Routledge Handbook of the Welfare State

Bent Greve

How to analyze welfare states and their development?

Publication details
Barbara Vis
Published online on: 02 Jul 2018

How to cite: - Barbara Vis. 02 Jul 2018, How to analyze welfare states and their development? from: Routledge Handbook of the Welfare State Routledge
Accessed on: 03 Aug 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
How to analyze welfare states and their development?

Barbara Vis

Introduction

For describing and understanding welfare states’ development, it is necessary to know how to analyze this development. This chapter discusses some of the main quantitative and qualitative approaches. Rather than describing in detail how to conduct the different kinds of analyses, this chapter focuses on the available approaches.

In this chapter, I discuss three broad categories of approaches: (1) quantitative, (2) qualitative comparative, and (3) single case studies, focusing on the most important methods available and illustrating these with one or two examples. This overview indicates that which approach is “best” depends on the researcher’s research question and goal.

The dependent variable and data

Before turning to the different approaches, let me first address two issues that all welfare state research has to deal with: (1) what is the study’s dependent variable?; (2) which type of data to use? The first issue relates to the so-called dependent variable problem (see Clasen and Siegel 2007). The dependent variable problem concerns how to conceptualize, for instance, the generosity of welfare states theoretically, how to operationalize it empirically and, finally, how to measure it. Since different authors adopt different definitions, their operationalizations differ and, consequently, their findings do too.

In the 1980s and early 1990s, the typical study of welfare state development – at that time focusing mostly on the welfare state’s expansion – used aggregate expenditure data to assess this development. Such data were the only ones available for a range of countries and periods at the time, which explains their widespread use. In his Three Worlds of Welfare Capitalism, Gösta Esping-Andersen (1990) argued that this focus was problematic. Because “it is difficult to imagine [that] anyone struggled for spending per se” (p. 21), Esping-Andersen considered expenditures to be epiphenomenal for assessing welfare state development. Scholars should rather examine the social rights of citizens. A key measure of these social rights that Esping-Andersen developed was the decommodification-index, which captured the degree of decommodification: the extent to which welfare states allow citizens to maintain a livelihood.
independently from the market. Many welfare state researchers concurred that focusing on social rights to assess welfare state development was a major step forward and therefore applauded the decommodification-index.

Esping-Andersen’s decommodification-index had two drawbacks: (1) it was based on data for only one point in time, and thus did not allow for analyses over time; and (2) given that this time point was around 1980 only, the index was already relatively out of date when Esping-Andersen published *The Three Worlds* in 1990. Two datasets that were built to capture the degree of social citizenship rights do not have these drawbacks. Korpi and Palme’s (2007) Social Citizenship Indicator Program (SCIP) provides data on social insurance programs’ (old age pensions, unemployment insurance and sick pay) replacement rates, benefit eligibility, coverage, waiting days, benefit duration and sources of financing for 18 OECD countries. The data are available for five-year intervals (1960, 1965, etc.), up to 2005 and going back to 1930 for some programs and categories. Scruggs and colleagues’ (2014) Comparative Welfare Entitlements Dataset (CWED), which is an update from Scruggs’ 2004 dataset, is in many respects similar to the SCIP dataset. The CWED focuses on the same social insurance programs (old age pensions, unemployment insurance and sick pay) and – like the SCIP dataset – provides data on replacement rates, benefit eligibility, coverage, waiting days and benefit duration. Different from the latter, the CWED includes time-series data, allowing for more fine-grained analyses over time. The time period covered by the CWED dataset is shorter than that of the SCIP dataset but more up to date. For most countries and programs, the data cover 1971 to 2010. The CWED dataset covers a larger set of countries: 27 OECD and 6 non-OECD countries.

Whereas the publication of these two datasets did spur a whole body of research, it did not put an end to questions relating to data. One of the questions is to what extent the SCIP and the CWED datasets measure the same concepts in the same way. Since both datasets aim to capture social citizenship rights, researchers might expect that the datasets would. Several contributions in a 2013 special issue of the *Journal of European Public Policy* addressed this question, concluding that there are clear differences across the two datasets in a key measure: the program’s replacement rates (see, e.g., Wenzelburger et al. 2013). This conclusion would certainly be a worrying one for the field. However, Danforth and Stephens (2013) demonstrate that the differences across the two datasets are more justifiable and, thereby, much less problematic. The datasets differ in three key respects (Danforth and Stephens 2013: 1286–1287): (1) they operationalize the core concept of social right differently; (2) they draw upon different data sources, and (3) the assumptions and methods of calculating the aggregate figures differ. Whereas the differences across the SCIP and the CWED datasets are, thus, largely justifiable, neither covers so-called new social risks, such as how to combine work and family life. Instead, both datasets focus on the risks faced by typical male breadwinner industrial workers with lifetime employment (cf. Danforth and Stephens 2013: 1296). There are different datasets that do focus on new social risks, such as the OECD’s (2017) *Family Policy Database*. In addition, the EU’s Mutual Information System on Social Protection database (MISSOC 2017) includes comparative data on a broad range of social protection systems, including new social risks’ areas (such as maternity and family).

