Establishing systematic evidence to assess the impact of power-sharing arrangements is important, both theoretically and politically. Whether these arrangements serve the long-term interests of peace-building, durable conflict management, and democratic consolidation in multiethnic societies remains an open question. Despite the desirability of establishing practical guidelines for crafting new constitutions, and the extensive literature which has developed, the evidence available to compare the performance of power-sharing on democracy and the management of ethnic conflict remains limited. Several approaches have been used in previous research, including in-depth treatment of selected national case studies, historical-institutional accounts of political development within particular countries, comparative cross-sectional approaches based on analyzing a subset of democracies (Lijphart) or minorities at risk (Gurr), and analogies drawn between the experience of the legislative underrepresentation of women with that of ethnic minorities. Unfortunately this literature remains inconclusive as a result of inconsistent results, with certain limitations in each approach. Given the inherent flaws of any single method taken in isolation, this study opts for a mixed research design, combining the virtues of rich and detailed studies of contrasting paired cases with the advantages of comparing large-N time-series cross-sectional data. Understanding both the methodology and its limitations is important in order to evaluate the results, but those wishing to focus upon the meat-and-potatoes findings may wish to skip ahead to Chapter 4.

Small-N Case Studies

Small-N national case studies have often been discussed to illustrate the pros and cons of power-sharing regimes. This approach is invaluable as a way to explore the complex processes of regime change, using historical narrative to examine detailed changes within each nation-state. Cases allow researchers to develop theories, to derive hypotheses, and to explore causal mechanisms. This
approach is particularly useful with outliers, such as the wealthy autocracy of Singapore and the poor democracy of Benin, which deviate from the generally observed pattern. This method fails to resolve the debate between proponents and critics of power-sharing, however, since the danger of potential selection bias means that different cases can be cited on both sides. There are also arguments about which cases best fit the consociational ideal type. The classic exemplars among established democracies are generally agreed to include the Netherlands, Belgium, and Austria (and possibly Switzerland), all plural societies containing distinct ethnic communities divided by language, religion, and region, with constitutions characterized by multiple veto-points and extensive power-sharing. Lijphart also highlights equivalent cases in many developing societies as diverse as Belgium, India, Switzerland, and South Africa but there are also many well-known ‘failed’ cases. The breakdown of consociational democracy is exemplified most clearly in Lebanon, where the 1943 National Pact divided power among the major religious communities, a system which collapsed in 1975 when civil war erupted. Other notable cases of malfunction include the consociational system in Cyprus, before civil war in 1963 and subsequent partition between the Greek and Turkish communities disrupted these arrangements. Another case which can be regarded as a potential failure concerns the intricate consociational arrangements for power-sharing along ethnic lines developed in the new constitution for Bosnia and Herzegovina set up by the Dayton Agreement, which seem to have reinforced ethnic homogeneity within each area, although also eliminating bloodshed. Czechoslovakia also experimented with these arrangements briefly in 1989–1993, before the ‘velvet revolution’ produced succession into two separate states. Studies have compared countries as diverse as Afghanistan, Cyprus, Kosovo, Macedonia, and South Africa, where these arrangements have been proposed or implemented. What often remains unclear from selected case study comparisons which focus upon states which have experienced power-sharing institutions or those which focus upon recent peace-building settlements, however, are the counterfactual cases which may have adopted majoritarian institutions decades or even centuries ago and thereby generated greater contemporary social cohesion. Consensus democracy may not be the root cause of the political problems experienced by societies such as Cyprus and Lebanon; nevertheless examples which have clearly failed temper strong claims that these arrangements, by themselves, are sufficient for managing ethnic conflict. Moreover it is also not clear whether consensus democracy is necessary for political stability in divided societies; we can also identify certain contemporary cases of newer democracies in multi-ethnic societies with majoritarian arrangements, including Mali and Botswana, which are classified by Freedom House as relatively successful at consolidating
Constitutional Reforms as ‘Natural Experiments’

An extension of the case study method uses historical explanations of constitutional reforms occurring within particular countries to provide an alternative way to examine changes in regimes, patterns of democratic consolidation, and the degree of conflict in multiethnic states. This is perhaps most effective with occasional ‘natural experiments’ allowing researchers to utilize ‘before’ and ‘after’ studies of the longitudinal impact within particular societies when political institutions change. The breakdown of regimes and the wholesale construction of new constitutional settlements, such as those which occurred in contemporary Iraq and Afghanistan, provide the most dramatic cases. But this approach is also illustrated in older democracies which have altered some fundamental aspects of their political arrangements, with a narrower reform which makes it easier to isolate specific institutional effects, as exemplified by comparison of the representation of the Maori community in New Zealand when the electoral system moved from majoritarian single-member districts to a combined system (mixed member proportional). Other natural experiments include monitoring the impact of varying degrees of regional autonomy on conflict in the Basque region, for example, whether this prompted ETA to renounce terrorist violence and end their bombing campaign in March 2006. Other cases include evaluating the impact of constitutional reforms on the linguistic communities and nationalist feelings in Belgium. Studies have also monitored changes in national identity in the United Kingdom and attitudes toward separatism following the establishment of the Scottish Parliament and Welsh Assembly.

