Economic Origins of Dictatorship and Democracy

DARON ACEMOGLU

JAMES A. ROBINSON

CAMBRIDGE
UNIVERSITY PRESS

1 Paths of Political Development

To understand why some countries are democracies whereas others are not, it is useful to distinguish between different characteristic paths that political institutions take over time. Only some of these paths end in democracy, at least at this moment in time. These stylized paths help us to orient ourselves among the complexities of real-world comparisons, and they illustrate the main mechanisms that we believe link the economic and political structure of a society to political institutions.

There are four main paths of political development. First, there is a path that leads from nondemocracy gradually but inexorably to democracy. Once created, democracy is never threatened, and it endures and consolidates. Britain is the best example of such a path of political development. Second, there is a path that leads to democracy but where democracy, once created, quickly collapses. Following this, the forces that led to the initial democratization reassert themselves, but then democracy collapses again and the cycle repeats itself. This path – where democracy, once created, remains unconsolidated – is best exemplified by the Argentinian experience during the twentieth century. Logically, a third path is one in which a country remains nondemocratic or democratization is much delayed. Because there are important variations in the origins of such a path, it is useful to split nondemocratic paths into two. In the first path, democracy is never created because society is relatively egalitarian and prosperous, which makes the nondemocratic political status quo stable. The system is not challenged because people are sufficiently satisfied under the existing political institutions. Singapore is the society whose political dynamics we characterize in this way. In the second of these nondemocratic paths, the opposite situation arises. Society is highly unequal and exploitative, which makes the prospect of democracy so threatening to political elites that they use all means possible, including violence and repression, to avoid it. South Africa, before the collapse of the apartheid regime, is our canonical example of such a path.

In this chapter, we illustrate these four paths and the mechanisms that lead a society to be on one or the other by examining the political history of the four

countries. We discuss the dynamics of political development in all cases, exploring why they ended in consolidated democracy in Britain, unconsolidated democracy in Argentina, and persistent nondemocracy – albeit of different forms – in Singapore and South Africa. Our discussion highlights many of the factors that subsequent analysis will show to be crucial in determining why a society moves onto one path rather than another.

1. Britain

The origins of democracy in Britain lie with the creation of regular Parliaments that were a forum for the aristocracy to negotiate taxes and discuss policies with the king. It was only after the Glorious Revolution of 1688 that Parliaments met regularly, and they did so with a very restrictive franchise. The membership of Parliament at this stage was inherited from feudal notions about the existence of different "estates" in society. These orders were the clergy and the aristocracy, who sat in the House of Lords by right, and the commons, who sat in the House of Commons. Members of the Commons were, in principle, subject to elections, although from the eighteenth century through the middle of the nineteenth century, most elections were unopposed so that no voting actually took place (Lang 1999, p. 12). Candidates tended to be proposed by the leading landowners or aristocrats and, because there was no secret ballot and voting was open and readily observed, most voters did not dare go against their wishes (Namier 1961, p. 83; Jennings 1961, p. 81).

Nevertheless, the constitutional changes that took place following the Civil War of 1642–51 and Glorious Revolution of 1688 led to a dramatic change in political and economic institutions that had important implications for the future of democracy (North and Thomas 1973; North and Weingast 1989; O'Brien 1993; Acemoglu, Johnson, and Robinson 2005). These changes emerged out of conflict between the Stuart monarchs intent on maintaining and expanding their absolutist powers and a Parliament intent on reigning them in. Parliament won. The outcome was a restructuring of political institutions that severely limited the monarchy's powers and correspondingly increased those of Parliament. The change in political institutions led to much greater security of property rights because people no longer feared predation by the state. In particular, it placed power into the hands of a Parliament in which was represented merchants and landowners oriented toward sale for the market. By the late eighteenth century, sustained economic growth had begun in Britain.

The first important move toward democracy in Britain was the First Reform Act of 1832. This act removed many of the worst inequities under the old electoral system, in particular the "rotten boroughs" where several members of Parliament were elected by very few voters. The 1832 reform also established the right to vote based uniformly on the basis of property and income.

The First Reform Act was passed in the context of rising popular discontent at the existing political status quo in Britain. Lang (1999, p. 26) notes

> Fear of revolution, seen as a particular risk given the growth of the new industrial areas, grew rather than diminished in the years after Waterloo, and Lord Liverpool's government (1821–1827) resorted to a policy of strict repression.

By the early nineteenth century, the Industrial Revolution was well underway, and the decade prior to 1832 saw continual rioting and popular unrest. Notable were the Luddite Riots of 1811–16, the Spa Fields Riots of 1816, the Peterloo Massacre in 1819, and the Swing Riots of 1830 (see Darvall 1934 and Stevenson 1979 for overviews). Another catalyst for the reforms was the July revolution of 1830 in Paris. The consensus among historians is that the motive for the 1832 reform was to avoid social disturbances. Lang (1999, p. 36) concludes that

> the level of unrest reinforced the case for immediate reform now, rather than later: it was simply too dangerous to delay any longer. Just as Wellington and Peel had granted emancipation to avoid a rising in Ireland, so the Whigs . . . should grant reform as the lesser of two evils.

The 1832 Reform Act increased the total electorate from 492,700 to 806,000, which represented about 14.5 percent of the adult male population. Yet, the majority of British people could not vote, and the aristocracy and large landowners had considerable scope for patronage because 123 constituencies contained fewer than one thousand voters. There is also evidence of continued corruption and intimidation of voters until the Ballot Act of 1872 and the Corrupt and Illegal Practices Act of 1883. The Reform Act, therefore, did not create mass democracy but rather was designed as a strategic concession. Unsurprisingly, the issue of parliamentary reform was still very much alive after 1832, and it was taken up centrally by the Chartist movement.

Momentum for reform finally came to a head in 1867, largely due to a juxtaposition of factors. Among these was a sharp business-cycle downturn that caused significant economic hardship and increased the threat of violence. Also significant was the founding of the National Reform Union in 1864 and the Reform League in 1865, and the Hyde Park Riots of July 1866 provided the most immediate catalyst. Searle (1993, p. 225) argues that

> Reform agitation in the country clearly did much to persuade the Derby ministry that a Reform Bill, any Reform Bill, should be placed on the statute book with a minimum of delay.

This interpretation is supported by many other historians (e.g., Trevelyan 1937; Harrison 1965).

The Second Reform Act was passed in 1867; the total electorate expanded from 1.36 million to 2.48 million, and working class voters became the majority in

all urban constituencies. The electorate was doubled again by the Third Reform Act of 1884, which extended the same voting regulations that already existed in the boroughs (urban constituencies) to the counties (rural constituencies). The Redistribution Act of 1885 removed many remaining inequalities in the distribution of seats and, from this point on, Britain only had single-member electoral constituencies (previously, many constituencies had elected two members – the two candidates who gained the most votes). After 1884, about 60 percent of adult males were enfranchised. Once again, social disorder appears to have been an important factor behind the 1884 act (e.g., Hayes 1982; Lang 1999, p. 114).

Following the Great War, the Representation of the People Act of 1918 gave the vote to all adult males over the age of twenty-one and women over the age of thirty who were ratepayers or married to ratepayers. Finally, all women received the vote on the same terms as men in 1928. The measures of 1918 were negotiated during the war and may reflect to some extent a quid pro quo between the government and the working classes who were needed to fight and produce munitions. Garrard (2002, p. 69) nevertheless notes that

> most assumed that, if the system was to survive and "contentment and stability prevail," universal citizenship could not be denied to men, perceived to have suffered so much and to have noticed Russia's Revolution.

Overall, the picture that emerges from British political history is clear. Beginning in 1832, when Britain was governed by the relatively rich, primarily rural aristocracy, strategic concessions were made during an eighty-six-year period to adult men. These concessions were aimed at incorporating the previously disenfranchised into politics because the alternative was seen to be social unrest, chaos, and possibly revolution. The concessions were gradual because, in 1832, social peace could be purchased by buying off the middle classes. Moreover, the effect of the concessions was diluted by the specific details of political institutions, particularly the continuing unrepresentative nature of the House of Lords. Although challenged during the 1832 reforms, the House of Lords provided an important bulwark for the wealthy against the potential of radical reforms emanating from a democratized House of Commons. This was so at least until just before the First World War, when the showdown with Herbert Asquith's Liberal government over the introduction of elements of a welfare state led to substantial limitations of the power of the Lords. After 1832, as the working classes reorganized through the Chartist movement and later the trade unions, further concessions had to be made. The Great War and the fallout from it sealed the final offer of full democracy. Although the pressure of the disenfranchised was more influential in some reforms than others, and other factors undoubtedly played a role, the threat of social disorder was the driving force behind the creation of democracy in Britain.

The emergence of democracy in Britain and its subsequent consolidation took place in a society that had long shed nearly all the remnants of medieval organization and that had successfully resisted the threat of absolutism. They also took place

in the context of rapid industrialization, urbanization, expansion of the factory system, rising inequality, and – in the period after the Repeal of the Corn Laws – rapid globalization of the economy.

2. Argentina

The beginnings of the modern Argentine Republic were in 1810 when it declared its independence. Following this period, the country was immersed in a chaotic series of civil wars and internal conflict over the structure of power and political institutions. The chaos finally abated in the 1860s. In 1853, a new constitution was written and, in 1862, Bartolomé Mitre was elected the first president of the unified republic. Mitre set about creating a state in the facilitating context of the first of a series of agricultural export booms that would sustain the Argentine economy until 1930. He created a national bureaucracy, taxation system, and legal system, and this period saw the foundation of electoral politics. However,

> The electoral law of 1853, which purported to allow popular participation in the political process, from the beginning proved itself a sham. Elections were invariably ritualistic parodies, staged-managed by lackeys of the powerful, with only a minute fraction of the electorate participating. (Rock 1987, p. 129)

After Mitre, Domingo Sarmiento became president and around him formed a party, the Partido Autonomista Nacional (PAN). Successive PAN presidents maintained power until 1916 by manipulating elections. However, they did so in the context of rising social discontent. After 1889, there was an effective opposition in the Unión Cívica, which in July 1890 launched a revolt against the government. After 1891, the Unión Cívica Radical (Radicals), under the leadership of Hipólito Yrigoyen, launched revolts in 1893 and 1905. However, despite the continuation of regimes based on the control and coercion of the electorate,

> Argentine elites were becoming aware of the unfolding similarities between Western European societies and their own, with the growing cities and the emergence of new social classes. Democracy's attractiveness lay in its promise of protecting political stability, for if political exclusion were maintained . . . the nation risked a repetition of the upheavals of the early 1890's. (Rock 1987, pp. 184–5)

In 1910, Roque Sáenz Peña, one of the leading advocates of political reform, became president. As Rock (1987, p. 188) put it:

> Radicals, socialists, and indirectly the anarchists helped fuel the movement for reform during the early years of the century. Progressives amongst the elite feared the growing popular support for the Radicals, wondering where their next revolt would come from.

The so-called Sáenz Peña Law was passed in 1912 when the secret ballot was introduced and fraudulent electoral practices outlawed. Universal male suffrage, originally introduced in the 1853 Constitution, finally became a reality. Smith

(1978, p. 10) argues that reform "was a calculated maneuver to salvage the prevailing system. Concerned with labor unrest and the apparent threat of violence."

Following these reforms, Yrigoyen was elected president in 1916.

> The reforms also brought surprises. Sáenz Peña and his supporters had espoused electoral reform in the belief that the old oligarchic factions would adapt to the new conditions and unite into a strong conservative party that would enjoy large popular support . . . instead, the conservatives repeatedly failed in their efforts at unity. (Rock 1987, p. 190)

As a consequence, the Radical party began to dominate Argentine politics, posing a severe threat to traditional interests. In 1916, Conservatives won 42 percent of the vote but by 1928 they had slipped to 25 percent. Smith (1978, p. 21) notes "this situation contrasts sharply with that in Sweden and Britain . . . where traditional elites continued to dominate systems after the extension of suffrage." Consequently, "by 1930 Yrigoyenists had a substantial delegation in the upper chamber and they threatened to gain a full majority in the upcoming elections" (Smith, 1978, p. 12). Thus, "the political system came to represent an autonomous threat to the socioeconomic system . . . Understandably enough, in view of their initial expectations, Conservatives came to see democracy as dysfunctional" (Smith 1978, p. 15; see also Potter 1981).

In September 1930, Yrigoyen was deposed by a military coup, followed in 1931 by a fraudulent election. "The election of 1931 restored power to the same broad complexion of groups that had controlled it before 1916 – the pampas' exporting interests and the lesser landowners of the provinces" (Rock 1987, p. 217). During the remainder of the 1930s, Conservatives continually used electoral fraud to maintain power, although by 1940 they were trying to reincorporate the Radicals to some extent. This sequence of Conservative administrations was ended by a military coup in 1943.