As indicated, there are sometimes good theoretical reasons to focus on expenditures (see, e.g., Siegel 2007: 50–54), such as when a researcher is interested in cost containment. Unfortunately, as De Deken and Kittel (2007) show, expenditure data in general are riddled with problems. A major problem is the comparability of the data, which results from the difficulty of having a consistent accounting edifice that can be systematically applied to all countries (p. 73). Another, related, problem is that missing data for disaggregated measures of social expenditure, such as spending on active labor market policies (ALMPs), lead – if uncorrected – to aggregate
social expenditure data that are not comparable across countries or over time (see Siegel 2007: 62–63). Excluding expenditures on ALMPs altogether from the aggregate social expenditure measure to circumvent this “patchy data” problem, as, for example, Castles (2004) does, is no panacea either. The rising importance of such policies in among other things the discourse of the Organization of Economic Co-operation and Development (OECD), the EU and governments alike makes ALMPs too relevant to exclude (on the problems and challenges of analyzing ALMPs comparatively, see Clasen et al. 2016).

Quantitative analyses

Ever since the development of welfare states became a research topic for scholars, quantitative analyses have been conducted. The sophistication of these analyses increased substantially over time. In the first decades, quantitative analyses were typically cross-sectional, focusing on the variation across countries and/or programs at one point in time. The seminal work by Flora and Heideheimer (1981) is an excellent example hereof. In the 1990s and early 2000s, the field turned en masse to conducting their quantitative analyses by means of pooled time-series cross-sectional analysis. Garrett and Mitchell’s (2001) oft-cited and re-analyzed study is characteristic of this approach, which combines analyzing the variation across cases (typically countries) and the variation over time. A third quantitative approach in the study of welfare state development is typology-building and re-analyzing existing typologies (see Chapters 27 and 28), an approach that spawned an extensive body of research, especially after the publication of Esping-Andersen’s The Three Worlds (see Emmenegger et al. (2015) for an overview).

Cross-sectional analysis

In the its early days, a lot of (quantitative) comparative welfare state research was mainly cross-sectional in that it examined the differences and similarities of welfare states at one point in time. Hereby, the findings of different cross-sectional analyses were sometimes compared. In Flora and Heideheimer’s (1981) edited volume which studies the development of the welfare state in Europe and the USA, for example, the contributors compared social insurance systems such as unemployment insurance (chs 2 and 5), the expenditures on programs like health services (ch. 3), political macro-characteristics at the time of Bismarck’s social insurance legislation in Germany in the 1880s (ch. 4), and the introduction of income tax systems (ch. 6). They also compared the sociopolitical paths of the adoption of social insurance legislation across time periods (ch. 2), hence contrasting different cross-sections. In terms of dependent variable, the contributions in Flora and Heideheimer focused mostly on the introduction of different programs and on how much was spent on them, which is in line with their research questions and goals.

A somewhat more recent cross-sectional analysis is Castles’ (2004) study that puts crisis, myths and measurement at its center. Specifically, Castles investigated whether processes of globalization, population aging and declining birth rates that allegedly endanger the welfare state “will lead to social policy disaster or whether they too will turn out to be mythical in character” (p. 2). Because these globalization and demographic crises hypotheses can be formulated more or less explicitly in terms of expenditure implications, it makes sense that Castles used measures of social expenditure as his dependent variables to test comparatively these hypotheses. The complexity of his measures increased throughout the book and showed that, in spite of their problems (see above), the disaggregated data in the social expenditures dataset of the OECD had much to offer to this end. Specifically, Castles first showed that a “race to the bottom” simply
did not materialize in aggregate social expenditure (ch. 2) and subsequently constructed an
expenditure-shares typology based on disaggregated spending data which revealed that the
structure of social provision and the priorities of the welfare types barely converged (ch. 3).
Although Castles’ cross-sectional analysis enabled him to debunk the globalization and population
hypotheses for most countries by examining the changes in the expenditure patterns
between 1980 and 1998, it provided no remedy for the aforementioned problem of the compara-
rability of the expenditure data.