The most common approach here has examined changes over time to electoral systems. These institutions used to be regarded as fairly stable, with only minor adjustments in the basic electoral formula translating votes into seats. In recent years, however, a range of minor and major reforms to electoral systems and processes have been more widely recognized. Hence Reynolds, Reilly, and Ellis noted that 27 countries worldwide changed the basic electoral formula used for their national parliamentary elections over the decade 1993–2004; for example, South Africa switched from first-past-the-post to list proportional representation (PR), Venezuela moved from list PR to a combined-dependent system, while Japan changed from single nontransferable vote to a combined-independent system. During a far longer period, Colomer’s study of electoral systems in democracies in 94 countries since the early nineteenth century monitored no fewer than 82 major changes in the electoral system used for legislative contests (for example, from majoritarian to mixed or from mixed to proportional formula). Considerable instability was evident in some countries; for example, Colomer reported that Greece experienced nine major changes in the electoral system, France had seven, and Italy and Portugal each had six changes.
A comparison by Goldner suggests that more than half the 125 countries that have held democratic elections since 1946 have experienced a change in their electoral system, based on a broader definition of what counts in this regard, including any shifts in district magnitude or assembly size of 20% of more, the introduction of presidential elections, alterations to electoral tiers or electoral formula, or the breakdown or reintroduction of democratic regimes. The estimated total number of reforms would rise even further if changes to the nomination and campaigning process were added to this list, including the widespread introduction of gender quotas into candidate recruitment processes or the use of new regulations governing party finance.

Electoral reforms represent natural experiments which are invaluable for understanding processes of political change in any one society over time. One important problem facing analysts, however, is that historically much else often alters simultaneously alongside these institutional reforms; for example, successive contests often produce shifts in patterns of party competition and in the composition of the governing coalition. These simultaneous changes make it difficult to isolate the specific impact of constitutional reforms, for example, on levels of democracy, the strength of ethnic identities, the degree of political stability, or patterns of multiethnic cooperation at community level. In addition, there are also problems in disentangling the direction of causality; negotiated peace-settlements which are the result of reduced intercommunal violence may also generate proportional electoral systems.

Understanding particular case studies, and processes of constitutional reform occurring within each nation, is therefore unlikely to provide definitive answers and more systematic analysis is required to generalize causal inferences across nation-states and over time. Even in large-N comparisons, however, involving multiple cases, the reliability and meaning of the results are heavily dependent upon the selection of the most appropriate sample of nation-states. Potential bias arises where the observations are restricted to a nonrandom set of observations. For example, Lijphart’s work provides the most extensive attempt to operationalize and measure ‘consensus’ democracies – defined in terms of their institutional characteristics. Lijphart classified three dozen ‘long-term democracies’, defined as those states which had been democratic from 1977 to 1996, in terms of 10 institutions and then compared their performance. Unfortunately by committing the sin of ‘selecting cases on the dependent variable’, this comparative framework is inherently flawed for any analysis of patterns of political stability and democratic consolidation. The universe excludes comparison of unstable states, whether they subsequently failed in violent internal wars (Lebanon, Yugoslavia, Cyprus), experienced a coup d’etat (Gambia, Fiji), split with peaceful succession (Czechoslovakia), or simply gradually became more repressive and authoritarian (Russia, the Maldives, Bhutan, Egypt, Liberia). Similar systematic sampling biases are often found in many other attempts to generalize about the institutions common in democracies, for example, Goldner’s and Colomer’s studies of electoral reform, which exclude any comparison of autocracies.
Evidence and Methods