After the coup in 1943, a series of military men assumed the presidency; however, the main feature of this period was the rise to power of Juan Domingo Perón, first as a member of the military junta and then as the elected president after 1946. Perón had moved the military regime onto a more radical and pro-labor path and organized a political machine around the state control of the labor movement. During his first presidency, Perón engineered a huge increase in wages and social benefits for the working classes. His policies were aimed at redistributing away from the rural sector toward the urban sector. Part of these policies included an aggressive pro-industrial policy of protection and import substitution (O'Donnell 1978, p. 147). Perón was reelected in 1951, albeit in an election tainted by corruption and the repression of the opposition, and he was subsequently removed from power by a coup in 1955. Between 1958 and 1966, civilian governments highly restricted by the military returned, only to be swept away by another coup in 1966 (see O'Donnell 1973; for the seminal analysis).

In 1966, General Juan Carlos Onganía became president, but his regime was quickly opposed by substantial social mobilization (Rock 1987, p. 349). Cavarozzi (1986, p. 36) notes the significance of "the popular insurrection of 1969 . . . [which] fused together blue and white collar workers, students and the urban poor." This revolt against the dictatorship was followed by more, particularly in 1971, and coincided with the emergence of several armed groups and guerillas dedicated to the overthrow of the regime.

Democracy was re-created in 1973 when Perón returned from exile and was elected president in the first truly democratic election since his first election in 1946. However, democratization unleashed the same distributional conflicts that it had before and "As in 1946, the kernels of his program were income redistribution in favor of labor, the expansion of employment, and renewed social reform" (Rock 1987, p. 361). In 1976, the Perónist government, led by Perón's third wife Isabel after his death in 1974, fell to a coup under the leadership of General Jorge Videla. "Once in power, the Army embarked on the conquest of any lingering resistance to a revolution in government whose aim was the total dismantlement of the Perónist state" (Rock 1987, p. 366). The regime that lasted until the Falklands (Malvinas) War of 1982–3 was the most repressive in Argentine history. Some ten thousands people "disappeared" and many thousands more were imprisoned without trial, tortured, and forced into exile. General Roberto Viola succeeded Videla in 1981 but was forced from office the same year by General Leopoldo Galtieri.

As the military became more and more beleaguered and popular protests against them rose, they launched the ill-fated invasion of the Falkland (Malvinas) Islands. Galtieri resigned when the Argentine forces surrendered in June 1982 and, the following year, democratic elections led to the election of Radical president Raúl Alfonsin. Argentina was a democracy again and it has stayed one with Alfonsin being followed by Carlos Menem in 1990, Fernando de la Rúa in 2000, and – after a bewildering succession of temporary presidents during the economic crisis of 2001–2 – by Néstor Kirchner in 2003.

The political history of Argentina therefore reveals an extraordinary pattern where democracy was created in 1912, undermined in 1930, re-created in 1946, undermined in 1955, fully re-created in 1973, undermined in 1976, and finally reestablished in 1983. In between were various shades of nondemocratic governments ranging from restricted democracies to full military regimes. The political history of Argentina is one of incessant instability and conflict. Economic development, changes in the class structure, and rapidly widening inequality, which occurred as a result of the export boom from the 1880s, coincided with pressure on the traditional political elite to open the system. But, the nature of Argentine society meant that democracy was not stable. Traditional interests were too threatened by the rise to power of the Radicals and continuously worked to undermine democracy. The economic changes of the 1930s only exacerbated this conflict. The

workers became stronger and more militant as they found a leader in Perón, and the distributional conflicts then became embedded in the pro-Perón, anti-Perón struggle. Dictatorial regimes collapsed because of social protests, and democracies collapsed because the radical, populist, and often unsustainable policies they adopted induced military coups.

3. Singapore

Sir Stamford Raffles acquired the island of Singapore from its local Malay ruler for the British East India Company in 1819 (Turnbull 1989; Huff 1994; Milne and Mauzy 1990; 2002). At that time, the island, comprising 622 square miles and lying just 176 kilometer north of the equator, was sparsely populated with just a few hundred inhabitants. It soon became an important trading port for the East India Company and expanded rapidly as a commercial center and entrepôt. This role continued even after the collapse of the East India Company (Singapore became a Crown Colony in 1867 as part of the Straits Settlement) and expanded with the British colonization of the Malayan peninsular after the 1870s and the development of an export economy in Malaya based on commodities such as tin and rubber.

After the Second World War and a traumatic occupation by the Japanese, a political awakening occurred in Singapore as in many other British colonies as they began to anticipate independence. The first elections for a legislative council were held in 1948 under a very restricted franchise in which a majority of the council was still appointed by the British Governor. The late 1940s and early 1950s were characterized by labor unrest, strikes, and demonstrations. In 1955, they forced the British to introduce a new constitution proposed by the Rendel Commission, in which a majority of seats of the legislative council were to be elected and the leader of the majority party would become chief minister. However, the 1955 elections were followed by more riots and social unrest, constitutional negotiations were reopened, and new elections were planned for 1959 with Singapore granted almost complete internal self-rule. The franchise was universal suffrage, and the People's Action Party (PAP) under Lee Kuan Yew won forty-three of the fifty-one seats in the 1959 election.

From the beginning, the PAP aggressively promoted industrialization. One of its strategies was taming the trade-union movement and creating a pliant labor force to attract multinational companies. In 1959, it began to reduce the power of unions, which was finally achieved in 1967 and 1968 when all unions were brought under government control. This was accomplished by the creation of a government body, the National Trade Union Congress, and strikes were made illegal. At the same time, Lee Kuan Yew and the leaders of the PAP distanced themselves from the more radical elements of the party. As a result, in 1961 the party split with thirteen parliamentary members resigning to form a new party, the Barisan Sosialis (BS). Despite this setback, the PAP bounced back and, even

before independence, began to show its skill at political maneuver:

> The PAP then strengthened its grip on power, harassing the BS and the trade unions. Most dramatically, prior to elections in 1963, the PAP used the police special branch to mount a sweep called Operation Cold Store, obliterating the BS's top level leadership. (Case 2002, p. 86)

As a result, in the 1963 elections the PAP took thirty-seven out of fifty-one seats, with the BS winning thirteen.

In this initial phase, the PAP saw integration with Malaya as part of its strategy of economic development because it would guarantee a large market for Singaporean firms. In 1963, Malaya, Singapore, Sabah, and Sarawak merged to form the Federation of Malaysia. However, in 1965, Singapore was expelled as a result of tensions between Malay and Chinese politicians (e.g., Lee Kuan Yew had campaigned in Kuala Lumpur in the 1964 Malaysian general elections, to the outrage of Malaysian politicians).

After the creation of the republic in 1965, the PAP began to harass its political opponents. As a consequence, all the BS members resigned their parliamentary seats and boycotted the 1968 elections. In these circumstances, the PAP won all fifty-eight seats, although fifty-one were uncontested. The PAP also won every seat in 1972, 1976, and 1980 against an assortment of opposition parties with the BS contesting elections again after 1972. Finally, a 1981 by-election resulted in the first opposition member since 1968. A second opposition member was elected in 1984 and by 1991 there were four. However, the opposition only ran candidates in a minority of seats; the PAP was consequently always guaranteed a majority in the Parliament. In 1997, the PAP won eighty-two of the eighty-three seats. In the 2001 elections, the PAP won eighty-one seats. To avoid a real opposition appearing in this period and to appease desires for some sort of alternative representation, the PAP introduced nonconstituency Members of Parliament who were allocated to those opposition losers who received the most votes. By 2001, there were nine of these members of the legislature. In 1990, Lee Kuan Yew retired as Prime Minister and was replaced by Goh Chok Tong, who was succeeded in 2004 by Lee's son, Lee Hsien Loong.

Throughout this period, the PAP extended its control over society, particularly through its control of the media. Case (2002, p. 89) demonstrates that "political activism in Singapore risks blacklisting, shunning, lawsuits, tax investigations, lost business opportunities, and detention without trial." To maintain its power, the PAP also engages in extensive gerrymandering to avoid losing any seats. Although the initial electoral system was based on British-style single-member districts, there is now a mix of these and multimember districts (called group representation constituencies). Rodan (1997, p. 178) notes that "single constituencies in which opposition parties came within striking distance of defeating PAP candidates in the last election have disappeared, usually subsumed under group representation constituencies comprising sitting PAP candidates."

When it comes to election time, the PAP also engages in blatant threats to the electorate to influence their votes. Rodan (1998, p. 179) notes that in 1997, the electorate

> . . . were given a stark choice: return government candidates and benefit from a range of expensive new public programs, or have this withheld or delayed in retaliation for electing PAP opponents. . . . Threats by Goh concerning the multimillion dollar housing upgrading program caused special concern. Given that around 86% of Singaporeans live in government built flats, the electorate is highly vulnerable to such intimidation. The announcement of a new system of vote counting enabling the government to ascertain voting preferences down to precinct levels of 5,000 voters reinforced the threat.

Given its size and colonial history, Singapore lacks an aristocracy – landed or otherwise – which has been important for Singaporean politics. It has an urbanization rate of 100 percent, and the ethnic composition of its population is approximately 75 percent Chinese, 15 percent Malay, and 8 percent from the Indian subcontinent. Prior to independence, Singapore also lacked large capitalists or business interests and, since independence, the largest capitalists involved in Singapore are foreigners, who are seemingly promoted by the PAP at the expense of indigenous business interests. Founded by English-educated professionals and middle-class people, the PAP recruits its politicians from the professions and the civil service, not through party members. Indeed, the party exists mostly as an electoral machine; otherwise, it works through the government rather than through some independent grassroots organization. Lee Kuan Yew said in 1984, "I make no apologies that the PAP is the Government and the Government is the PAP" (quoted in Milne and Mauzy 1990, p. 85).

Overall, we see that Singapore moved to democracy and independence as its citizens protested against British colonial rule, but the PAP rapidly established one-party rule after 1963. Since then, the economy has boomed, inequality has been low, and the PAP has maintained power through relatively benign means, fostering popularity through extensive social welfare programs as well as engaging in threats and coercion. Although there has been imprisonment and harassment, there have been no "disappearances" and there is apparently little opposition to PAP rule and little pressure for political change.

4. South Africa

The European presence in South Africa began in 1652 when the Dutch East India Company founded a colony in Table Bay. Its aim was to grow food and provisions for its ships sailing around the Cape of Good Hope from Europe to Asia. The Dutch settlements gradually expanded at the expense of the indigenous Khoikhoi but only extended about 100 miles inland by the end of the eighteenth century. The strategic position of the Cape Colony meant that it became an important prize in geopolitical competition. During the Napoleonic Wars, it was seized by the British first in 1795 and then again – this time conclusively – in 1806, and the colony was amalgamated into the British Empire.

The British, like the Dutch East India Company, initially had no designs on the interior and were more concerned with the safety of the shipping routes to India and Asia. However, the colonial policies of the British alienated many of the Dutch settlers, who became known as Boers or Afrikaners. In response, the Boers moved inland en masse, founding the Orange Free State in 1854 and the Transvaal in 1860.

The British government had formalized the political institutions of the Cape Colony in 1853 when it introduced a bicameral parliament that could legislate on domestic matters, although subject to a veto from London. The executive branch of government consisted of officials appointed by the colonial office. The franchise for the legislature did not specifically disenfranchise people based on racial origins but, instead, adopted the British system of property and income restrictions (Thompson 1995, p. 65).

The political balance between the British Empire and the Boer Republics was altered by the discovery of diamonds in Kimberley and gold on the Witwatersrand in the 1870s. The labor relations in these areas quickly exhibited a pattern that would subsequently become known as "apartheid," with blacks being unable to dig for diamonds, forced to carry passes to impede labor mobility, banned from desirable occupations that became reserved for whites, and forced to live in segregated communities and camps. The British annexed the diamond fields in 1871, the Transvaal in 1877, and, in 1879, finally vanquished the powerful Zulu Kingdom. However, the Transvaal successfully rebelled in 1881 and it was only after the South African War of 1899–1902 that the British government conquered all of the Boer Republics. The British moved the colonies toward a union and, in 1910, the Cape Colony, Natal, Orange Free State, and Transvaal were joined to form the Union of South Africa.

That first government, run by Louis Botha and Jan Smuts, gradually began to reinforce many of the vast inequalities in South African society, a process that culminated in the creation of full-blown apartheid with the election of the National Party (NP) under D. F. Malan in 1948. For example, in 1913, the Natives Land Act stopped Africans from purchasing land outside of "native areas," which were reserves set aside for Africans, consisting in 1939 of about 12 percent of the land area (Africans represented 70 percent of the population in this period; see Thompson 1995, Table 1, p. 278).

