross-case contextual analysis, Ebbinghaus finds that it is not only the incentives provided by welfare state pathways to early retirement that explain the cross-national variations but also the particular strategies of employers and workplace representatives in coping with particular economic pressures caused by different production strategies. He uses outliers as special cases to investigate the interaction between "pull" and "push" factors. For instance, the Swedish case shows that a generous welfare state does not always produce high early retirement, while Japanese firms use mandatory retirement but also provide re-employment for older workers, explaining their high employment rate in a coordinated market economy.

In chapter 9, Adam Przeworski considers the crucial issue of selection bias in macrocomparative analysis. The concern is that what appears to be a causal effect of some institution or policy might rather be an effect of whatever gave rise to that institution or policy. For example, generous family policies are found mainly in the Nordic countries. We observe a positive association between family policy generosity and female employment rates across countries, but the true cause of the latter might be some feature of the Nordic societies or their policy-making processes that led them to adopt generous family policies, rather than the family policies themselves. As Przeworski puts it: "The generic problem in identifying causal effects is how to answer the counterfactual question: what would have occurred had the cause been absent?" In this example, the counterfactual hypothesis is that female employment rates would not be comparatively high in the Nordic countries had those countries not implemented generous family policies. Przeworski notes that "Whether we can successfully solve such problems is … largely a matter of luck, namely whether history has been kind enough to generate observations that can be used to inform us about the plausible counterfactuals."

The substantive question Przeworski explores is the impact of political regime – conceptualized dichotomously as democracy or autocracy – on labor force participation. Because the affluent countries are all democracies and because the approach Przeworski uses to estimate selection bias requires a relatively large number of countries, he includes not only rich nations but developing ones as well. His chapter nicely illustrates the use of appropriate techniques for addressing the selection bias worry.

In analyses confined to the rich countries, selection bias is, unfortunately, both more likely to be present and less amenable to statistical

estimation. The only recourse for the macrocomparative analyst is careful, and explicit, counterfactual thinking (Fearon, 1991; Esping-Andersen and Przeworski, 2001). For the most part the contributions to this book set this issue aside, but it ought to be an increasingly prominent concern in such analyses.

Onward

The chapters in this volume attempt to highlight the advantages and drawbacks of some prominent methodological approaches to macrocomparative analysis. The principal aim is to help researchers – ourselves included – to make more informed choices about which approach(es) to use in their research and to make better use of whichever one(s) they choose. We hope the book succeeds in this endeavor.

Notes

1. Thus far, however, relatively few macrocomparative studies have made use of more than one of these methods in analyzing a particular research question. Regression and small-N analysis are combined in Boix (1998), Huber and Stephens (2001), Swank (2002), and Kenworthy (2008). Ebbinghaus and Visser (1999) and Hicks (1999) couple regression with QCA.
2. It is sometimes thought that an analysis of a single country, usually referred to as a "case study," is not comparative. But most such studies are comparative (Rueschemeyer, 2003; Gerring, 2005). The comparison is not across countries but rather over time and/or across sub-units (regions, localities) within the country. Although small-N analyses tend to be qualitative, they can be quantitative as well; the distinction between small-N and large-N analysis is not the same as that between quantitative and qualitative analysis.

References

Ayer, A. J. 1956. *The Problem of Knowledge*. London: Penguin.

Baccaro, Lucio and Diego Rei. 2007. "Institutional Determinants of Unemployment in OECD Countries: Does the Deregulatory View Hold Water?" *International Organization* 61: 527–569.

Barro, Robert J. 2000. "Inequality and Growth in a Panel of Countries." *Journal of Economic Growth* 5: 5–32.

Blau, Francine D. and Lawrence M. Kahn. 2002. *At Home and Abroad: U.S. Labor Market Performance in International Perspective*. New York: Russell Sage Foundation.

Boix, Carles. 1998. *Political Parties, Growth, and Equality*. Cambridge: Cambridge University Press.

Brady, Henry E. and David Collier, eds. 2004. *Rethinking Social Theory*. Lanham, MD: Rowman and Littlefield.