Minorities at Risk

Another common approach utilizes the comparative dataset derived from the Minorities at Risk (MAR) project. Initiated by Ted Robert Gurr, this project compares politically active communal groups.\(^1\) The evidence has been widely used in the literature on ethnic conflict, including incidents of nonviolent protest, violent protest, and political rebellion (the latter ranging from sporadic acts of terrorism to cases of protracted civil war). For example, Frank Cohen employed this data to compare patterns of behavior among 233 ethnic groups in 100 countries. The study concluded that both federalism and PR electoral systems were significantly related to lower levels of rebellion by ethnic groups, confirming that consensus democracies are more effective at managing ethnic conflict. Nevertheless, the comparison of ‘minorities at risk’ remains problematic for this task mainly because sample bias is also inherent in the methodology employed in selecting cases for observation. A ‘minority at risk’ is defined by the codebook as an ethno-political group that (i) “collectively suffers, or benefits from, systematic discriminatory treatment vis-à-vis other groups in a society” and/or (ii) “collectively mobilizes in defense or promotion of its self-defined interests.”\(^1\) Religious, linguistic, and regional minorities are therefore systematically excluded from the dataset if they are successfully integrated or assimilated into society so that they do not organize separately as a political association. Multiple cleavages exist in society, but MAR only recognizes those which become and remain politically salient. The traditional deep division between Catholics and Protestants remains critical to the Good Friday Peace Agreement in Northern Ireland, for example, and the imprint of this cleavage continues to mark party politics in many nation-states in continental Europe. In Britain, by contrast, although many people continue to express a religious affiliation, the political difference between Protestants and Catholics, which was critical to British politics from the Reformation until the late-nineteenth century, gradually faded as a salient party cleavage and electoral cue during the early twentieth century.\(^2\) Unfortunately, therefore, the MAR data also suffers from serious problems of systematic selection bias, by monitoring contemporary cases of ethno-political conflict but excluding the most successful cases of ethnic accommodation, where minorities have been politically integrated or assimilated into the majority population.

Gender and Ethnicity

Rather than examining direct indicators, an alternative strategy for studying ethnic minority representation seeks to generalize by analogies with the experience of the role of women in elected office, where this is regarded as a proxy indicator of ‘minority’ representation in general.\(^3\) Certainly both women and ethnic minorities are commonly some of the most underrepresented groups in legislative office in most established democracies.\(^4\) A considerable body of evidence has also now accumulated suggesting that female representation is...
commonly greater under PR party lists compared with majoritarian electoral systems. But is it legitimate to generalize from the experience of women to the representation of ethnic minorities? In fact, there are many reasons why this strategy may prove seriously flawed. After reviewing the literature, Bird concludes that substantial differences exist in the reasons underlying the underrepresentation of women and ethnic minorities.\footnote{In particular, concentrated ethnic communities are clustered geographically within certain areas, allowing territorial groups to make local gains in particular minority constituencies within majoritarian electoral systems, even in heterogeneous plural societies; for example, African-American representatives are elected to Congress in minority-majority districts in New York City, Detroit, and Los Angeles.} By contrast, the male-to-female ratio in the population is usually fairly uniformly distributed across different electoral constituencies. Htun also points out that positive action strategies often differ substantially in the opportunities they provide for women and ethnic minorities. Statutory gender quotas mandating the minimal proportion of women that parties adopt as legislative candidates are common in many established democracies, for example, although none of these nation-states have adopted similar quota laws to compensate for the underrepresentation of ethnic minority candidates. By contrast, reserved seats in parliament are used for ethnic, racial, national, linguistic, or religious communities in more than two dozen countries, including guaranteeing the inclusion of specified minorities in legislatures through the creation of separate electoral rolls, the allocation of special electoral districts, or provisions for direct appointment to the legislature. Most countries adopting these mechanisms have majoritarian electoral systems which are unable to guarantee adequate minority representation through other means. For all these reasons, direct comparisons of women and ethnic minority representation are a flawed research strategy, and alternative approaches are needed to explore the cross-national evidence on a more systematic basis.

Cross-Sectional Large-N Comparisons

Another common large-N approach within the quantitative tradition has sought to classify political institutions in many countries worldwide, developing typologies of electoral systems, public sector bureaucracies, patterns of interest group politics, party systems, or presidential and parliamentary executives, and then comparing how far the contemporary performance of democracy varies systematically by institutional types. Ever since Aristotle, analytical typologies have always been a vital part of comparative politics, with recent attempts to classify membership of each type using the logic of fuzzy sets. By selecting a larger range of cases using a random sample of countries, or the universe of independent nation-states worldwide, the cross-sectional comparative strategy overcomes many of the potential problems of selection bias which limit case-study approaches, but it also raises important methodological questions. The next chapter discusses the challenge of establishing appropriate operational
measures of core concepts, notably how we develop indicators of democracy which are generally regarded as valid and reliable, and which are genuinely independent from the variables thought to explain democracy. Serious difficulties also relate to many other aspects which are central to comparative studies of the relationship between institutions and democratization, including measures of ethnic fractionalization, federalism, and corruption. In addition, cross-sectional comparisons taken at one point in time face the challenge of establishing the endogeneity of political institutions.