At the same time, the first organized black political consciousness began to emerge with the founding of the African National Congress (ANC) in 1912. At first, it was a modest movement organized by middle-class Africans but, following the Second World War, the ANC became radicalized because of the failure to liberalize the system. In 1943, the ANC adopted a statement called *Africans' Claims in South Africa*, demanding for the first time universal adult suffrage.

The application of apartheid after 1948 reached its apogee during the prime-ministership of Hendrik Verwoerd between 1958 and 1966. The government attempted to move all Africans into eight (then ten) homelands, and only Africans whose labor was needed in the white economy could be present in "European areas." They had to carry "passes," proving that they were legally outside of the tribal areas.

The apartheid regime was sustained by massive infringements on political and civil rights. The government established tight control over the media and had a monopoly on radio and television. The police were given vast powers to arrest people without trial and hold them indefinitely in solitary confinement. Under the Public Safety Act of 1953, the government could declare a state of emergency and rule by proclamation.

Throughout the 1950s, the ANC continually contested in the streets and in the law courts the policies of the NP. In one such demonstration in Sharpeville in 1960, a riot exploded and police fired into the crowd, killing eighty-three people. After this incident, the government moved to finally eradicate the ANC and, in 1964, Nelson Mandela and other top leaders were imprisoned on Robben Island. Despite losing much of their leadership to South African prisons or exile, the ANC continued to be the focus of opposition to the regime. The NP pressed ahead with its goal of creating independent homelands (or bantustans), where all Africans would be citizens. In 1976, the Transkei and Bophuthatswana were declared independent nations by the government (although they were never recognized by any other national government or international agency).

In 1976, a riot in Soweto, a large African township just outside Johannesburg, ended in 575 deaths (Thompson 1995, pp. 212–13). Soweto marked a turning point. In the 1960s, the apartheid government had managed to crush the ANC leadership, but

> after the Soweto uprising, a protest culture pervaded the black population of South Africa. Students and workers, children and adults, men and women, the educated and the uneducated became involved in efforts to liberate the country from apartheid. (Thompson 1995, p. 228)

The apartheid government had no choice but to make some concessions. It immediately announced the cessation of the creation of homelands; however, as soon as the turmoil subsided, the government reneged and two more homelands were created in the early 1980s. More significant, the government moved to legalize African trade unions and in 1984 introduced a new constitution in which both Indians and Coloureds had their own legislatures. The whites remained in a solid majority in the legislature. After P. W. Botha was elected president, he had only one Indian and one Coloured in his cabinet, neither with a specific portfolio. After 1984, the government also removed job reservations, which stopped Africans from undertaking specific occupations.

Nevertheless, the basic philosophy or structure of apartheid was unaltered. These concessions were, therefore, not sufficient to prevent the strikes, riots, and social unrest that became more widespread. For instance, in 1985, 879 people were killed in political violence, and there were 390 strikes involving 240,000 workers. The African trade unions, whose legitimization had been a concession after Soweto, were in the forefront of antistate activities. In June 1986, the Botha government responded to these events by declaring a state of emergency and sending the army into the townships to restore order.

The situation got worse for the apartheid regime in October 1986 when the United State imposed sanctions. From the mid-1980s onward, sensing the infeasibility of continuing with the same set of institutions, many members of the South African white elite started to make overtures to the ANC and black leaders. The industrial chaos caused by the strikes was severely damaging to profits and, from the late 1970s onward, there was sustained capital outflow from South Africa (Wood 2000, Figure 6.3, p. 154). Prominent white businessmen met with the ANC in London and other places, and Mandela himself was moved from Robben Island and had many discussions with different members of the Botha government.

> As Mandela recognized, if there was to be peaceful transition, a way would have to be found to reconcile the ANC demand for majority rule with "the insistence of whites on structural guarantees that majority rule will not mean domination of the white minority by the blacks." (Thompson 1995, p. 244)

In February 1989, L. W. de Klerk took over from P. W. Botha as the head of the NP and was elected president in September.

> De Klerk . . . understood that domestic and foreign pressures were undermining the racial order. De Klerk concluded that the best hope for his people was to negotiate a settlement from a position of strength, while his government was still the dominant force in the country. (Thompson 1995, p. 244)

At the beginning of 1990, he lifted the ban on the ANC and released Mandela from prison. Intense negotiations started over the nature of the transition from the apartheid era and what sort of society would follow it. Constitutional negotiations began in December 1991 with the NP proposing a series of measures to weaken the threat of black majority rule.

> South Africa was to become a confederation of states with vast and irremovable powers. Its central executive was to be a coalition of every party that won a substantial number of seats in an election, the chairmanship was to rotate among party leaders, and all decisions were to be made by consensus or special majorities. (Thompson 1995, p. 248)

Such stipulations were unacceptable to the ANC and in June 1992 the negotiations broke down. In September, they were restarted and, by February 1993, there was an agreed-upon timetable for transitions to the April 1994 election.

An interim constitution was agreed upon with the first new Parliament elected in 1994, charged with devising a permanent constitution. The interim constitution incorporated thirty-four basic principles and dictated that no subsequent amendment would be valid if it contradicted them; whether it did so was to be determined by a constitutional court appointed by President Mandela. Other amendments required a two-thirds majority of both Houses of Parliament. The main concession to the NP was that there had to be compulsory power-sharing in the cabinet, with any party that won at least twenty seats in the national assembly getting representation in the cabinet in proportion to its seats. The ANC received 62.7 percent of the vote in the 1994 election.

From its roots, like many colonial societies, South Africa was a society of great inequalities, both economic and political. In the twentieth century, this inheritance led to a highly undemocratic polity in which only whites were enfranchised. After the Second World War, Africans began to successfully mobilize against this political status quo, and they were able to exert increasing pressure, rendering the existing apartheid regime infeasible and threatening mass revolt. Attempts by the regime to make concessions, although leaving the system basically unaltered, failed to achieve this objective, and the apartheid regime maintained power through the use of extensive repression and violence. In 1994, the regime was forced to democratize rather than risk potentially far worse alternatives.

5. The Agenda

We see four very different paths of political development in these narratives. Britain exemplifies the path to consolidated democracy, without any significant reversals in the process. Argentina illustrates the possibility of a transition to an unconsolidated democracy, which then reverts back to nondemocracy, with the process potentially repeating itself multiple times. Singapore is an example of a society in which a nondemocratic regime can survive a long time with relatively minor concessions but also without significant repression. South Africa before the collapse of apartheid exemplifies a nondemocratic regime that survives by using repression. We now propose a framework to understand these various paths and develop predictions for when we expect to see one path versus another.

2 Our Argument

Why did Britain, Argentina, Singapore, and South Africa follow different political paths? More generally, why are some countries democratic whereas others are ruled by dictatorships or other nondemocratic regimes? Why do many nondemocracies transition into democracy? What determines when and how this transition takes place? And, relatedly, why do some democracies, once created, become consolidated and endure whereas others, like many of those in Latin America, fall prey to coups and revert back to dictatorship?

These are central questions for political science, political economy, and social science more generally, but there are neither widely shared answers nor an accepted framework to tackle them. The aims of this book are to develop a framework for analyzing these questions, provide some tentative answers, and outline future areas for research. As part of our investigation, we first provide an analysis of the role of various political institutions in shaping policies and social choices, emphasizing how politics differs in democratic and nondemocratic regimes. To do so, we model the attitudes of various individuals and groups toward different policies and, therefore, toward the political institutions leading to these policies.

To facilitate the initial exposition of our ideas, it is useful to conceive of society as consisting of two groups – the elites and the citizens – in which the latter are more numerous. Our framework emphasizes that social choices are inherently conflictual. For example, if the elites are the relatively rich individuals – for short, the rich – they will be opposed to redistributive taxation; whereas the citizens, who will be relatively poor – for short, the poor – will be in favor of taxation that would redistribute resources to them. More generally, policies or social choices that benefit the elites will be different from those that benefit the citizens. This conflict over social choices and policies is a central theme of our approach.

Who is the majority and who is the elite? This depends to some extent on context and the complex way in which political identities form in different societies. In many cases, it is useful to think of the elite as being the relatively rich in society, as was the case in nineteenth-century Britain and Argentina. However, this is not always the case; for instance, in South Africa, the elites were the whites and, in

many African countries, the elites are associated with a particular ethnic group. In other societies, such as Argentina during some periods, the elite is the military.

It may not be a coincidence that in many situations the elite and the rich coincide. In some cases, those who are initially rich may use their resources to attain power, perhaps by bribing the military or other politicians. In other circumstances, power may be attained by people who are not initially rich. Nevertheless, once attained, political power can be used to acquire income and wealth so that those with power naturally tend to become rich. In either case, there is a close association between the elite and the rich.

Our theory of which societies will transit from dictatorship to democracy and under what circumstances democracy will be consolidated is related to the conflict between the elite and the citizens over politics. These groups have opposing preferences over different political institutions, democracy and dictatorship, which they recognize lead to different social choices. However, we also emphasize that political institutions do not simply determine the extent of redistribution or who benefits from policies today, they also play the role of regulating the future allocation of political power. In democracy, the citizens have more power both today and in the future than they would in nondemocratic regimes because they participate in the political process.

The framework we develop is formal, so our exposition emphasizes both the concepts that we believe are essential in thinking about democracy as well as how those concepts and issues can be formally modeled using game theory.

1. Democracy versus Nondemocracy

At the outset, we have to be clear about the precise questions that we tackle and the basic building blocks of our approach. In building models of social phenomena, an often-useful principle is the so-called Occam's razor. The principal, popularized by the fourteenth-century English philosopher William of Occam, is that one should not increase the number of entities required to explain a given phenomenon beyond what is necessary. In other words, one should strive for a high degree of parsimony in formulating answers to complex questions. Given the complexity of the issues with which we are dealing, we frequently make use of this principle in this book not only to simplify the answers to complex questions but, perhaps even more daringly, to also simplify the questions. In fact, in an attempt to focus our basic questions, we use Occam's razor rather brutally and heroically. We abstract from many interesting details and also leave some equally important questions out of our investigation. Our hope is that this gambit pays off by providing us with relatively sharp answers to some interesting questions. Of course, the reader is the judge of whether our strategy ultimately pays off.

Our first choice is about the classification of different regimes. Many societies are today governed by democratic regimes, but no two democracies are exactly alike and most exhibit a number of marked institutional differences. Consider, for

instance, the contrast between the French presidential system and the British parliamentary system, or that between the majoritarian electoral institutions as used in the United States and the system of proportional representation used in much of continental Europe. Despite these differences, there are some important commonalities. In a democracy, the majority of the population is allowed to vote and express their preferences about policies, and the government is supposed to represent the preferences of the whole population – or, using a common description, "democracy is the government by the people for the people." In contrast, many other countries are still ruled by dictators and nondemocratic regimes.[1] There are even more stark differences between some of these nondemocratic regimes than the differences between democracies. For example, reflect on the contrast between the rule of the Chinese Communist Party since 1948 and that of General Pinochet in Chile between 1973 and 1989. When we turn to other nondemocratic regimes, such as the limited constitutional regimes in Europe in the nineteenth century, the differences are even more marked.

Nevertheless, these nondemocratic regimes share one common element: instead of representing the wishes of the population at large, they represent the preferences of a subgroup of the population: the "elite." In China, it is mainly the wishes of the Communist Party that matter. In Chile, most decisions were made by a military junta; it was their preferences, and perhaps the preferences of certain affluent segments of the society supporting the dictatorship, that counted. In Britain before the First Reform Act of 1832, less than 10 percent of the adult population – the very rich and aristocratic segments – was allowed to vote, and policies naturally catered to their demands.

From this, it is clear that democracies generally approximate a situation of *political equality* relative to nondemocracies that, in turn, represent the preferences of a much smaller subset of society and thus correspond more to a situation of *political inequality*. Our focus is to understand the social and economic forces pushing some societies toward regimes with greater political equality versus those encouraging the development of more nondemocratic systems. In our models, except in Chapter 8, we work with a dichotomous distinction between democracy and nondemocracy. Nevertheless, in deciding how democratic actual regimes are and in empirical work, it is more useful to think of various shades of democracy. For example, none of the nineteenth-century reform acts in Britain introduced universal adult suffrage, but they were all movements in the direction of increased democracy. We want to understand these movements; to do so, we begin by simply considering a move from nondemocracy to full democracy (universal adult suffrage). Our definition is "Schumpetarian" (Schumpeter 1942) in the sense that we emphasize that a country is democratic if a certain political process takes place – if certain key institutions, such as free and fair elections and free entry into

[1] In the text, despite the title of our book, we prefer to use the term *nondemocracy* to alternatives, such as *dictatorship* or *authoritarian regime*, because it has fewer specific connotations than any of the other terms.

politics, are in place. To the extent that democracy is associated with particular outcomes, it will be because they stem from its institutional features.