A final example of a cross-sectional study is Marx’s (2015) study on how the experience of
temporary employment influences political preferences and behavior. Marx tests both existing
theories on and approaches to this relationship, particularly insider–outsider theory and risk-
based explanations of political behavior, and develops new arguments, such as that temporary
workers are especially likely to support small green parties. Marx’s main source of data is the
European Social Survey (ESS), from which he pools data for 17 European countries for the
three ESS waves between 2008 and 2013. This pooling implies that the study “is mainly con-
cerned with effects of temporary employment that are relatively uniform across countries”
(p. 8) – something that Marx considers “a first step to advance empirical knowledge” (p. 8).
Because the ESS does not include all variables to test his hypotheses, Marx complements the
ESS data with the German Longitudinal Election Studies (GLES) of 2009 and 2013 (Rattinger
et al. 2012, 2014) to analyze voting behavior. Furthermore, Marx collected original survey data
in four countries that are similar in having particularly high shares of temporary employment but
that are different in many other respects: the Netherlands, Poland, Spain and Sweden. The latter
enables Marx to assess the generalizability of the political behavior effects. By combining various
types of survey data, Marx was able to avoid each type’s shortcoming and systematically test his
hypotheses.

**Pooled time-series cross-sectional analysis**

Since the mid-1990s, many welfare state scholars, as well as political scientists more generally,
turned to pooled time-series cross-sectional analysis as their preferred quantitative approach.
This technique has two main advantages which explains why they did so. A first one is that
pooling cross-sectional and time-series data reduces the so-called small-n problem. This problem,
which often bedevils comparative welfare state research, pertains to the problem of having too
small a number of observations to make all relevant inferences; that is, to test all relevant hypo-
theses. The rule is that one observable implication can only give independent information about
one other fact; that is, each observation allows for one inference at most. If the observations are
not independent, as is often the case in comparative welfare state research, we need (many)
more than n observations to make n inferences. This small–n problem of more inferences than
observations is often present in qualitative case studies, but also in cross-national research that
studies, for instance, 18 OECD countries. Including both cross-sectional and time-series data
reduces this problem by increasing the number of observations. For example, adding time-series
data for 10 years to 18 OECD countries increases the number of observations from 18 to 180.
This allows for more fully specified models to be estimated and thus for more inferences to be
drawn. A second advantage is that pooling allows for the control of exogenous shocks that all
units of observation share, such as an oil crisis (by controlling for time effects), and limits omitted
variable bias (by controlling for unit effects).

Unfortunately, pooling data also introduces a wide range of (new) problems. Specifically, the
potential problems of the time-series dimension, such as autocorrelation (that is, error terms
that are not independent from one time period to another) and non-stationarity (that is, the
Welfare states and their development

...persistence of variables over time), as well as of the cross-sectional dimension like heteroskedasticity (that is, the variance of the error terms varies across units), are often present and regularly reinforce one another. Because of the way these problems are (not) handled, the results of pooled time-series cross-sectional analyses are regularly not very reliable (Kittel and Winner 2005; Plümper et al. 2005). Wilson and Butler’s (2007) review of 195 published articles in political science illustrates the problems involved well. Wilson and Butler showed that crucial specification issues are usually not discussed or considered and that sensitivity analyses are even rarer. This conclusion relates to the observation of Beck—who in 1995 with Katz told the political science community “what to do (and not to do) with time series cross-section data” (Beck and Katz 1995)—that “many analysts thought that the P in PCSE [panel corrected standard errors] stood for Panacea and not Panel” (Beck 2007: 97). Technically, it is possible to test for all potential problems in time-series cross-sectional analysis and to remedy the occurring ones to arrive at an econometrically sound model. However, because fixing problems usually means that the model is re-specified, the regression model to be estimated also changes. A result may be that the findings of the estimated model no longer provide an answer to the research question posed.