Political institutions can be understood as the formal rules and the informal social norms which structure the workings of any regime. Formal rules include the legislative framework governing constitutions, as embodied in official documents, constitutional conventions, legal statutes, codes of conduct, and administrative procedures, authorized by law and enforceable by courts. Institutions provide incentives and sanctions which constrain human behavior. It is neither necessary nor sufficient for rules to be embodied in the legal system to be effective; social norms, informal patterns of behavior, and social sanctions also create shared mutual expectations among political actors. Nevertheless this study focuses most attention upon the formal rules as these represent the core instruments of public policy. These rules are open to reform and amendment by the political process, whether by legislation, executive order, constitutional revision, administrative decision, judicial judgment, or bureaucratic decree. Although there is a ‘gray’ overlapping area, by contrast most social norms are altered gradually by informal processes such as social pressures, media campaigns, and cultural value shifts located outside the formal political arena. Drawing upon the distinction first suggested by Duverger, institutional changes to the formal rules can be understood to exert both mechanical effects, including those brought about by the legal system, as well as psychological effects, shaping social norms and beliefs. In societies characterized by rule of law, the impact of mechanical effects can be expected to be manifested relatively fast, for example, the introduction of a lower formal vote threshold for elections to the legislature. By contrast, the psychological effects of such a legal change may only become evident over a long-term process, as people gradually learn to adapt their behavior to the new rules, for example, in their perceptions about whether casting a ballot for a minor party represents a wasted vote.

If institutions are regarded as relatively durable, exemplified by the basic division of powers between the executive and legislative branches established in the US Constitution, then the formal rules governing the political system can be understood as shaping the incentives governing the behavior of lawmakers and presidents. Nevertheless there are also complex issues of reciprocal causality; political actors make the rules which then tie their hands. The problem of institutional endogeneity is least serious in established democracies which have not altered their basic constitutional arrangements for many decades, perhaps for more than a century, even if there are a series of more minor adjustments in administrative procedures and political roles. In this regard, institutions can be regarded as persistent constraints influencing subsequent patterns of
political behavior and social norms. The difficulty of disentangling cause and effect arises with greater urgency in established democracies which have experienced more fundamental reforms, such as devolution in Scotland and Wales, decentralization in Catalonia and Galicia, and electoral reform toward combined systems in New Zealand and Italy. Here the conventional story suggests, for example, in Scotland, that growing nationalist identities in the postwar era led gradually toward the creation of the Scottish Parliament in 1999, but that in a reciprocal effect, devolution has also served to slightly strengthen Scottish identities.\footnote{Similarly, the standard account emphasizes that discontent with the power of the executive generated pressures for electoral reform in New Zealand, and, once the electoral system altered from first-past-the-post (FPTP) to mixed member proportional (MMP), this facilitated fragmentation in the party system.} Attention to developments over time is critical to understanding the causal process under analysis.

The problem of treating political institutions as endogenous factors causing democratization is most acute in countries which have undergone fundamental constitutional changes where institutions are not sticky, including in many newer democracies. As Colomer emphasizes, parties make electoral rules.\footnote{It follows that plural societies engaging multiple parties and factions in peace-negotiations, for example, are probably more likely to adopt PR electoral systems and federal power-sharing, since this arrangement makes establishing agreement easier among many participants and minor parties.} If we assume that political actors are risk-averse, and that new rules are established under conditions of considerable uncertainty about their outcome, power-sharing agreements reduce the risks of losing.\footnote{Given their recent history and experience of internal conflict, some countries using power-sharing institutions may therefore have a far worse historical legacy of political stability and internal conflict, a pattern which spills over into contemporary politics, in comparison with homogeneous societies which may have adopted power-concentrating rules. Although there is a growing literature on the causes of electoral system change, the reasons why countries select one or another constitutional arrangement remain poorly understood. Rational calculations about the potential consequences of these choices are hampered by the limited information and uncertainty surrounding the consequences of each constitutional option. A ‘mixed scanning’ approach may be adopted, where policymakers consider a limited range of constitutional options, based on past histories and comparisons with neighboring states. Countries continue to reflect an imprint from their colonial legacies; for example, about 60% of ex-British countries adopted first-past-the-post for the lower house of parliament, while about the same proportion of ex-Spanish colonies adopted PR.} Cultural proximity, influential regional models, and learning across national borders are also part of the process, as illustrated by the way that many Latin American countries adopted presidential executives, while by contrast Central European states were more likely to adopt parliamentary systems.
Problems of establishing the direction of causality are equally severe when understanding the relationship between ethnic identities and political institutions. On the one hand, essentialist perspectives assume that ethnic identities reflect fixed and stable psychological orientations. These assumptions lie at the heart of the consociational thesis that political institutions will reduce ethnic tensions most effectively by incorporating stakeholders from rival communities into decision-making processes. Ethnic identities and separate communal boundaries are regarded in consociational arguments as the enduring social base upon which political institutions arise to form the superstructure. The boundaries drawn between Protestant and Catholic communities in Belfast, black-white racial tensions in Los Angeles, or violent conflict between Arab militias and black African farmers in Darfur, for example, are seen to reflect long-standing cultural dissimilarities which precede the construction of the modern nation-state, making them difficult to alter and manipulate, for example, through assimilation policies. Hence the logic of the post-Dayton peace agreement which sought to balance the interests of the Bosniak, Serbian, and Croatian communities by a careful allocation of seats in the new parliament governing Bosnia and Herzegovina, for example, with the office of the speaker rotating among members drawn from each community.