Collier, David, Henry E. Brady, and Jason Seawright. 2004. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." Pp. 229–66 in *Rethinking Social Inquiry*, edited by Henry E. Brady and David Collier. Lanham, MD: Rowman and Littlefield.

Daly, Mary. 2000. "A Fine Balance: Women's Labor Market Participation in International Comparison." Pp. 467–510 in *Welfare and Work in the Open Economy, Volume II: Diverse Responses to Common Challenges*, edited by Fritz W. Scharpf and Vivien A. Schmidt. Oxford: Oxford University Press.

Dingeldey, Irene. 2002. "European Tax Systems and Their Impact on Family Employment Patterns." *Journal of Social Policy* 30: 653–72.

Ebbinghaus, Bernhard. 2005. "When Less Is More: Selection Problems in Large-N and Small-N Cross-National Comparisons." *International Sociology* 20: 133–52.

Ebbinghaus, Bernhard. 2006. *Reforming Early Retirement in Europe, Japan and the USA*. Oxford: Oxford University Press.

Ebbinghaus, Bernhard and Jelle Visser. 1999. "When Institutions Matter: Union Growth and Decline in Western Europe, 1950–1995." *European Sociological Review* 15: 135–58.

Eckstein, Harry. 1975. "Case Study and Theory in Political Science." In *Strategies of Inquiry: The Handbook of Political Science*, vol. 7, edited by F. Greenstein and N. Polsby. Reading, MA: Addison-Wesley.

Esping-Andersen, Gøsta. 1999. *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.

Esping-Andersen, Gøsta, with Duncan Gallie, Anton Hemerijck, and John Myles. 2002. *Why We Need a New Welfare State*. Oxford: Oxford University Press.

Esping-Andersen, Gøsta and Adam Przeworski. 2001. "Quantitative Cross-National Research Methods." Pp. 12649–55 in *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Elsevier.

Fearon, James D. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43: 169–95.

Ferrarini, Tommy. 2003. *Parental Leave Institutions in Eighteen Post-War Welfare States*. Swedish Institute for Social Research Dissertation Series, No. 58.

Ferrera, Maurizio, Anton Hemerijck, and Martin Rhodes. 2000. "The Future of Social Europe: Recasting Work and Welfare in the New Economy." Report prepared for the Portuguese Presidency of the European Union.

Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." Pp. 131–50 in *Political Analysis*, edited by James A. Stimson. Ann Arbor: University of Michigan Press.

George, Alexander L. and Andrew Bennett. 2004. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Gerring, John. 2004. "What Is a Case Study and What Is It Good For?" *American Political Science Review* 98: 341–54.

Goertz, Gary and Harvey Starr. 2003. "Introduction: Necessary Condition Logics, Research Design, and Theory." Pp. 1–23 in *Necessary Conditions*, edited by Gary Goertz and Harvey Starr. Lanham, MD: Rowman and Littlefield.

Gornick, Janet C. and Marcia K. Meyers. 2003. *Families That Work: Policies for Reconciling Parenthood and Employment*. New York: Russell Sage Foundation.

Greene, William H. 2003. *Econometric Analysis*, 5th edition. New York: Prentice Hall.

Griffin, Larry J., Pamela Barnhouse Walters, Phillip O'Connell, and Edward Moor. 1986. "Methodological Innovations in the Analysis of Welfare-State Development: Pooling Cross Sections and Time Series." Pp. 101–38 in *Futures for the Welfare State*, edited by Norman Furniss. Bloomington: Indiana University Press.

Hall, Peter A. 2003. "Aligning Ontology and Methodology in Comparative Research." Pp. 373–404 in *Comparative Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer. Cambridge: Cambridge University Press.

Hicks, Alexander. 1999. *Social Democracy and Welfare Capitalism*. Ithaca, NY: Cornell University Press.

Hicks, Alexander and Lane Kenworthy. 1998. "Cooperation and Political Economic Performance in Affluent Democratic Capitalism." *American Journal of Sociology* 103: 1631–72.

Hicks, Alexander and Lane Kenworthy. 2003. "Varieties of Welfare Capitalism." *Socio-Economic Review* 1: 27–61.

Huber, Evelyn and John D. Stephens. 2001. *Development and Crisis of the Welfare State*. Chicago: University of Chicago Press.