Our approach means that we are not simply interested in when universal adult suffrage was introduced but rather in understanding all movements in the direction of increased democracy. For example, in Argentina, universal male suffrage was introduced by the constitution of 1853, but electoral corruption was so endemic that democracy was not a reality until after the political reforms under President Sáenz Peña in 1912. In this case, we consider 1912 to be a key movement toward democracy. In the case of Britain, the reforms of 1867 greatly extended voting rights, but universal male suffrage was not conceded until 1919. However, electoral corruption was eliminated and secret voting was introduced in 1872. In this case, we see 1867 as representing an important step toward political equality in Britain.

We have less to say on the extension of suffrage to women. In almost all European countries, voting rights were first given to adult men and subsequently extended to women. This reflected the then-accepted gender roles; when the roles began to change as women entered the workforce, women also obtained voting rights. It is likely, therefore, that the mechanisms that we propose better describe the creation of male suffrage than the extension of voting rights to women.

Our dichotomous distinction between democracy and nondemocracy makes sense and is useful only to the extent that there are some important elements central to our theory and common to all democracies but generally not shared by nondemocracies. This is indeed the case. We argue that democracy, which is generally a situation of political equality, looks after the interests of the majority more than nondemocracy, which is generally dominated by an elite and is more likely to look after its interests. Stated simply and extremely, nondemocracy is generally a regime for the elite and the privileged; comparatively, democracy is a regime more beneficial to the majority of the populace, resulting in policies relatively more favorable to the majority.

We claim that nondemocracy represents political inequality relative to democracy. In democracy, everybody has a vote and, at least potentially, can participate in one way or another in the political process. In nondemocracy, an elite, a junta, an oligarchy, or – in the extreme case – just one person, the dictator, is making the decisions. Hence, the contrast in terms of political equality makes sense. This, of course, does not mean that democracy corresponds to some ideal of political equality. In many successful democracies, there is one-person-one-vote, but this is far from perfect political equality. The voices of some citizens are louder, and those with economic resources might influence policies through nonvoting channels, such as lobbying, bribery, or other types of persuasion. Throughout the book, when we discuss political equality in democracy, it is always a *relative* statement.

Overall, the outlines of our basic approach are taking shape. We think of regimes falling into one of two broad categories: democracy and nondemocracy.

Democracy is thought of as a situation of political equality and characterized by its relatively more pro-majority policies. Often pro-majority policies coincide with pro-poor policies, especially a greater tendency to redistribute income away from the rich toward the poor. In contrast, nondemocracy gives a greater say to an elite and generally opts for policies that are less majoritarian than in a democracy.

2. Building Blocks of Our Approach

We have now determined the basic focus of our investigation: to understand why some societies are democratic, why some societies switch from nondemocracy to democracy, and why some democracies revert back to dictatorships. We have already mentioned some of the building blocks of our approach; it is now time to develop them more systematically.

The first overarching building block for our approach is that it is economic.[2] By this term, we do not mean that individuals always act rationally according to some simple postulates. Nor do we mean that there are only individuals, and no social groups, in society. Instead, we mean that individuals have well-defined preferences over outcomes or the consequences of their actions; for example, they prefer more income to less and they may prefer peace, security, fairness, and many other things. Sometimes masses of individuals have interests in common or even act collectively. However what matters is that individuals *do* have well-defined preferences that they understand. They evaluate various different options, including democracy versus nondemocracy, according to their assessments of their (economic and social) consequences. In such situations, the economic approach suggests that people often behave strategically and that their behavior should be modeled as a game. Game theory is the study of situations with multiple decision makers, interacting strategically. The basic tenet of game theory is that individuals choose between various strategies according to their consequences. Our economic focus and the presence of important interactions between various political actors render all the situations analyzed herein essentially "game theoretic." We, therefore, make heavy use of game theory in modeling preferences over different regimes and transitions between these regimes.

To see the implications of these assumptions, consider a group of individuals for whom democracy and nondemocracy have the same consequences in all spheres, except that democracy generates more income for them; they naturally prefer more income to less. Therefore, we expect these individuals to prefer democracy to nondemocracy. At some level, this postulate is very weak; but, at another level, we are buying a lot with our economic focus. Most important, we are getting a license to focus on the consequences of the regimes, and preferences over regimes are derived from their consequences. Such an approach is consistent with many

[2] In political science, such an approach is often called "rational choice."

historical accounts of the motivations of different actors. For example, in 1839, the Chartist J. R. Stephens argued:

> The question of universal suffrage . . . is a knife and fork question, a bread and cheese question . . . by universal suffrage I mean to say that every working man in the land has a right to a good coat on his back, a good hat on his head, a good roof for the shelter of his household, a good dinner upon his table. (quoted in Briggs 1959, p. 34)

The alternative would have been simply to assume that one group dislikes democracy whereas another group likes democracy – for example, because of certain ideological preferences or biases (Diamond 1999). Indeed, Diamond (1992, p. 455) argues that

> democracy becomes truly stable only when people come to value it widely not solely for its economic and social performance but intrinsically for its political attributes.

We are not denying that such ideological preferences exist, but we believe that individuals' and groups' preferences over regimes derived from the economic and social consequences of these regimes are more important. Later in the book, we discuss how introducing ideological preferences affects our results, and the general message is that – as long as these do not become the overriding factors – they do not affect our conclusions.

Our second building block is that politics is inherently conflictual. Most policy choices create distributional conflict; one policy benefits one group whereas another benefits different individuals. This is a situation of *political conflict* – conflict over the policies that society should adopt. These groups – for example, the rich and the poor – have conflicting preferences over policies, and every policy choice creates *winners* and *losers*. For instance, with high taxes, the rich are the losers and the poor are the winners, whereas when low taxes are adopted, the roles are reversed. In the absence of such conflict, aggregating the preferences of individuals to arrive at social preferences would be easy; we would simply have to choose the policy that makes everybody better off. Much of political philosophy exists because we do not live in such a simple world, and situations of conflict are ubiquitous. Every time society (or the government) makes a decision or adopts a policy, it is implicitly siding with one group, implicitly resolving the underlying political conflict in one way or another, and implicitly or explicitly creating winners and losers.

Although the economic approach emphasizes individual preferences and motivations, many individuals often have the same interests and sometimes make the same decisions. Moreover, groups of individuals may be able to act collectively if there are no collective-action problems or if they can solve any that exist. If this is the case, then we can usefully discuss conflict and who is in conflict with whom in terms of groups of individuals. These groups may be social classes, somewhat similar to Marxist accounts of history and politics, or they may be urban agents, ethnic or religious groups, or the military. Our focus on social groups as key political actors is motivated by our sense that the most important forces in political conflict and change are groups of individuals.

Leaving aside issues of political philosophy related to how a just or fair society should reconcile these conflicting preferences, how does society resolve political conflict in practice? Let us make this question somewhat more concrete: suppose there are two policies, one favoring the citizens and the other favoring the elites. Which one will the society adopt? Because there is no way of making both groups happy simultaneously, the policy choice has to favor one group or the other. We can think that which group is favored is determined by which group has *political power*. In other words, political power is the capacity of a group to obtain its favorite policies against the resistance of other groups. Because there are always conflicting interests, we are always in the realm of political conflict. And, because we are always in the realm of political conflict, we are always under the shadow of political power. The more political power a group has, the more it will benefit from government policies and actions.

What is political power? Where does it come from? In thinking of the answers to these questions, it is useful to distinguish between two different types of political power: *de jure political power* and *de facto political power*. Imagine Thomas Hobbes's (1996) state of nature, where there is no law and man is indistinguishable from beast. Hobbes considered such a situation to argue that this type of anarchy was highly undesirable, and the state, as a leviathan, was necessary to monopolize force and enforce rules among citizens. But, how are allocations determined in Hobbes's state of nature? If there is a fruit that can be consumed by one of two individuals, which one will get to eat it? The answer is clear: because there is no law, whoever is more powerful, whoever has more brute force, will get to eat the fruit. The same type of brute force matters in the political arena as well. A particular group will have considerable political power when it has armies and guns to kill other groups when policies do not go its way. Therefore, the first source of political power is simply what a group can do to other groups and the society at large by using force. We refer to this as de facto political power. Yet, and fortunately so, this is not the only type of political power. Today, key decisions in the United Kingdom are made by the Labour Party, not because it can use brute force or because it has acquired de facto power through some other means but rather because political power has been allocated to it by the political system (i.e., it was voted into office in the last general election). As a result, among policies with conflicting consequences, the Labour Party can choose those that are more beneficial to its constituency or to its leaders. We call this type of political power, allocated by political institutions, de jure political power. Actual political power is a combination of de jure and de facto political power, and which component matters more depends on various factors – a topic that we discuss later.

Finally, we refer to the social and political arrangements that allocate de jure political power as *political institutions*. For example, an electoral rule that gives

the right to decide fiscal policies to the party that obtains 51 percent of the vote is a particular political institution. For our purposes, the most important political institutions are those that determine which individuals take part in the political decision-making process (i.e., democracy versus nondemocracy). Therefore, a major role of democracy is its ability to allocate de jure political power. In democracy, the majority has relatively more de jure political power than it does in nondemocracy. That democracies look after the interests of the majority of citizens more than nondemocracies is simply a consequence, then, of the greater de jure political power of the majority in democracy than in nondemocracy.

3. Toward Our Basic Story

Armed with the first two basic building blocks of our approach, we can now start discussing preferences over different regimes. Typically, there is political conflict between the elites and the citizens, and democracies look after the interests of the citizens more than nondemocracies. It is, therefore, natural to think that the citizens have a stronger preference for democracy than the elites. So, if there is going to be conflict about which types of political institutions a society should have, the majority of citizens will be on the side of democracy and the elites will be on the side of nondemocracy. This is a good starting point.

We could add more empirical content to this structure by assuming that the elites were the relatively rich and the majority the relatively poor. Indeed, in many instances, the transition from nondemocracy to democracy was accompanied by significant conflict between poorer elements of society, who were hitherto excluded and wanted to be included in the political decision-making process, and the rich elite, who wanted to exclude them. This was most clearly the case in nineteenth-century Europe, particularly Britain, as we saw in Chapter 1, when initially the middle classes and subsequently the working classes demanded voting rights. Their demands were first opposed by the rich elite, who then had to concede and include them in the political system.

In line with this account of political developments in nineteenth-century Europe, Aminzade (1993, p. 35) describes the arrival of universal male suffrage to French politics as follows:

> French workers, mainly artisans, constituted the revolutionary force that put the Republican party in power in February 1848 . . . and working class pressure from the streets of Paris forced liberal Republican leaders . . . to reluctantly concede universal male suffrage.

Perhaps, more tellingly, the key players in the process of democratization saw it as a fight between the rich and the poor. Viscount Cranborne, a leading nineteenth-century British Conservative, described the reform struggle as

> . . . a battle not of parties, but of classes and a portion of the great political struggle of our century – the struggle between property . . . and mere numbers. (quoted in Smith, 1966, pp. 27–8)

The conflict between the poorer and richer factions of society was also a defining characteristic of most instances of the introduction of universal suffrage in Latin America in the first half of the twentieth century – including the experiences in Argentina in 1912, as we saw in Chapter 1, but also in Uruguay in 1919, in Colombia in 1936, and in Venezuela in 1945. The arrival of democracy in South Africa and Zimbabwe similarly followed a conflict between the rich whites and poor blacks.

This discussion, therefore, highlights how the majority of citizens want democratic institutions because they benefit from them and, therefore, will strive to obtain them. Given our definition of political power, we can say that the citizens are more likely to secure a transition to democracy when they have more de facto political power. Thus, we have already constructed a simple theory of democratization: the citizens want democracy and the elites want nondemocracy, and the balance of political power between the two groups determines whether the society transits from nondemocracy to democracy (and perhaps also whether democracy, once created, becomes consolidated or reverts back to nondemocracy later).

This could be viewed as a simplified version of our theory of democratization. But, in fact, it is so simplified that some of the essential features of our theory are absent. Most important, the role that democracy or, more generally, political institutions play is trivialized.

The theory says that democracy leads to social choices more favored by the majority of citizens; hence, the citizens prefer democracy to nondemocracy, and democracy results when the citizens have sufficient political power. However, if the citizens have sufficient political power, why don't they use this power to simply obtain the social choices and policies that they prefer rather than first fight for democracy and then wait for it to deliver those policies to them? Is democracy simply a not-so-necessary intermediate step here? One could argue so.

This is only a feature of the simple story we have told so far, and it is a characteristic of neither real-world political institutions nor of our theory. In practice, political institutions play a much more fundamental role than being a simple intermediating variable: they regulate the future allocation of political power between various social groups. They play this role because we do not live in a static world like the one described in the previous narrative but rather in a dynamic world, where individuals care not only about policies today but also about policies tomorrow. We can capture this important role of political institutions and obtain a more satisfactory understanding of democracy and democratization by incorporating these dynamic strategic elements, which is what our theory of democratization attempts to do.