By contrast, constructivist perspectives challenge the basic assumption that ethnic identities and community boundaries are fixed and stable. This view emphasizes that the design of political institutions and public policies can either heighten or weaken latent ethnic identities, so that social cleavages based on language, religion, race, or nationality should not be regarded as endogenous to the political system. In Serbia, for example, Gagnon shows how Slobodan Milosevic was elected to office on a public platform which deepened ethnic hatred and intercommunity rivalry, manipulating symbols and myths to maximize popular support. Snyder argues that elites in democratizing states have incentives to manipulate and exploit nationalist conflict, a process exacerbated by electoral pressures. Colonial powers have drawn boundaries creating nation-states and provinces which cut across or which reflect existing community boundaries. Byman argues that policies of assimilation are capable of breaking down communal boundaries, although they also conflict with recognition of minority rights, for example, when Turkey attempts to repress the Kurdish language in all schools and mass media, or where states attempt to encourage more positive opportunities for minority integration into the majority culture. Given the constructivist perspective, studies examining the relationship between ethnic identities and constitutional arrangements therefore need acknowledge the potential problem of treating the former as endogenous to the political system. To disentangle the direction of causality involved in this relationship, ideally time-series data is needed, as well as historical case studies tracing the dynamics of the process of democratization, institutional change, and the waxing and waning strength of ethnic identities within particular countries.
THE MIXED RESEARCH DESIGN

The unresolved debate between consociationalism and its critics has therefore received widespread attention in the extensive body of scholarly research, as well as within the policymaking community. The cases of Benin and Togo suggest the need to explore these issues further. Subsequent chapters go on to compare countries around the globe more systematically to see whether institutions matter for democracy, controlling for a wide range of social and cultural factors which may also contribute toward democratization. To overcome the shortcomings of each of the methods already discussed, this study supplements large-N quantitative comparisons with in-depth paired case studies.

**Time-Series Cross-Section (CSTS) Data**

The period of analysis focuses upon patterns of regime change occurring during the third wave of democratization, for 30 years from 1973 to 2004 (or the end point of the series, if slightly earlier). The research design relies on cross-sectional time-series (CSTS) data facilitating comparison over space and time. This sort of data consists of repeated observations on a series of fixed (nonsampled) units (all contemporary nation-states worldwide), where the units are of interest in themselves. The country-year is the basic unit of analysis, generating 5,680 observations across the whole period. This produces a large enough time-series to model the dynamics for each unit. The comparison covers 191 contemporary nation-states (excluding dependent territories and states which dissolved during this period).

The analysis of cross-sectional longitudinal data needs to address certain important challenges and analysis of any panel datasets raises complex issues. Ordinary least squares regression estimates assume that errors are independent, normally distributed, and with constant variance. This type of data violates these assumptions and raises potential problems of heteroskedasticity, autocorrelation, robustness, and missing data. Heteroskedasticity is produced if the range of variations in the scatter of nation-states around the regression line is not uniform across different levels of democracy. Autocorrelations are generated because, with time-series data, the same countries are being counted repeatedly and the additional observations do not provide substantially new information. The danger is that the study will fail to identify the lack of independence between cases and will subsequently reach false conclusions. The beta coefficients in any regression analysis will remain unbiased but the disturbance terms from the errors (i.e., omitted variables) are likely to be correlated. The use of ordinary least squares regression models leads to estimates of standard errors, used for evaluating the significance of the relationship, which are less accurately measured than they appear. In other words, the significance of any coefficients will be inflated, suggesting that significant relationships exist when they do not. Various options are available to overcome the problem of both autocorrelated and heteroskedastic disturbances found in cross-sectional...
time-series datasets, such as feasible generalized least squares that estimate errors with an ARH\textsubscript{1} model, or the use of robust regression.