Iversen, Torben and Anne Wren. 1998. "Equality, Employment, and Budgetary Restraint: The Trilemma of the Service Economy." *World Politics* 50: 507–46.

Jahoda, Marie. 1982. *Employment and Unemployment: A Social Psychological Analysis*. Cambridge: Cambridge University Press.

Janoski, Thomas and Alexander Hicks, eds. 1994. *The Comparative Political Economy of the Welfare State*. Cambridge: Cambridge University Press.

Jaumotte, Florence. 2003. "Female Labour Force Participation: Past Trends and Main Determinants in OECD Countries." Working Paper 376, OECD Economics Department. Paris: OECD.

Katz, Aaron, Matthia vom Hau, and James Mahoney. 2005. "Explaining the Great Reversal in Spanish America: Fuzzy-Set Analysis versus Regression Analysis." *Sociological Methods and Research* 33: 539–73.

Kenworthy, Lane. 2004. *Egalitarian Capitalism*. New York: Russell Sage Foundation.

Kenworthy, Lane. 2006. "Reconsidering the Effect of Public Opinion on Social Policy Generosity in Affluent Countries." Unpublished. Department of Sociology, University of Arizona.

Kenworthy, Lane. 2007. "Toward Improved Use of Regression in Macrocomparative Analysis." *Comparative Social Research* 24: 343–50.

Kenworthy, Lane. 2008 (forthcoming). *Jobs with Equality*. Oxford: Oxford University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton, NJ: Princeton University Press.

Kittel, Bernhard. 1999. "Sense and Sensitivity in Pooled Analysis of Political Data." *European Journal of Political Research* 35: 225–53.

Kok, Wim, Carlos Dell'Aringa, Federico Duran Lopez, Anna Eckström, Marta João Rodrigues, Christopher Pissarides, Annette Roux, and Günther Schmid. 2003. *Jobs, Jobs, Jobs: Creating More Employment in Europe*. Report of the European Commission's Employment Taskforce. Brussels: European Commission.

Korpi, Walter. 2000. "Faces of Inequality: Gender, Class, and Patterns of Inequalities in Different Types of Welfare States." *Social Politics* 7: 127–91.

Lieberman, Evan S. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99: 435–52.

Lieberson, Stanley. 1991. "Small *N*'s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." *Social Forces* 70: 307–20.

Mahoney, James. 2000. "Strategies of Causal Inference in Small-*N* Analysis." *Sociological Methods and Research* 28: 387–424.

Mahoney, James. 2003. "Strategies of Causal Assessment in Comparative Historical Analysis." Pp. 337–72 in *Comparative Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer. Cambridge: Cambridge University Press.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14: 227–49.

Mahoney, James and Dietrich Rueschemeyer, eds. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

Mandel, Hadas and Moshe Semyonov. 2006. "A Welfare State Paradox: State Interventions and Women's Employment Opportunities in 22 Countries." *American Journal of Sociology*. 111: 1910–49.

Manicas, Peter T. 2006. *A Realist Philosophy of Science: Explanation and Understanding*. Cambridge: Cambridge University Press.

Meyers, Marcia K., Janet C. Gornick, and Katherin E. Ross. 1999. "Public Childcare, Parental Leave, and Employment." Pp. 117–46 in *Gender and Welfare State Regimes*, edited by Diane Sainsbury. Oxford: Oxford University Press.

Morgan, Kimberly and Kathrina Zippel. 2003. "Paid to Care: The Origins and Effects of Care Leave Policies in Western Europe." *Social Politics* 10: 49–85.

OECD (Organization for Economic Cooperation and Development). 1994. *The OECD Jobs Study*. Paris: OECD.

OECD. 2001. "Balancing Work and Family Life: Helping Parents Into Paid Employment." Pp. 129–66 in *OECD Employment Outlook*. Paris: OECD.

OECD. 2005. *Extending Opportunities: How Active Social Policy Can Benefit Us All*. Paris: OECD.

OECD. 2006. *OECD Employment Outlook: Boosting Jobs and Incomes*. Paris: OECD.

Orloff, Ann. 2002. "Women's Employment and Welfare Regimes." Social Policy and Development Programme Paper 12. United Nations Research Institute for Social Development.