4. Our Theory of Democratization

Consider the simplest dynamic world we can imagine: there is a "today" and a "tomorrow," and the elites and the citizens care about policies both today and tomorrow. There is nothing that prevents society from adopting a different policy tomorrow from the one it chose today. Thus, it is not sufficient for the citizens

to ensure policies they prefer today; they would also like similar policies to be adopted tomorrow. Suppose we are in a nondemocratic society, which generally looks after the interests of the elites. Citizens have de facto political power today, so they can obtain the policies they like, but they are unsure whether they will have the same political power tomorrow. Given that we are in a nondemocratic society, tomorrow the elites may become more powerful and assertive and the citizens may no longer have the same political power. Can they ensure the implementation of the policies they like both today and tomorrow?

This is where political institutions may be important relative to the static world described previously. Institutions, by their nature, are *durable* – that is, the institutions of today are likely to persist until tomorrow. A democratic society is not only one where there is one-person-one-vote today but also one that is expected to remain democratic at least in the near future. This durability was already implicit in our definition of political institutions as a means of allocating political power: they regulate the *future* allocation of political power. For example, democracy means that tomorrow there will be a vote to determine policies or to decide which party will rule and the whole population will participate. Nondemocracy means that much of the population will be excluded from collective decision-making processes.

Imagine now that the citizens do not simply use their de facto political power today to obtain the policies they like now, but they also use their political power to change the political system from nondemocracy to democracy. If they do so, they will have effectively increased their de jure political power in the future. Instead of nondemocracy, we are now in a democratic regime where there will be voting by all. With their increased political power, the citizens are therefore more likely to secure the policies they like tomorrow as well.

We have now moved toward a richer theory of democratization: transition to democracy – or, more generally, a change in political institutions – emerges as a way of regulating the future allocation of political power. The citizens demand and perhaps obtain democracy so that they can have more political say and political power tomorrow. Returning to the beliefs of the Chartist J. R. Stephens (quoted in Briggs 1959), we can now see that he was correct in demanding universal suffrage as a means of securing the "right to a good coat . . . a good hat . . . a good roof . . . [and] a good dinner" for working men rather than directly demanding the coat, the roof, and the dinner. Those would have been only for today, whereas universal suffrage could secure them in the future as well.

Notice an important implicit element in the story: *the transitory nature of de facto political power*. The citizens are presumed to have political power today but uncertain about whether they will have similar power tomorrow. The balance between the elites and the citizens or, more generally, between various social groups is not permanent, is not set in stone, is not the same today as it will be tomorrow; it is transitory. This is reasonable in the dynamic and uncertain world in which we live. It will be even more compelling when we think of the sources of

political power for the disenfranchised citizens in nondemocracy. First, let us try to understand why the transitory nature of political power matters. Suppose that the citizens have the same political power tomorrow as they have today. Why should they need political institutions to help them? If their political power is sufficient to obtain the policies they like (even to obtain the institutions they like) today, then it will be so in the future as well, and there will be no need to change the underlying political institutions. It is precisely the transitory nature of political power – that the citizens have it today and may not have it tomorrow – that creates a demand for change in political institutions. The citizens would like to lock in the political power they have today by changing political institutions – specifically, by introducing democracy and greater representation for themselves – because without the institutional changes, their power today is unlikely to persist.

So why do the citizens have political power in nondemocracy? The answer is that they have de facto rather than de jure political power. In nondemocracy, the elites monopolize de jure political power but not necessarily de facto political power. The citizens are excluded from the political system in nondemocracy, but they are nonetheless the majority and they can sometimes challenge the system, create significant social unrest and turbulence, or even pose a serious revolutionary threat. What is there to stop the majority of the population overwhelming the elite, which constitutes a minority, and taking control of society and its wealth, even if the elites have access to better guns and hired soldiers? After all, the citizens successfully occupied Paris during the Paris Commune, overthrew the existing regime in the 1917 Russian Revolution, destroyed the dictatorship of Somoza in Nicaragua in 1979, and in many other instances created significant turbulence and real attempts at revolution. However, a real threat from the citizens requires the juxtaposition of many unlikely factors: the masses need to solve the collective-action problem necessary to organize themselves,[3] they need to find the momentum to turn their organization into an effective force against the regime, and the elites – who are controlling the state apparatus – should be unable to use the military to effectively suppress the uprising. It is, therefore, reasonable that such a challenge against the system would only be transitory: in nondemocracy, if the citizens have political power today, they most likely will not have it tomorrow.

Imagine now that there is an effective revolutionary threat from the citizens against nondemocracy. They have the political power today to get what they want and even to overthrow the system. They can use their political power to obtain "the coat, the roof, and the dinner," but why not use it to obtain more, the same things not only for today but also in the future? This is what they will get if they can force a change in political institutions. Society will make a transition to democracy and, from then on, policies will be determined by one-person-one-vote, and the

[3] That is, individuals should be convinced to take part in revolutionary activity despite the individual costs and the collective benefits to them as a group.

citizens will have more political power, enabling them to obtain the policies they desire and the resulting coat, roof, and dinner.

In practice, however, changes in political institutions do not simply happen because the citizens demand them. Transitions to democracy typically take place when the elite controlling the existing regime extend voting rights. Why would they do so? After all, the transfer of political power to the majority typically leads to social choices that the elite doesn't like – for instance, higher taxes and greater redistribution away from it in the future, precisely the outcomes it would like to prevent. Faced with the threat of a revolution, wouldn't the elite like to try other types of concessions, even giving the citizens the policies they want, rather than give away its power? To answer this question, let us return to the period of effective revolutionary threat. Imagine that the citizens can overthrow the system and are willing to do so if they do not get some concessions, some policies that favor them and increase their incomes and welfare.

The first option for the elite is to give them what they want today: redistribute income and more generally adopt policies favorable to the majority. But, suppose that concessions today are not sufficient to dissuade the citizens from revolution. What can the elite do to prevent an imminent and, for itself, extremely costly revolution? Well, it can promise the same policies tomorrow. Not only a coat, a roof, and a dinner today but also tomorrow. Yet, these promises may not be *credible.* Changing policy in the direction preferred by the citizens is not in the immediate interest of the elite. Today, it is doing so to prevent a revolution. Tomorrow, the threat of revolution may be gone, so why should it do so again? Why should it keep its promises? No reason and, in fact, it is unlikely to do so. Hence, its promises are not necessarily credible. Noncredible promises are worth little and, unconvinced by these promises, the citizens would carry out a revolution. If it wants to save its skin, the elite has to make a credible promise to set policies that the majority prefer; in particular, it must make a credible commitment to future pro-majority policies. A credible promise means that the policy decision should not be the elite's but rather placed in the hands of groups that actually prefer such policies. Or, in other words, it has to transfer political power to the citizens. A credible promise, therefore, means that it has to change the future allocation of political power. That is precisely what a transition to democracy does: it shifts future political power away from the elite to the citizens, thereby creating *a credible commitment to future pro-majority policies.* The role that political institutions play in allocating power and leading to relatively credible commitments is the third key building block of our approach.

Why, if a revolution is attractive to the citizens, does the creation of democracy stop it? This is plausibly because revolution is costly. In revolutions, much of the wealth of a society may be destroyed, which is costly for the citizens as well as the elite. It is these costs that allow concessions or democratization by the elite to avoid revolution. In reality, it will not always be the case that democracy is sufficiently

pro-majority that it avoids revolution. For example, the citizens may anticipate that, even with universal suffrage, the elite will be able to manipulate or corrupt political parties or maybe it will be able to use its control of the economy to limit the types of policies that democracy can implement. In such circumstances, anticipating that democracy will deliver few tangible rewards, the citizens may revolt. However, to limit the scope of our analysis, we normally restrict our attention to situations where the creation of democracy avoids revolution. Historically, this seems to have been typical, and it means that we do not delve deeply into theories of revolution or into the modeling of post-revolutionary societies.

We now have our basic theory of democratization in place. In nondemocracy, the elites have de jure political power and, if they are unconstrained, they will generally choose the policies that they most prefer; for example, they may choose low taxes and no redistribution to the poor. However, nondemocracy is sometimes challenged by the citizens who may pose a revolutionary threat – when they temporarily have de facto political power. Crucially, such political power is transitory; they have it today and are unlikely to have it tomorrow. They can use this power to undertake a revolution and change the system to their benefit, creating massive losses to the elites but also significant collateral damage and social losses. The elites would like to prevent this outcome, and they can do so by making a credible commitment to future pro-majority policies. However, promises of such policies within the existing political system are often noncredible. To make them credible, they need to transfer formal political power to the majority, which is what democratization achieves.

This story of democratization as a commitment to future pro-majority policies by the elites in the face of a revolutionary threat and, perhaps more important, as a commitment made credible by changing the future distribution of political power is consistent with much historical evidence. As illustrated by the British, Argentinian, and South African political histories discussed in Chapter 1, most transitions to democracy, both in nineteenth- and twentieth-century Europe and twentieth-century Latin America, took place amid significant social turmoil and revolutionary threats. In addition, the creation of democratic societies in most former European colonies in the 1950s and 1960s was the result of pressure by the disenfranchised and relatively poor colonials against the colonizing power. Such threats of turmoil and social disorder similarly accompanied the recent spate of democratizations in Africa (Bratton and van der Walle 1997) and Eastern Europe (Bunce 2003). To quote a classic European example, in presenting his electoral reform to the British Parliament in 1831, Prime Minister Earl Grey was well aware that this was a measure necessary to prevent a likely revolution. He argued:

> There is no-one more decided against annual parliaments, universal suffrage and the ballot, than I am. My object is not to favour, but to put an end to such hopes and projects . . . The principle of my reform is, to prevent the necessity of revolution . . . reforming to preserve and not to overthrow. (quoted in Evans 1996, p. 223).

Lang's (1999, pp. 38–9) conclusion mirrors Grey's:

> The Whigs were aware of the support among working people for the bill. . . . However, they were also quite determined not to allow the working classes to hold any sort of dominant position in the new electoral system. Passing the bill therefore saved the country from risings and rebellion; the content of the bill saved the country from the "evils" of democracy. Needless to say, disappointment among the working classes was likely to be intense once they realized how little they had gained from the bill, but by then they would have lost their middle class allies, won over to the system by the bill, and would be powerless to do anything about it.

The same considerations were also determining factors for the later reforms. For example,

> as with the First Reform Act, the *threat* of violence has been seen as a significant factor in forcing the pace [of the 1867 Reform Act]; history . . . was repeating itself. (Lee 1994, p. 142).

Similarly, the threat of revolution was the driving force behind democratization in the French, German, and Swedish cases. For example, Tilton (1974, pp. 567–8) describes the process leading to the introduction of universal male suffrage in Sweden as follows:

> neither [of the first two reform acts] passed without strong popular pressure; in 1866 crowds thronged around the chamber while the final vote was taken, and the 1909 reform was stimulated by a broad suffrage movement [and] a demonstration strike . . . Swedish democracy had triumphed without a revolution – but not without the *threat* of a revolution. (italics in original)

The threat of revolution and social unrest played an equally important role in the establishment of voting rights for the populace in Latin America. We saw in Chapter 1 how in Argentina, universal male suffrage was effectively institutionalized in 1912 by President Roque Sáenz Peña when the secret ballot was introduced and fraudulent electoral practices outlawed. The movement toward a full democracy was driven by the social unrest created by the Radical Party and the rapid radicalization of urban workers. In Colombia, the creation of universal suffrage during the administration of Liberal President Alfonso López Pumarejo in 1936 was similarly inspired; leading Colombianist historian David Bushnell (1993, p. 185) describes it as follows:

> López . . . was a wealthy man . . . yet he was well aware that Colombia could not go on indefinitely ignoring the needs and problems of what he once described as "that vast and miserable class that does not read, that does not write, that does not dress, that does not wear shoes, that barely eats, that remains . . . on the margin of [national life]." In his opinion such neglect was not only wrong but also dangerous, because the masses would sooner or later demand a larger share of the amenities of life.

Similarly, the reinstatement of democracy in Venezuela in 1958 was a response to intense uprisings and unrest. In describing the situation, Kolb (1974, p. 175) wrote:

> . . . in dramatic intensity and popular violence, the events on January 21 and 22 in Caracas . . . was a true popular revolution of Venezuelan citizens . . . armed with rocks, clubs, home-made grenades, and Molotov Cocktails, against a ferocious and well-trained Police force.

The evidence is, therefore, consistent with the notion that most moves toward democracy happen in the face of significant social conflict and possible threat of revolution. Democracy is usually not given by the elite because its values have changed. It is demanded by the disenfranchised as a way to obtain political power and thus secure a larger share of the economic benefits of the system.