Following the advice of Betz and Katz, when comparing relationships across countries, the study uses ordinary least squares (OLS) linear regression with panel-corrected standard errors (PCSE) to measure the impact of the independent variables on levels of democratization across each nation.\textsuperscript{43} This approach is particularly suited to the dataset as the number of countries under comparison (\(N = 191\)) is far greater than the number of years (\(T = 30\)). Nevertheless this form of data analysis needs considerable care to double-check the robustness of the results given different specification issues, alternative models, and diagnostic tests.\textsuperscript{46} When computing the standard errors and the variance-covariance estimates, OLS regression analysis models with PCSE assume that disturbances are, by default, heteroskedastic and contemporaneously correlated across panels. One important advantage of this approach is that the results are relatively easily interpretable, as OLS regression is the most widely familiar statistical technique in the social sciences. The use of panel-corrected standard errors is the most appropriate approach where the data contains all contemporary countries worldwide, rather than a sample of countries drawn from a larger universe, where estimating the random effects may be more suitable.\textsuperscript{47} Moreover with a large time-series, the results of fixed and random effects models usually converge, so there is no substantial difference in the use of either approach. The use of fixed effects has its costs, however, since it forces us to drop any independent variables from the model that are unchanging attributes of each country, such as region. Fixed effects models also make it hard for any slowly changing variables to appear substantively or statistically significant, making a rigorous test for estimating the impact of any institutional reforms.\textsuperscript{48}

To check the robustness of the results under different specifications, the regression analysis models are presented for the Freedom House data. The results are then replicated using the other three standard indicators of democracy discussed in the next chapter, provided by Polity IV, Vanhanen, and Cheibub and Gandhi, to examine whether the key findings are confirmed and the results hold irrespective of the particular measure selected for analysis. The large-N comparison produces systematic patterns which hold across many instances but this process loses the depth which is derived from examining the process of democratization through intensive case studies. Nevertheless all regression analysis is subject to certain well-established pitfalls, such as unit heterogeneity, autocorrelation, and the use of different time lags, and conclusions relying solely upon this method would not be wholly satisfactory or convincing.

Case Studies

To put more flesh on the bones, the large-scale econometric analysis is supplemented by selected case-study narratives, illuminating the dynamics of regime change from a more qualitative perspective. No single method or technique
can be regarded as wholly convincing but the combination of approaches provides a more plausible story. To illustrate the theory and analysis with more concrete examples, paired cases are selected within each region to compare leaders and laggards in the process of democratization during the third wave. This process is exemplified by the West African cases already described briefly, including the contrasts between the end of military rule and the transition to stable multiparty democracy and a free press experienced in Benin, compared with persistent repression found in Togo. Beninese and Togolese shared a common history for many centuries and both countries remain among the poorest nation-states in the world. Both societies are divided ethnically by cleavages of religion and language. Despite all that these neighboring states share, during the last decade Benin has established the conditions for the peaceful rotation of governing power, while Togo remains an unreconstructed autocracy. Similar paired case-study comparisons are discussed to illustrate contrasting political developments in other regions, including contrasts among Costa Rica and Venezuela, Taiwan and Singapore, Hungary and Belarus, and Turkey and Egypt.

Classifying Institutions

The comparison also needs to classify the core institutions at the heart of consociationalism. In practice, executive power-sharing can take many institutional forms, making the ideal type notoriously difficult to test empirically; for example, it may involve a coalition of ethnic parties in cabinet (as in South Africa), the allocation of ministerial portfolios based on explicit recognition of major religious or linguistic groups (as in Belgium), a presidency made up of a committee of three representing each nation, with a rotating chair (as in Bosnia and Herzegovina), or (as in Lebanon) the division of the presidency (Maronite Christian), prime minister (Sunni Muslim), and speaker (Shi’a Muslim). Other arrangements used to secure the election of minorities to the legislature, even within majoritarian electoral systems, include reserved seats (used in New Zealand), overrepresentation of minority districts (such as smaller electoral quota used for Scottish constituencies at Westminster), and minority redistricting (exemplified in the United States). Territorial autonomy can also take multiple complex forms, with the powers and responsibilities for services such as education, taxation, and domestic security divided among multiple layers of government and administrative units.