The advantages as well as the potential problems of pooled time-series cross-sectional analyses are visible in the well-known study by Garrett and Mitchell (2001) on the effect of globalization on welfare state effort (measured by government spending, government consumption and social security transfers) and taxation. By adding a time dimension of 33 years to the data on 18 developed democracies, they have a maximum of around 600 observations in their regressions. This allows Garrett and Mitchell to estimate more complex models; that is, to include more explanatory variables than would have been the case if they had conducted cross-sectional analyses. Moreover, the inclusion of a lagged dependent variable allowed them to assess the stickiness of government spending over time, another theoretical expectation. However, the re-analyses by Kittel and Winner (2005) and Plümper and colleagues (2005) of Garrett and Mitchell’s study show that the latter suffered from many of the potential problems of pooled time-series cross-sectional analyses identified above. Such conclusions have led to a somewhat less enthusiastic stance towards this approach within the field and have been, at least partly, responsible for the rise of the use of other approaches like configurational comparative analysis (see below). They have also resulted in new, more sophisticated ways to address (especially) the statistical problems involved. Overall, however, pooled time-series cross-sectional analysis remains a main quantitative approach for studying welfare state development. Despite its drawbacks, it remains a useful approach when a researcher is interested in the variation in welfare state development across countries and over time, and can employ relatively straightforward dependent variables, such as social expenditures or generosity data.

Typology-building and re-analyzing existing typologies

A third and final quantitative approach to analyzing welfare state development is building typologies and re-analyzing existing typologies. This approach uses the fact that individual welfare states vary a great deal but that some are more similar than others; similarities (and differences) that are relatively persistent over time. Specifically, welfare states are grouped together in “worlds”, “regimes” or “families” and a typology can capture how they are grouped together. Typologies are useful empirical and heuristic devices because they enable comparative analysis by grouping separate entities into simpler categories, thereby reducing the existing complexity and offering a first step towards explaining the cross-national variation. Building on the work of Titmuss (1958), Esping-Andersen (1990) developed his well-known typology of three welfare state regimes: a social democratic regime, a conservative or continental regime, and a liberal...
regime (see also Chapters 11 to 14, this volume). For building the typology, Esping-Andersen focused particularly on the degree of decommodification; that is, the situation “when a person can maintain a livelihood without reliance on the market” (p. 22), the degree of stratification, and the primary location of welfare production (the state, the market or the family).

Especially since the late 1990s, re-analyzing the empirical value of existing typologies has become a growth industry in welfare state research (for overviews and discussions, see Arts and Gelissen 2010: 574–577; Powell and Barrientos 2011; Barrientos 2015; Emmenegger et al. 2015). These analyses are usually done by means of cluster analysis or principal component analysis and, less often, by factor analysis and qualitative comparative analysis. Often, the research question in such analyses is how well cases “fit” the typology (Van Kersbergen and Vis 2015). Interestingly, this question is actually irrelevant in relation to a typology such as Esping-Andersen’s because a typological approach differs from an ideal-type approach (see Van Kersbergen and Vis 2014; see also Chapter 28, this volume). An ideal-type is a theoretical construct to which empirical cases may or may not correspond. A typology, conversely, is a classification device in which all empirical cases are located as belonging to one, and only one, of the types. In a typology, a country belongs thus to either the social democratic regime or to the conservative regime or to the liberal regime. In ideal-typical analysis, conversely, the question is to what extent an empirical case belongs to the ideal-type. A country may, for example, have the closest fit to the social democratic ideal-type but also have features of the liberal regime type. The goodness-of-fit question thus matters only when a researcher works with ideal-types.

Comparative qualitative analyses

There are numerous studies on welfare state development that adopt a comparative qualitative approach. Such studies are often inspired by empirical puzzles, such as why welfare state retrenchment happened in country A but not in country B. Here, I focus on a “traditional” small-\(n\) case study by Green-Pedersen (2002) and a study applying configurational comparative approaches by Vis (2010). While this selection is admittedly arbitrary, these two studies exemplify two different ways of conducting a comparative analysis qualitatively.