Why does the creation of democracy act as a commitment when we know that democracy often collapses once created? This is because although coups sometimes occur, it is costly to overthrow democracy, and institutions, once created, have a tendency to persist. This is mostly because people make specific investments in them. For instance, once democracy has been created, political parties form and many organizations, such as trade unions, arise to take advantage of the new political circumstances. The investments of all these organizations will be lost if democracy is overthrown, giving citizens an incentive to struggle to maintain democracy. Moreover, once democracy has been created, the majority may have greater control over the military than they had under a nondemocratic regime, which changes the underlying balance of de facto power.

Finally, the trade-off for the elite, facing the threat of revolution by the citizens, is not simply between policy concessions and democratization. A further alternative would be to use force and repression. For example, the white South African regime rejected calls for democracy and kept itself in power for decades by using the military to repress demonstrations and opposition. Similarly, Argentine military regimes of the 1960s and 1970s killed thousands of people to avoid reintroducing democracy; this has been the pattern in many other Latin American countries including Guatemala and El Salvador. In Asia, nondemocratic regimes in China and Burma have used force to block demands for democracy. This was also true in Eastern European countries during the dominance of the Soviet Union – for example, in Hungary in 1956 and Czechoslovakia in 1968. It is clear why repression is attractive for elites because it allows them to maintain power without having to make any concessions to the disenfranchised. Nevertheless, repression is both costly and risky for elites. It leads to loss of life and destruction of assets and wealth, and – depending on the international climate of opinion – it may lead to sanctions and international isolation, as happened in South Africa during the 1980s. Moreover, repression may fail, which could cause a revolution – the worst possible outcome for the elites. These considerations imply that only in certain circumstances will repression be attractive. When we incorporate this into the

analysis, we see that democracy arises when concessions are not credible and repression is not attractive because it is too costly.

5. Democratic Consolidation

A theory of democratization is not sufficient to understand why some countries are democratic whereas some others are ruled by dictatorships. Many countries become democratic but eventually revert back to a nondemocratic regime as a result of a military coup. This has been an especially common pattern in Latin America. As we saw in Chapter 1, Argentina is a striking example of the instabilities of Latin American democracy. Similarly, the path to democracy has been marred by switches to dictatorships in Brazil, Chile, Guatemala, Peru, Venezuela, and Uruguay. Why has democracy been so hard to consolidate in much of Latin America?

To answer this question, we need to develop a theory of coups or, alternatively, a theory of democratic consolidation. What is a consolidated democracy? A democracy is consolidated if the set of institutions that characterize it endure through time. Our theory of democratic consolidation and coups builds on the different attitudes of the elites and the citizens toward democracy. Once again, the citizens are more pro-democratic than the elites (because democracy is more pro-citizen than nondemocracy). Consequently, when there is a situation with the military on the side of the elite and sufficient turbulence to allow a military takeover, the elites might support or sponsor a coup to change the balance of power in society.

The reason that the elites might want to change political institutions, from democracy to nondemocracy, is similar to the reason that the citizens want democratization. What the elites care about is changing policies in their favor, and political turbulence and the alignment between their interests and those of the military might give them the opportunity to do so. However, there is the issue of the transitory nature of de facto political power. They will have this opportunity today but not necessarily tomorrow. Any promise by the citizens to limit the extent to which policy is pro-majority in the future is not credible within the context of democratic politics. Tomorrow, the threat of a coup may be gone and democratic politics will again cater to the needs of the majority, therefore choosing the policies it prefers without worrying about the elite undermining its power via a coup. However, this is precisely what made democracy so costly for the elite in the first place. To change future policies in a credible way, the elites need political power. A coup is their way of increasing their de jure political power so they can pursue the policies they like. In other words, a coup enables the elites to turn their transitory de facto political power into more enduring de jure political power by changing political institutions.

A related reason that a coup may arise is that, in the midst of political and social turbulence, the military and the elite segments of society may be, perhaps rightly, worried about the future sustainability of democracy and even of the capitalist system and want to preempt a potential move farther toward the left or even a revolution.

6. Determinants of Democracy

Now that we have a theory of democratization, we can ask which factors make the emergence and consolidation of democracy more likely. We have so far explained how our theory can account for transitions from nondemocracy to democracy and possibly back again to nondemocracy. However, just as important are *the comparative statics* of the equilibrium, meaning how the equilibrium changes when some underlying factors change. These comparative statics enable us to explain why some countries transition to democracy whereas others do not, and why some countries remain democracies whereas democracy collapses in other countries. These comparative statics can then guide empirical and historical work in understanding the incidence of democracy.

6.1 Civil Society

6.1.1 Democratization

Our framework implies that a relatively effective threat of revolution from the citizens is important for democratization. When the citizens are not well organized, the system will not be challenged and transition to democracy will be delayed indefinitely. Similarly, when civil society is relatively developed and the majority is organized, repression may be more difficult. Therefore, some degree of development in civil society is also necessary for democratization. We take such development as given in this book and it plausibly represents the outcomes of long-run historical processes (e.g., Putnam 1993).

6.1.2 Consolidation

The strength and nature of civil society is as important for the consolidation of democracy as it is for its creation in the first place. Not only is a well-organized civil society necessary to push for democracy, it is also necessary to protect it. When civil society is better organized, coups are easier to resist, more costly to undertake, and less likely to succeed. Hence, democracy is more likely to be consolidated.

6.2 Shocks and Crises

6.2.1 Democratization

In our theory, democratizations occur because of the transitory nature of de facto political power. In some situations, the collective-action problem is easier to solve, opponents to the regime are easier to coordinate, and revolutions are easier and less costly to carry out. These are typically times of crises – for example,

harvest failures, economic depressions, international financial or debt crises, and even wars. Such crises and macroeconomic shocks are intrinsically transitory and lead to short-term fluctuations in de facto political power. Our theory, therefore, predicts that democratizations are more likely to arise in a situation of economic or political crisis. A clear example is the democratization in Argentina after the Falklands (Malvinas) War in 1983.

6.2.2 Consolidation

Just as opponents of dictatorship can gain temporary de facto power when there are political or economic crises, so can opponents of democracy. Our analysis suggests that, as with democratizations, coups are more likely to arise in situations of crisis. An illustrative example is the coup against Allende in Chile in 1973, which came during the first big rise in oil prices and a large economic depression.

6.3 Sources of Income and Composition of Wealth

6.3.1 Democratization

Another important determinant of the trade-off between democracy and repression is the source of income for the elites. In some societies, the elites are heavily invested in land, whereas in others, the elites are those with investments in physical and human capital. There are likely to be three major differences in the attitudes of landowners and (physical and human) capital owners toward democracy and nondemocracy. First, land is easier to tax than physical and human capital. Therefore, landowners have more to fear from democracy than nondemocracy, which makes them more averse to democracy. Second, social and political turbulence may be more damaging to physical and human capital owners who have to rely on cooperation in the workplace and in the trading process, which makes landowners more willing to use force to preserve the regime they prefer. Third, different sets of economic institutions are feasible in a predominantly agrarian economy, which influence the relative intensity of elites' and citizens' preferences over different regimes. For instance, labor-repressive institutions, such as slavery, are relatively more efficient with agricultural technology than in industry (Eltis 2000). This implies that democracy is worse for elites because the changes in collective choices that it brings undermine their preferred set of economic institutions. All three considerations imply that democratization is more likely in a more industrialized society where the elite own significant physical and human capital than a more agricultural society where the elites are mainly invested in land. Stated differently, democracy is more likely when the elites are industrialists rather than landowners.

Although the nature of revolutions is not the focus of this book, these ideas also have interesting implications for the incidence of revolutions. For example, they can help account for why most revolutions – for example, in Russia, Mexico, China, Vietnam, Bolivia, and Nicaragua – take place in primarily agrarian societies. We suggest that this is because landed elites favor repression rather than concessions and, when repression fails, revolutions take place. In more urbanized and industrialized societies, where the elites are invested in capital, concessions are favored and revolutions are observed less often.

6.3.2 Consolidation

The source of income for the elites also impacts the decision of whether to mount a coup. If the elites are heavily invested in land, then coups may tend to be less costly. More important, democracy is relatively worse for such individuals given that land can be taxed at higher rates than capital, and also that economic institutions under democracy are further from those preferred by the elites. In contrast, when the elites' wealth is mostly in the form of physical and human capital, coups are more expensive for them and democracy is less threatening. As a result, democracy is less likely to consolidate when the elites are landowners than when they are capitalists.

6.4 Political Institutions

6.4.1 Democratization

Our framework also suggests that the nature of democratic political institutions may be crucial for explaining why some societies democratize but others do not. In particular, when the elites can use repression to avoid democratizing, they do so because they anticipate that democracy will be harmful for their interests. So far, our characterization of democracy as the rule of the majority has been overly stylized in order to communicate the main elements of our analysis. In reality, one person's vote may be worth more than another's and, in particular, the elites may be able to exercise more or less influence over what happens in a democracy – even though their influence is relatively less than it is in a dictatorship.

One way they can do this is through the design of democratic institutions. In his 1913 book, *An Economic Interpretation of the U.S. Constitution*, Beard argued that the constitution was written by rich property holders with an eye to maintaining the worth of their assets (including, one should add, their slaves) in the face of likely radical democratic pressures.[4] Beard argued that

> inasmuch as the primary object of a government, beyond the mere repression of physical violence, is the making of the rules that determine the property relations of society, the dominant classes whose rights are thus to be determined must perforce obtain from the government such rules as are consonant with the larger interests necessary to the continuance of their economic processes, or they must themselves control the organs of government. In a stable despotism the former takes place; under

[4] Although many details of Beard's arguments are now contested, the general thrust of his argument is accepted by many scholars. For instance, Wood (1969, p. 626) notes in his seminal book that the constitution "was intrinsically an aristocratic document designed to check the democratic tendencies of the period." See McGuire (1988) for partially supporting statistical evidence.

> any other system of government, where political power is shared by any portion of the population, the methods and nature of this control become the problem of prime importance – in fact, the fundamental problem in constitutional law. The social structure by which one type of legislation is secured and another prevented – is a secondary or derivative feature arising from the nature of the economic groups seeking positive action and negative restraint. (1913, p. 13)

Even the notion of representative democracy, as opposed to participatory or direct democracy, can be seen as an attempt to dilute populist pressures and undermine the power of the majority (as argued by Manin 1997).

Clearly, then, democratic political institutions can be structured to limit the power of the majority. A more recent example is the constitution written during the dictatorship of General Pinochet in Chile, which attempted to minimize the threat of socialism in Chile by engaging in systematic gerrymandering and the underrepresentation of urban areas, and which otherwise attempted to cement the veto of the military over democratic decision making (Londregan 2000; Siavelis 2000).

Another example, discussed in Chapter 1, is the way that the South African constitution was written in an attempt to protect the interests of whites under democracy.

If a nondemocratic regime or elite can design or manipulate the institutions of democracy so as to guarantee that radical majoritarian policies will not be adopted, then democracy becomes less threatening to the interests of the elites. Less threatened, the elites are more willing to create democracy in the first place. For instance, when democracy is less threatening, it will be less attractive to use repression to avoid it. Thus, Pinochet's constitution, according to our framework, facilitated democratization in Chile. It may even be the case that, as in South Africa, the majority of citizens are themselves willing to restrict their policy options to facilitate a transition to democracy. As we discuss in Chapter 6, the ANC realized that it had to make concessions to the whites about the structure of democratic institutions. For the ANC, this was better than carrying on with the fight against the apartheid regime. By giving the elite credible guarantees, a process of democratization is facilitated that might otherwise not take place.

6.4.2 Consolidation

Just as the structure of democratic institutions influences democratization in the first place, so it helps to determine whether democracy consolidates. In particular, institutions that place limits on pro-majoritarian policies in democracy are likely to help consolidation. In fact, the elites may be quite influential in democracy because they control a strong upper house, like the Prussian Junkers in nineteenth-century Germany, or the British aristocracy in the House of Lords, or because they control the party system. Knowing that in democracy they will be able to insure against the most excessively majoritarian policies, the elites will be less willing to undertake action against democracy.

An interesting example in this context is the links between the elite and both traditional ruling parties in Colombia. Throughout the twentieth century, the Liberal and Conservative Parties managed successfully to avoid the entry of left-wing parties by manipulating electoral institutions, particularly the form of proportional representation. Without a left-wing party, highly redistributive political agendas did not emerge in Colombia. Interestingly, Colombia has one of the most consolidated democracies in Latin America, although there are often complaints that the system does not represent the interests of the majority.

Another example of the connection between political institutions and democratic consolidation is the claim that presidential democracies may be more unstable than parliamentary democracies and more prone to coups (Linz 1978, 1994). This idea makes sense in our framework because, whereas in a legislature checks and balances and lobbying may allow the elites to block radical policy proposals, a directly elected president is more likely to represent the preferences of the majority in society and, therefore, to be more populist. Hence, presidential systems may be more threatening to the interests of the elites and thus induce more coups.