Passmore, John. 1967. *A Hundred Years of Philosophy*. London: Macmillan.

Pettit, Becky and Jennifer Hook. 2002. "The Structure of Women's Employment in Comparative Perspective." Working Paper 330. Luxembourg Income Study. Available at: www.lisproject.org.

Phelps, Edmund S. 1997. *Rewarding Work*. Cambridge, MA: Harvard University Press.

Plantenga, Janneke and Johan Hansen. 1999. "Assessing Equal Opportunities in the European Union." *International Labour Review* 138: 351–79.

Pruss, Alexander R. 2006. *The Principle of Sufficient Reason*. Cambridge: Cambridge University Press.

Ragin, Charles C. 1987. *The Comparative Method*. Berkeley: University of California Press.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2005. "Core versus Tangential Assumptions in Comparative Research." *Studies in Comparative International Development* 40: 33–8.

Ragin, Charles C. 2006. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." *Political Analysis* 14: 291–310.

Ragin, Charles C. and Benoît Rihoux. 2004. "Qualitative Comparative Analysis (QCA): State of the Art and Prospects." *Qualitative Methods Newsletter*, Fall: 3–13.

Rubery, Jill, Mark Smith, and Colette Fagan. 1999. *Women's Employment in Europe*. London: Routledge.

Rueschemeyer, Dietrich. 2003. "Can One or a Few Cases Yield Theoretical Gains?" Pp. 305–36 in *Comparative Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer. Cambridge: Cambridge University Press.

Ruhm, Christopher J. 1998. "The Economic Consequences of Parental Leave Mandates: Lessons from Europe." *Quarterly Journal of Economics* 113: 285–317.

Sainsbury, Diane. 1999. "Taxation, Family Responsibilities, and Employment." Pp. 185–209 in *Gender and Welfare State Regimes*, edited by Diane Sainsbury. Oxford: Oxford University Press.

Scharpf, Fritz W. and Vivien A. Schmidt, eds. 2000. *Welfare and Work in the Open Economy. Volume 1: From Vulnerability to Competitiveness*. Oxford: Oxford University Press.

Schettkat, Ronald. 2005. "Is Labor Market Regulation at the Root of European Unemployment? The Case of Germany and the Netherlands." Pp. 262–83 in *Fighting Unemployment: The Limits of Free Market Orthodoxy*, edited by David R. Howell. Oxford: Oxford University Press.

Seawright, Jason. 2002. "Testing for Necessary and/or Sufficient Causation: Which Cases Are Relevant?" *Political Analysis* 10: 178–93.

Seawright, Jason. 2005. "Qualitative Comparative Analysis vis-à-vis Regression." *Studies in Comparative International Development* 40: 3–26.

Shalev, Michael. 2007. "Limits and Alternatives to Multiple Regression in Comparative Research." *Comparative Social Research* 24: 261–308.

Skocpol, Theda and Margaret Somers. 1980. "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22: 174–97.

Stier, Haya, Noah Lewin-Epstein, and Michael Braun. 2001. "Welfare Regimes, Family-Supportive Policies, and Women's Employment Along the Life-Course." *American Journal of Sociology* 106: 1731–60.

Swank, Duane. 2002. *Global Capital, Political Institutions, and Policy Change in Developed Welfare States*. Cambridge: Cambridge University Press.

Visser, Jelle. 2002. "The First Part-Time Economy in the World: A Model To Be Followed?" *Journal of European Social Policy* 12: 23–42.

Visser, Jelle and Anton Hemerijck. 1997. *"A Dutch Miracle": Job Growth, Welfare Reform, and Corporatism in the Netherlands*. Amsterdam: Amsterdam University Press.

Wilson, Sven E. and Daniel M. Butler. 2007. "A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications." *Political Analysis* 15: 101–23.

Wilson, William Julius. 1996. *When Work Disappears*. New York: Knopf.

Winegarden, C. R. and Paula M. Bracy. 1995. "Demographic Consequences of Maternal-Leave Programs in Industrial Countries: Evidence from Fixed-Effects Models." *Southern Economic Journal* 6: 1020–35.