In *Democracies* and *Patterns of Democracy*, Lijphart refined and built upon his earlier work in the attempt to operationalize consociationalism as an ideal type and to classify established democracies into two categories: ‘consensus’ (power-sharing) or ‘majoritarian’ (power-concentrating) democracies. The major institutions are understood to cluster into two main dimensions. The ‘parties-executive’ dimension for consensus democracies rests on the existence of proportional representation elections, multiparty competition, coalition governments, executive-legislative balance, and interest group corporatism. The
‘federal-unitary’ dimension for consensus democracies includes federalism and decentralization, balanced bicameralism, constitutional rigidity, judicial review, and central bank independence. Rather than attempting to test the impact of all the institutions of consensus democracy, this study focuses upon comparing the role of four important pillars, namely, proportional representation electoral systems (leading, in turn, toward multiparty legislatures and coalition governments), the type of parliamentary or presidential executive (leading toward horizontal patterns of unified or divided government), federalism and decentralization (leading toward vertical patterns of regional autonomy and the protection of minority rights for territorial groups), and the regulation of channels of political communication (as one of the primary channels available for opposition movements in civil society). Many other institutional reforms have been associated with strengthening democracy; for example, beyond the electoral process, Carothers notes that the standard democracy assistance template usually includes improving the integrity of governance and reducing corruption; expanding the capacity of public sector management; strengthening the rule of law, human rights, and the role of an independent judiciary; and building an active range of civil society organizations and NGOs. These are all potentially important reforms which are worth exploring in a broader study of the full impact of democracy assistance, but this study focuses upon a narrower set of institutions. Reforms designed to achieve power-sharing in executive-legislative relations and in the decentralization of decision-making are likely to have a strong impact on promoting democratic consolidation in divided societies. The conceptualization, classification, and measurement of each of the main types of institutions compared in this study are discussed in subsequent chapters.

Some have sought to examine the impact of power-sharing arrangements by constructing single indicators. The concept of ‘veto points’, developed most fully by Tsebelis, is an attempt to identify the functionally equivalent checks and balances within any political system and to aggregate these into a single index. Linder and Bachtiger adopt a similar measure, the power-sharing index, which includes a horizontal dimension through the use of proportional representation electoral systems as well as institutional veto points (such as a strong separation of powers and multiparty governing coalitions) and a vertical dimension based on the degree of federalism. The study found that the horizontal dimension of power-sharing, but not the vertical dimension, was more strongly related to democratization in Africa and Asia than economic development. These combined measures may be useful for some analytic objectives, but they can be problematic from a policymaking perspective, since any analysis of the impact of power-sharing arrangements on democracy will be unable to identify the specific institutions which produced these effects. The measure may be particularly sensitive to the aggregation process and it is not at all obvious how best to proceed in this regard, for example, whether a political system with a strict division between the judiciary and executive is functionally equivalent in power-sharing terms to another system which has minority...
vetoes guaranteed in the constitution. Both provide a form of power-sharing but how should these rules be counted? Another potential problem is that the veto point indicators may be measuring some of the same components that are used for the construction of certain indicators of democracy, such as the criteria used by Polity IV, contaminating the independent and dependent variables. By contrast, examining the separate effects of the types of electoral systems, executives, federalism, and the mass media on democracy provides a cleaner and more policy-relevant interpretation of the results of the analysis. These institutions are also independent of the selected indicators of democracy. For example, as discussed in the next chapter, although Polity IV incorporates the use of elite competition or direct elections as part of its measure of limitations on executive authority, the project does not monitor the particular type of electoral system used for this purpose. In the same way, Cheibub and Gandhi define regimes as democratic if those who govern are selected through contested elections, but the type of electoral rules, whether majoritarian or proportional, is not part of their measure.

Measuring Ethnic Fractionalization

One of the most complex issues facing empirical research on consociationalism concerns the most appropriate concept and measurement of ethnic fractionalization. Cross-national studies of the evidence have been hindered by the difficulties of establishing robust and consistent measures of ethnic identities that are applicable across many different types of societies. States often contain multiple cultural cleavages and forms of social identity, some overlapping, and studies need to choose the one that is most salient politically and most relevant theoretically to the issue under consideration. Ethnic groups are defined here as a community bound by a belief in common ancestry and cultural practices, whether based on religion, language, history, or other cultural customs and ties. Debate continues to surround the origins and nature of ethnic identities. The essentialist perspective regards ethnic identities as largely fixed at birth or in early childhood, due to the physical characteristics of groups, such as their racial skin color or facial features, or based on enduring social conventions and cultural norms. By contrast, as discussed earlier, the constructivist perspective regards ethnic identities as socially created, where the salience of alternative identities is open to manipulation. In this view, community differences can be exacerbated (for example, where politicians preach the heated rhetoric of ethnic hatred and nationalism to maintain their popularity) or ameliorated (where there are successful attempts to assimilate groups). In practice, there are substantial difficulties in comparing ethnic identities across nation-states. Nigeria, for example, contains an estimated 250 tribal groups, as well as sharp regional divisions between northern Muslims and southern Christians. Language is important in Switzerland, which splits into the predominant German and the minority French- and minority Italian-speaking regions. Belgium divides into the Fleming and Walloon segments. By contrast, the United Kingdom divides
Evidence and Methods

by nationalist identities into England, Scotland, Wales, and Northern Ireland (the latter subdivided into Protestant and Catholic communities), while the United States uses the idea of racial characteristics and language as the defining basis of the major ethnic cleavages. Given that the meaning and form of identities are so culturally diverse, it remains unclear whether cross-national studies can compare like with like, or whether they can even compare functionally equivalent groups, across societies.