“Traditional” small-\(n\) comparative case studies

A popular way of assessing welfare state development is by conducting a comparative analysis of a small number of cases – typically two to four – sometimes alongside a quantitative analysis of a larger set of cases. Which cases to select for a small-\(n\) comparison is an important question; the case selection should offer the best possible test of the study’s theoretical and/or empirical argument(s). In general, small-\(n\) welfare state scholars employ either Mills’ most similar systems design (MSSD) or, somewhat less often, his most different systems design (MDSD) (for a description of these designs, see Landman and Carvalho 2017: ch. 2). In a MSSD, the researcher selects two (or more) cases that have all variables in common – that is, are most similar – except for the outcome and the one or two variables of theoretical interest. Hereby an MSSD mimics experimental designs in which similar groups are compared that vary only in whether they receive the treatment or not. A MDSD, conversely, implies the selection of cases that differ on all variables – that is, are most different – except for the outcome and the one or two variables of theoretical interest. The idea of this latter design is that if a similar factor (e.g., leftist partisanship) produces a particular outcome (such as welfare state development) in cases that differ in all other ways that might influence the outcome, the finding likely holds across cases that are more similar as well. What is often considered a drawback of a MDSD is that there is no variation on
Welfare states and their development

the dependent variable, hence causing selection bias. A problem when employing a MSSD is finding cases that are sufficiently similar.

An example of a small-\(n\) comparative analysis that applies a MSSD is Green-Pedersen’s (2002) study of the role of party competition in welfare state retrenchment. Green-Pedersen compared retrenchment in old age pensions, unemployment benefits and disability pensions/early retirement benefits in Denmark and the Netherlands between the early 1980s and the late 1990s. Green-Pedersen argued that two factors are crucial for explaining the variation in retrenchment across these two countries: (1) the type of political party, and (2) the type of party competition. Denmark and the Netherlands, similar in many other respects, differed particularly in terms of party competition. Green-Pedersen proposed that when party competition allowed for a consensus among parties, implementing unpopular retrenchment initiatives (and getting away with it at the polls) would be easier. In so-called bloc systems in which a left-wing bloc and a right-wing bloc alternate in office, as in Denmark, consensus would emerge only if leftist parties reign. In such a system, vote seeking and office seeking coincide. In pivot systems like the Netherlands, conversely, a center party – usually a Christian democratic one – is pivotal in that it dominates both government formation and legislation. The pivot party can focus on vote seeking (because it will automatically be in office). The major party of the left, however, faces a dilemma between vote seeking and office seeking. This produces consensus when the center party opts for retrenchment, because the left-wing party then has to accommodate the center party if it wants to gain office. Green-Pedersen’s comparative analysis of retrenchment supported this argument.

Green-Pedersen’s study is also interesting owing to his dependent variable. He defined retrenchment as “a change in the [social security] scheme making it less attractive to the (potential) claimants” (2002: 58). The quality of benefits and the eligibility criteria determined the attractiveness of a scheme, measured by the reform’s budgetary implications. This narrow conceptualization and operationalization of retrenchment fitted well with Green-Pedersen’s interest in how the electorate responds to retrenchment. If there are no budgetary implications, it is unlikely that the electorate cares much about the reform, because it does not hurt their own (financial) position.

Configurational comparative approaches

A second (mostly) qualitative approach to study welfare state development is a family of approaches known collectively as configurational comparative approaches. These techniques, introduced to the social sciences by Ragin in 1987, have since been developed (e.g. Ragin 2008; Rihoux and Ragin 2009; Schneider and Wagemann 2012) and have become increasingly popular in welfare state research since the late 2000s. Configurational comparative approaches can formalize case-oriented analysis and thereby offer tools to improve comparative research. The approaches are particularly apt for identifying the minimally necessary and/or minimally sufficient (combinations of) conditions that bring about an outcome, such as welfare state development. Configurational comparative approaches may be applied to studies with a large-\(n\) (see e.g., Ragin and Fiss (2016) for a recent example). Still, the approaches have an edge on studies with an intermediate number of cases (between, say, 5 and 50). The choice for configurational comparative approaches should, however, be rooted in the researcher’s interest in identifying the combinations of minimally necessary and/or minimally sufficient conditions.