Paradoxically, then, this perspective might also help explain why the consolidation of democracy in Chile may have run smoothly after the systematic gerrymandering that General Pinochet arranged in the electoral rules. This manipulation underrepresented urban areas at the expense of more conservative rural areas, thus reducing the political power of the left. The consequence was a less redistributive but more stable democracy. Turkey and Thailand provide other examples in which constitutions written or commissioned by the military may have helped democratic consolidation. Haggard and Kaufman (1995, p. 110) note:

> Ironically, the greater security for the armed forces during the initial years of the transition probably *reduced* the threat to civilian authority in Chile, Turkey, and Korea.

However, whereas increasing the power of the elites in democracy may promote democracy, giving the elites too much power will undermine it. In our framework, democracy arises from conflict between elites and disenfranchised majorities who are prepared to accept democracy rather than something more radical because it gives them more political power than nondemocracy. If the elites have too much power in democracy, democracy will do little to improve the welfare of the majority. In this case, democracy is not a solution to social conflict, and the result will either be revolution or an elite that keeps itself in power through repression.

6.5 The Role of Inter-Group Inequality

6.5.1 Democratization

Our framework makes predictions about the effect of *inter-group inequality* – inequality between groups – on the creation and consolidation of democracy. For convenience, we outline these using the word *inequality* to refer to inter-group inequality. However, these predictions about inter-group inequality may

not translate into statements about standard measures of inequality and income distribution (e.g., the labor share or the Gini coefficient). This is particularly true when political conflict is not rich versus poor but rather along other lines, perhaps between ethnic or religious groups.

Everything else being equal, greater inter-group inequality makes revolution more attractive for the citizens: with revolution, they get a chance to share the entire income of the economy (minus what is destroyed in revolution), whereas in nondemocracy, they obtain only a small fraction of these resources. Because an effective threat of revolution is the spark that ignites the democratization process, greater inter-group inequality should be associated with a greater likelihood of democratization.

There is also another reason why inter-group inequality might contribute to democratization. Recall that democratization occurs as a credible commitment to future redistribution, when the promise of redistribution is not sufficient to stave off the threat of revolution. The stronger the threat of revolution, the more likely it is that this promise will be insufficient and that the elite will be forced to create democracy. Because greater inter-group inequality contributes to the strength of the threat of revolution, it makes democratization more likely via this channel as well.

This discussion of the role of inter-group inequality is one-sided, however. It highlights how greater inequality increases the threat of revolution and thus the demand for democracy by the citizens. However, inter-group inequality may also affect the aversion that the elites have to democracy. To see why consider a standard model of redistributive taxation as in Meltzer and Richard (1981). Note that as the gap between the elites and the citizens rises (i.e., as inter-group inequality increases), the burden placed on the elites, even *at a constant tax rate*, rises. This is because with greater inequality, a larger share of total tax revenues will be raised from the elites, who now command a greater fraction of the resources in the economy. Therefore, greater inter-group inequality typically increases the burden of democracy on the elites, even if the tax rate remains constant or changes little. Moreover, many approaches suggest that greater inter-group inequality should increase the tax rate, contributing to this effect. If this is so, there would be another reason for greater inequality to increase the burden of democracy on the elites. With greater inequality, the benefits from redistribution increase, inducing the citizens to prefer higher levels of taxation.[5] Overall, therefore, it seems compelling that the costs of redistributive taxation and democratic politics to the elites and, hence, their aversion to democracy should be generally higher for the elites in a society where the difference in incomes between the elites and the citizens is greater.

[5] As discussed in Chapter 4, there are theoretical and empirical arguments for why the relationship between inequality and redistribution may be more complex (e.g., greater inequality may enable the elites to lobby more effectively against redistribution in democracy). Nevertheless, it is generally the case that with greater inter-group inequality, democracy imposes a greater burden on the elites than nondemocracy does.

How does this affect the relationship between inter-group inequality and transition to democracy? The most important implication is that as inequality increases and democracy becomes more costly for the elites, repression becomes more attractive. Therefore, greater inter-group inequality may also discourage democratization.

Putting these two pieces of the story together, we find that there is a nonmonotonic (i.e., inverted U-shaped) relationship between inter-group inequality and the likelihood of transition to democracy. In the most equal societies, revolution and social unrest are not sufficiently attractive for the citizens; either there are no challenges to nondemocratic systems or any challenges can be met by temporary measures, such as some limited redistribution. In other words, in these fairly equal societies, the citizens are already benefiting from the productive resources of the economy or even perhaps from the growth process, so they do not make further strong demands. This may be the reason why democracy arrived late in a number of equal and rapidly growing economies, such as South Korea and Taiwan, and has yet to fully arrive in Singapore. In stark contrast, in the most unequal societies (e.g., South Africa prior to 1994), the citizens have great reason to be unhappy and often try to rise up against the authority of nondemocracy. Now, however, the elites have a lot to lose from abandoning the system that looks after their own interests and transitioning into one that will place a greater redistributive burden on them. Thus, instead of democracy, a highly unequal society is likely to result in a repressive nondemocracy – or, sometimes when repression is not enough, perhaps even experience a revolution. This mechanism can also explain the persistence of nondemocratic regimes in the highly unequal countries of Latin America, such as El Salvador and Paraguay. This account, then, suggests that democracy has the best chance to emerge in societies with middle levels of inequality. Here, the citizens are not totally satisfied with the existing system, and the elites are not so averse to democracy that they resort to repression to prevent it. This is the situation we find in Britain and Argentina in the late nineteenth and early twentieth centuries.

6.5.2 Consolidation

Inequality also critically influences the propensity of a democracy to consolidate. Because the main threat against democracy comes from its redistributive nature, the greater redistribution away from the elites the more likely they are to find it in their interest to mount a coup against it. Therefore, greater inequality is likely to destabilize democracy because, as observed previously, the burden of democracy on the elites is increasing in the income gap between them and the citizens.

This comparative static result with respect to inequality offers a potential explanation for why democracy may have been more difficult to consolidate in Latin America than in Western Europe. Latin American societies are considerably more unequal and, therefore, suffer more from distributional conflict between the elites and the citizens. Our framework predicts that in highly unequal societies, democratic policies should be highly redistributive but then abruptly come to an end

with a coup that reverts back to much less redistributive policies. This pattern is reminiscent of the oscillations of many Latin American countries between the highly redistributive but unsustainable populist policies of short-lived democracies and the fiscally more conservative approach of subsequent nondemocratic regimes. Tellingly, Kaufman and Stallings (1991, p. 27) also emphasize a close connection between unconsolidated democracy and populist redistribution:

> ... established democracies (Venezuela, Colombia and Costa Rica in our study) were also associated with orthodox macro policies. ... It was the transitional democracies (Peru, Argentina and Brazil) that followed populist policies.

Combining the effects of inequality on democratization and coups, we can see that equal societies never democratize in the first place. This helps to account for Singapore's path of political development. Higher but still relatively low levels of inter-group inequality lead societies to democratize and, once created, democracy is consolidated because it is not so costly for the elites that a coup is desirable. This may capture Britain's path of political development. Even higher levels of inequality still lead to democratization, but democracy does not consolidate because coups are attractive. As a result, the outcome is unconsolidated democracy, which is the path that Argentina followed in the twentieth century. Finally, at the highest levels of inequality, democracy is so threatening for the elites that they use repression to avoid it, a situation that characterized South Africa until 1994.[6]

6.6 The Middle Class

6.6.1 Democratization

Perhaps the most famous treatise on the origins of democracy is Moore's (1966) *Social Origins of Dictatorship and Democracy*. Our work owes a natural intellectual debt to Moore, especially because we paraphrased his title. In our theory, the major factor that distinguishes democracy from nondemocracy is the greater political equality of democracies; so far, we have only distinguished between two groups: the elites and the citizens. This was mainly for simplification (again, an application of Occam's razor). Nevertheless, in many circumstances, a third group between the elite and the great mass of citizens may be of significance. In general, this group could be identified in different ways but, following the emphasis of many scholars, it is useful to think of this group as the middle class forming a distinct political actor. When the middle class is brought into our framework, we obtain a range of interesting results, some of them vindicating the emphasis that Moore and other scholars placed on the middle class.

The first role that the middle class can play in the emergence of democracy is as the driver of the process. Recall that in our framework, democracy emerges in

[6] These are all, of course, statements where "other things are held equal." Inter-group inequality is not the only thing that determines whether a society democratizes or a democracy consolidates.

response to a serious revolutionary threat or significant social unrest. The middle class can be the driver in this process by playing a key role in the revolutionary movement or by fueling and maintaining it. Almost all revolutionary movements were led by middle-class actors and, more important, a number of the major challenges to the existing regime; for example, the uprisings that helped induce the First Reform Act in Britain or those during the Paris Commune in France or the revolts of the Radical Party in Argentina were largely middle-class movements (see O'Donnell and Schmitter 1986, pp. 50–2, on the crucial role of the middle class in contemporary democratizations). Therefore, the middle class – by virtue of its more comfortable economic situation and the greater education of its members – can be a critical catalyst in the process toward democracy. This might also explain why many of the early moves toward democracy in Europe were only partial. If the middle class is the key actor, it may be sufficient for the elites to co-opt the middle class rather than concede a comprehensive democracy to all those who are excluded from the political system. The resulting picture resembles the gradual move toward democracy experienced in much of Western Europe: first, the middle classes are included in the political process and then the franchise is extended to the mass of citizens.

Perhaps the more important role of the middle class is that of a *buffer* in the conflict between the elites and the citizens. Recall that when the elites expect democracy to adopt policies highly unfavorable to them, they prefer repression to democratization. The presence of a large and relatively affluent middle class ensures that they play an important role in democratic politics and, because they are more prosperous than the citizens, they will typically support policies much closer to those that the elites prefer. Therefore, by limiting the amount of policy change induced by democracy, a large and affluent middle class may act like a buffer between the elites and the citizens in democracy. It does this by simultaneously making democratization more attractive for the elites than repression and changing policy enough that the citizens are content not to revolt.

The role of the middle class in the transition to democracy might help us understand the contrast between the political histories of Costa Rica and Colombia on the one hand and Guatemala, El Salvador, and Nicaragua on the other. Despite many similarities in their colonial histories and economic structures, these five countries have had very different political trajectories (Paige 1997; Nugent and Robinson 2002). Costa Rica and Colombia have become stable albeit restricted democracies since the middle of the nineteenth century and successfully made the transition to effective universal suffrage in 1948 and 1936, respectively. Guatemala, El Salvador, and Nicaragua, on the other hand, were dominated by dictators in the nineteenth century and initial moves toward democracy – for example, in El Salvador in the late 1920s and in Guatemala between 1945 and 1954 – were snuffed out by coups and repression. These three societies made the transition to democracy very late. One important difference among these countries is that there is a relatively large and affluent middle class, especially smallholder

coffee producers, in Costa Rica and Colombia but not in the other three. Perhaps as a consequence, democratic politics, once installed, has been much more conflict-ridden in Guatemala, El Salvador, and Nicaragua than in Costa Rica and Colombia.

6.6.2 Consolidation

The middle class may play an important role in consolidating democracy by limiting redistribution. A society with a large and affluent middle class will engage only in limited redistribution away from the elites toward the citizens and, therefore, provide a much smaller threat to the interests of the elites. This might be useful in understanding why many Western European and some Latin American societies, like Costa Rica and Colombia, with comparatively large middle classes have also had relatively stable democracies, whereas El Salvador and Guatemala, which lack such a middle-class buffer, have had difficulty consolidating democracy.

6.7 Globalization

There is no doubt that there are stronger economic links between nations today than forty years ago. Countries are more closely linked internationally today, with economic organizations such as the European Union, NAFTA, Mercosur, and Asean; there are much larger volumes of goods and services being traded, and much larger cross-border financial transactions. Do these major economic and political changes have implications for the circumstances under which democracy will arise and consolidate?

6.7.1 Democratization

Globalization might contribute to democratization in a number of distinct ways. First, international financial integration means that capital owners, the elites, can more easily take their money out of a given country. This makes it more difficult to tax the elites and reduces the extent to which democracy can pursue populist and highly majoritarian policies. International financial integration, therefore, makes the elites feel more secure about democratic politics and discourages them from using repression to prevent a transition from nondemocracy to democracy.

Second, international trade affects factor prices and, via this channel, modifies redistributive politics. Countries differ in their factor endowments, and the relative abundance of factors of production determines patterns of specialization and the impact of trade on relative prices. One implication of increased international trade is an increase in the rewards to the relatively abundant factor in each country. In the case of less developed nations – which are typically those still in nondemocracy today and, therefore, the main candidates for democratization – this means an increase in the rewards to labor. Intuitively, before the advent of significant trade flows, less developed countries had an excess of labor and a shortage of capital, depressing the rewards to labor and increasing those to capital. Trade opening will pull these rewards toward those prevailing in the rest of the world, thus increasing the rewards to labor and potentially reducing the return to capital. Trade opening will, therefore, reduce the gap between the incomes of labor and capital, thus changing the extent of inequality between capital owners and labor owners.