Moreover the available data to estimate ethnic identities is often limited and unreliable. Aggregate sources drawn from official population censuses, household surveys, and general social surveys facilitate analysis of the distribution of religious, linguistic, national origin, or racial groups in each country. But not all surveys seek to incorporate these items, in part because of their cultural sensitivity, and, unlike measures of occupational class and socioeconomic status, no standard international practices maintain consistency across sources. Where religious, linguistic, racial, national, or other forms of ethnic identities are systematically monitored in official surveys, the data usually allows us to monitor the distribution of these populations, but far fewer survey questions seek to measure the salience or meaning of these identities. What may matter is less ethnic fractionalization, as commonly measured, than ethnic polarization, meaning the distance between groups. Clearly, given these limitations, the measurement of ethnic diversity needs further research.

Many previous studies of linguistic cleavages have also had to rely upon badly flawed aggregate sources, exemplified by the Soviet-era Atlas Narodov Mira (1964), the original dataset used to construct the ethno-linguistic fractionalization (ELF) index. The index has been widely employed by economists, following Easterly and Levine’s study showing that economic growth was negatively related to ethno-linguistic fractionalization. ELF is computed as 1 minus the Herfindahl index of ethno-linguistic group shares, and it estimates the probability that two randomly selected individuals from a population belong to different groups. Nevertheless it is now recognized that the atlas contains some basic coding inaccuracies, the material is also badly dated, and the linguistic cleavage represents only one dimension of ethnic identities, and not necessarily the most important one.

This project compares systematic cross-national evidence worldwide to classify nation-states according to the degree of ethno-linguistic fractionalization, based on a global dataset created by Alesina and his colleagues. This study classifies 201 countries or dependencies on the basis of the share of the population speaking each language as their ‘mother tongues’. The dataset was usually derived from census data, as collated in most cases by the Encyclopaedia Britannica, with a few cases of missing data supplemented by the CIA World Factbook. Religious fractionalization is also calculated for 215 nation-states and dependencies, drawing upon the same sources. The relationship between these indicators, and the distribution of countries under comparison, is illustrated by the scatter-plot in Figure 2.1. As expected, certain societies in the top-right quadrant emerge as highly heterogeneous on both measures, notably many
FIGURE 2.1. Linguistic and Religious Fractionalization. Note: For the linguistic and religious fractionalization indices, see Alesina et al. 2003. For the classification of the major type of electoral system, see Norris 2004.
sub-Saharan African nation-states (such as South Africa, Uganda, Kenya, and Nigeria). At the same time, many Arab states such as Saudi Arabia, Libya, and Jordan, located in the bottom-left quadrant, are very homogeneous according to these indicators. Because the underlying data sources used for constructing these indicators are fairly imprecise, and they depend heavily upon the categorization scheme used and the underlying population estimates, modest differences in the position of countries on the ethnic fractionalization indices are probably unreliable. Nevertheless the indices can be used to make broad classifications and each is therefore dichotomized, with scores from 50 and above on the 100-point scales defined as plural societies.

It should be noted that this index estimates the objective distribution of different linguistic and religious groups in the population, but it does not seek to measure the subjective meaning or societal importance of these forms of identity. In this regard, analogies can be drawn between the ‘objective’ indicators of occupation and income used to gauge socioeconomic status and the ‘subjective’ identifications which respondents offer when asked in surveys where they feel that they belong in terms of social class. In the same way, ‘objective’ and ‘subjective’ indicators of ethnic identities may coincide or they may differ sharply. For example, Canada is classified as relatively heterogeneous in religion, divided between Protestants and Catholics, but as this society has become fairly secular, these forms of formal religious identity may carry few significant consequences beyond the completion of official forms for the government census or birth certificates. On the other hand, France is classified as fairly homogeneous in religious identities, as it remains predominantly Catholic, although the expression of religious identities has aroused heated debate in recent years, exemplified by legal bans passed against the wearing of Muslim headscarves in schools. The evidence within this study only seeks to compare the existence of objective indicators of religious and linguistic cleavages, not their subjective salience. A constructivist perspective emphasizes that people possess multiple social identities, and the salience of these latent characteristics may rise or fall in response to situational factors, including the role of parties and politicians competing for power, and how far they ‘play the ethnic card’ in whipping up ethnic hatred through manipulating symbols and myths. These are important insights but nevertheless the strength of ethnic identities remains extremely difficult to measure and compare with any degree of accuracy in the absence of representative surveys, especially any changes occurring over time in response to the cues provided by political elites, and hence beyond the limitations of this study. On this basis, we can go on to consider the alternative indicators of democracy and what these reveal about trends in democratization during the third wave.