An example of a configurational comparative analysis of the fuzzy-set variant (fsQCA) is Vis’ (2010) study of the combinations of the minimally necessary and/or sufficient conditions for governments’ pursuit of different types of welfare state reform. Vis examined over 20 British,
Danish, Dutch and German governments between the late 1970s and mid-2000s. With this design, Vis combined a MDSD and a MSSD design (see above). The four countries differed with respect to a number of characteristics that possibly influence governments’ pursuit of reform, such as the type of welfare state, the party competition and the electoral system. In this sense, the countries were most different and focusing on them helped establish the robustness of the findings by controlling for the influence of contextual factors. The comparison of different governments within a country, conversely, was the most similar part. In terms of the dependent variable, or outcome in the terminology of configurational comparative approaches, Vis focused on two measures of so-called unpopular reform: a broad, more qualitative measure and benefit cutbacks. Her analysis showed that governments accepted the risk of electoral punishment involved in unpopular welfare state reform only when they found themselves confronted with losses (necessary condition) in the form of a deteriorating socio-economic situation and/or a deteriorating political position. Only then were they prepared to confront the risk of losing votes involved in unpopular reform in a, so to speak, final attempt to try to recoup (some of) the losses incurred. A deteriorating socio-economic situation was only sufficient for triggering reform when combined with one or two other conditions: a worsening political position or a rightist government composition. Testing Vis’ theoretical argument by means of a “traditional” quantitative approach like pooled time-series cross-sectional analysis or a “traditional” case study approach would have been more difficult, to say the least.

Single case studies

A final body of literature analyzes welfare state development by means of single case studies. The studies typically ask questions such as: why was a particular reform, against all odds, successfully implemented in country X? Or why did a particular reform fail to come about in that country? What is the detailed process, including the actors involved, their positions and ideas, behind this failure? Single case studies generally define and operationalize their dependent variable as broad and encompassing, tailored to the country they focus on and the specific research question.

Still, many studies in welfare state research that focus on one country and that labeled single case studies are actually comparative, because they examine the development within this country over time, for instance, by comparing different reforms (or the absence thereof) or across welfare state programs. Stiller’s (2010) study of welfare state reform in Germany is no exception here. Stiller examined four so-called structural reforms that took place in the 1990s and 2000s: two healthcare reforms, the Hartz IV reform that merged unemployment assistance and social assistance, and the 2001 pension reform. For all these reforms, Stiller assessed to what extent “ideational leadership” – ‘leadership with the help of ideas’ (p. 17) – was observable in the reform process and how this ideational leadership relates to the adoption of the reform. By carefully tracing the process of these reforms and, in particular, the role played by the minister involved in the reform (ideational leadership displayed or not?), Stiller showed that ideational leadership mattered for getting one of the healthcare reforms adopted but not for the other. In the pension reform and the Hartz IV reform, ideational leadership in combination with concession-making were important for getting the reforms adopted. By studying different reform processes in one country, Stiller employed a MSSD. An advantage of using this design within a single country context is that the context (e.g. institutional structure) could be kept relatively identical.

Lipset and Marks’ (2000) analysis of why socialism – and social democracy more broadly – failed in the United States (US) offers another example of a single case study that also includes comparative elements. Lipset and Marks assessed a series of explanations from the literature
about why the Socialist Party could not turn into a viable third party in the US, such as the impediments for third parties in the American two-party system and the split between the Socialist Party and the trade unions. While thus not focusing directly on welfare state development, the absence of socialism is one of the factors explaining why the American welfare state is so lean. The comparative part of Lipset and Marks’ work comes in when they assess the hypotheses. Regarding the influence of the party system, for instance, they argued that in countries with electoral systems similar to the American system, like Sweden up until 1907, Denmark up until 1915, the United Kingdom (UK) and Canada, socialist or social democratic parties did become durable and often influential parties. This indicates that the party system by itself cannot account for socialism’s failure in the US.

**Conclusion**

There are many approaches available for studying welfare state development, of which this chapter has discussed some of the most prominent ones. None of these approaches is “best” in that it trumps the others in all respects. Which approach is most suitable or best rather depends crucially on one’s research question and goal. When a researcher, for example, wants to get an in-depth understanding of how a process of welfare state development evolved in a particular country, a single case study would be the appropriate choice. If a researcher is interested more in how welfare states develops in general, that is, if he or she wants to arrive at a (more) generalizable theory about the factor(s) influencing this development, a quantitative approach would be a better choice. Whether he or she in that case should opt for a cross-sectional design or a pooled time-series cross-sectional design depends in turn on the variation the researcher wants to explain (only across countries or also over time?) and on the available data. Researchers who are interested in explaining welfare state development in a limited set of countries, possibly with the ambition of generalizing the findings to a larger set of countries, are probably best off conducting a comparative qualitative study. Depending on the precise research question, hypotheses, available data and the number of cases, either a small-n comparative analysis or a configurational comparative analysis would be the appropriate choice. None of the approaches discussed only has strengths and scholars need to be aware of the weaknesses of the approach they select. Sometimes, these weaknesses can be reduced by employing different approaches into one study – although this is also not achieved without difficulties.

**References**