The specific implications of our framework depend on three things: (1) the nature of relative factor abundance; (2) the nature of political identities; and (3) where a country is on the inverted U-shaped relationship between inter-group inequality and democratization. Imagine that nondemocratic countries are labor abundant, political conflict is between rich capital-owning elites and poor labor-owning citizens, and inequality is sufficiently high that the elites use repression to stay in power. In this case, increased trade integration will reduce the extent of inequality between the elites and the citizens and will make democracy less redistributive. Because democracy will then be less threatening to the elites, they will be less inclined to use repression to avoid democracy. In such circumstances, globalization promotes democracy. Nevertheless, our framework does not imply that the impact of globalization on factor prices always promotes democracy. Let's continue to postulate that conflict is between the rich and the poor and that we are on the part of the inverted U-shaped relationship where the rich use repression to stay in power. Now consider Latin American countries such as Argentina, Brazil, and Uruguay in the late nineteenth century. In these countries, the elites owned a lot of land and they were also land abundant. As predicted by the theory of international trade, pre–First World War globalization led to large increases in returns to land (O'Rourke and Williamson 1999). In our framework, this increases inter-group inequality and makes the elites less likely to democratize. It also increases the proportion of elite wealth invested in land, another factor that we suggest makes democracy more threatening to the elites. By the converse of these arguments, in this case, globalization would impede democratization (as long as we are on the part of the inverted U-shaped relationship where inequality discourages democratization, as assumed previously).

Third, increased international trade also means that disruption of economic activity may become more costly for many less developed nations that are now integrated into the world economy and, therefore, repression may now be much more costly for the elites, again favoring democracy.

Finally, increased political integration and the end of the Cold War (if not hijacked by the war against terrorism) might imply that countries that repress their citizens can perhaps expect stronger sanctions and reactions from the democratic world. This effectively increases the costs of repression, promoting democracy. This might be especially important because a number of nondemocratic regimes in the Cold War Era, such as Mobutu's disastrous dictatorship in Zaire, were kept alive by the explicit or implicit support of the international community.

6.7.2 Consolidation

Just as globalization can induce democratization, so it can aid democratic consolidation. Indeed, all of the mechanisms listed that link increased globalization to democratization also imply that coups will be less likely. This is either because coups become more costly in a more integrated world or because globalization implies that democracy is less threatening to the elites.

7. Political Identities and the Nature of Conflict

Most of the comparative static results discussed so far do not depend on the identity of the elite; they apply even in societies where the nature of political conflict is not along class lines. In South Africa, race may be more salient, although race and socioeconomic class overlap to a large extent. In Rwanda, it may be more plausible to think of groups forming along the lines of ethnicity: Hutu or Tutsi. In Mauritius, political conflict has been between people of East Indian descent and a heterogeneous coalition of others, some of whom are rich (i.e., the white sugar planters and Chinese business elites) and some very poor (i.e., mostly the descendents of African slaves). In the latter case, there is no simple overlap between ethnicity or race and class (Bowman 1991).

As long as one accepts the premise that the interests of individuals are partly about economic outcomes, our basic analysis remains unaltered. Consider our ideas about political institutions. Here, we showed that if political institutions were such as to limit the type of policies that could occur in democracy, they tended to induce consolidated democracy. This result applies even in Mauritius. If institutions limit democracies, then they limit what the East Indian majority can do to the Creole minority. Hence, they reduce the incentive of a Creole dictatorship to repress democracy and, once democracy has been created, they make coups less attractive – exactly as in our previous analysis.

Next, consider the ideas we developed about the connection between the composition of the wealth of the elite and democratization or coups. These ideas apply immediately in this case. Even when politics is East Indian against Creole, as the economy develops and capital becomes more important than land, repression and coups become more costly and (pro–East Indian) democracy becomes less redistributive. As in our baseline analysis, this tends to create a consolidated democracy, even in Mauritius. Interestingly, Mauritius has been a consolidated democracy since independence, and this process of consolidation has taken place in the context of the radically declining importance of land, the rapid development of industry, and the expansion in the importance of human capital.

The nature of political identities may undoubtedly influence the form of collective choice under democracy, which ties our analysis to several important traditions in political science. For example, contrast a society where political identities and cleavages are on the basis of class with one where there are many cross-cutting cleavages or race, ethnicity, religion, or region. The pluralist model of

democracy is one in which society is indeed divided into many different groups. A standard claim about a pluralistically based society is that it generates less income redistribution and smaller welfare states because the many different cleavages stop a broad coalition in favor of redistribution emerging. In consequence, for instance, pluralistic societies do not have strong socialist parties (Lipset and Marks 2000). If this is the case, then our theory suggests that such societies would be more likely to have consolidated democracy because elites would have little to fear from majority rule. This helps explain the longevity and stability of democracy in the United States, often thought to be the epitome of a pluralistic society.

8. Democracy in a Picture

The previous discussion illustrates the various empirical implications of our theory for the circumstances under which a society becomes and stays a democracy. To fix ideas, it is useful to use simple pictures to illustrate the circumstances under which different regimes arise. To map our comparative statics about inter-group inequality into the data, we do this in the context in which the elites are the rich and the citizens are the poor. Recall the four "paths" of political development that we sketched in Chapter 1. The first, the British path, was one of fully consolidated democracy. The second, the Argentine path, was that of unconsolidated democracy. The third path, that of Singapore, was persistent nondemocracy in which the political status quo can be sustained without serious repression. The fourth path, that of South Africa, was persistent nondemocracy with repression. The comparative statics of our theory allow us to depict these different outcomes in a picture.

In essence, the different political outcomes occur because these societies differ fundamentally in their underlying economic structures, and it is this that motivates the title of our book. In addition, we also emphasize differences in political institutions, to some extent historically determined, to some extent consciously chosen with the nature of the regime in mind. To keep the pictures simple, we assume that conditions are such that a revolution never occurs in equilibrium, and we also abstract from the use of concessions (e.g., they are always insufficiently credible to stop revolt) so that if revolution is a threat, a nondemocratic regime must choose between repression or conceding democracy.

Consider Figure 2.1, which captures the predictions of our theory for democratization. On the horizontal axis, we plot inequality, with moves from left to right corresponding to greater inequality. The origin represents a completely equal society. On the vertical axis, we plot the historically determined costs of repression that are exogenous, such as the extent to which repression destroys assets in society. We have divided the resulting square into different regions that represent how different structures lead to different paths of political development. When

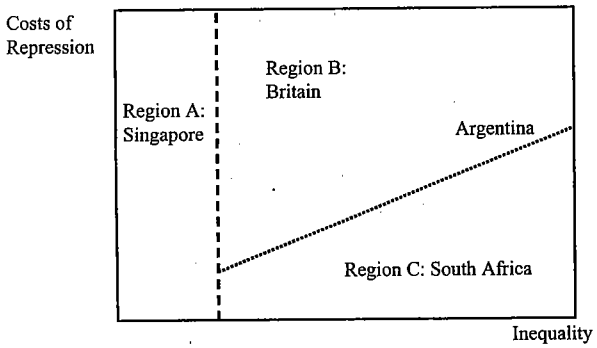

Figure 2.1. Democratization.

inequality is sufficiently low, the cost of repressing or mounting coups is irrelevant because the poor are sufficiently content under the political status quo not to rock the boat. This corresponds to Region A, where there is nondemocracy that remains unchallenged, and in it we place Singapore. In Region B, inequality is higher and revolution becomes a threat. However, the cost of repression is sufficiently high that democracy is created. In this region, we place both Britain and Argentina. Finally, in Region C, inequality is so high that revolution is a threat to nondemocracy but the cost of repression is sufficiently low that democracy can be avoided. This is the case of South Africa until 1994. The cost of repression in South Africa might have been lower because the disenfranchised groups were black Africans and Coloureds, and exclusion and repression were justified by an explicitly racial philosophy.

To study the consolidation of democracy, we must turn to Figure 2.2. Here, there are just two regions separated by an upward-sloping curve. When the cost of a coup is zero, the rich are always willing to undertake a coup. However, as the cost of a

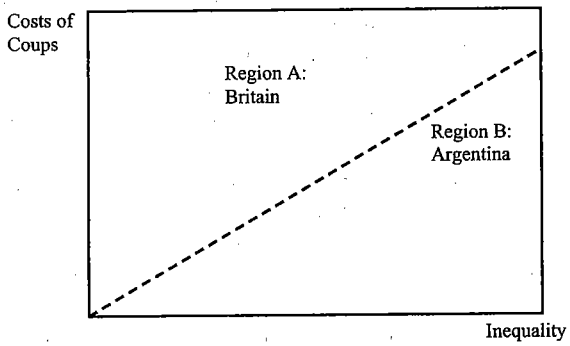

Figure 2.2. Democratic Consolidation.

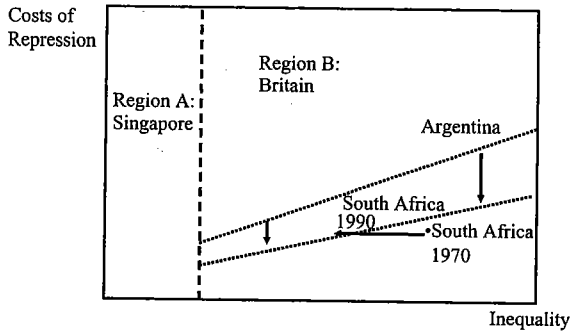

Figure 2.3. Democratization in South Africa.

coup rises, inequality must be sufficiently high (i.e., democracy must be sufficiently costly to the rich) for it to be worthwhile. Figure 2.3 has just two regions. We have placed Britain in Region A. Once created, democracy will consolidate if it is not too redistributive and if coups are sufficiently costly. However, when inequality is very high, the costs of a coup may be sufficiently low that it is attractive. This is the case in Region B, where democracy is unconsolidated; here, we have placed Argentina. Singapore is, of course, not in this picture because it has yet to make the transition to democracy.

These simple pictures also allow us to trace out the paths of political development of different countries. For example, we look at the history and future of democracy in South Africa in Figure 2.3. Why did South Africa finally move so belatedly to a democratic regime? The arrows in Figure 2.3 capture part of the story. From the mid-1970s onward, inequality fell in South Africa, making democracy less threatening for the white elite. At the same time, the industrial sector rose at the expense of the agricultural sector, and human and physical capital became more important. In terms of the picture, this means that at a given level of inequality, the elites are less willing to repress. This moves the boundary between Region C and Region B downward. Changes in the global environment, particularly globalization, also have the effect of moving the same boundary downward, implying that for fixed levels of inequality, the cost of repression had to be lower to justify the persistence of dictatorship. Thus, some time between 1970 and 1994, South Africa moved out of Region C into Region B, and democracy was created.

What does the future hold for South Africa? To see this, we must turn to Figure 2.4, which asks whether democracy will consolidate after apartheid. Because inequality is still very high, one might conjecture that South Africa would be in Region B and, therefore, an unconsolidated democracy. Nevertheless, the impact of development of the South African economy, increasing importance of physical and human capital, and increased globalization has the effect of moving the boundaries between Regions A and B down. Now, for a given cost of a

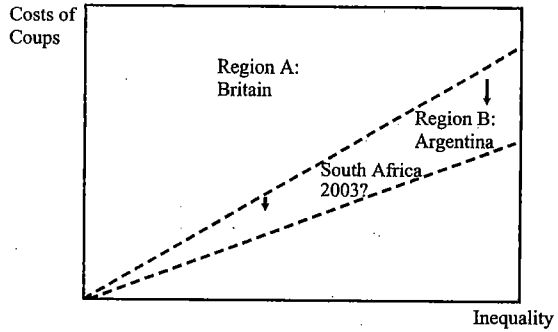

Figure 2.4. Democratic Consolidation in South Africa?

coup, inequality must be higher to justify mounting a coup against democracy. Moreover, as discussed in Chapter 1, the structure of political institutions after apartheid was designed specifically to protect the interests of the whites, a factor that again moves this line down. Thus, although one cannot be certain of the future (witness the evolution of democracy in Zimbabwe since 1980), one might hope that South Africa had transitioned into Region A rather than Region B.

Prediction in the case of Singapore seems much easier. Figures 2.1 and 2.2 suggest that if and when Singapore becomes a democracy, it is very likely to consolidate.

9. Overview of the Book

The remainder of our book develops the arguments outlined in this chapter. The remainder of this part continues to lay the scene. In Chapter 3, we survey the empirical evidence about cross-country patterns of democracy. We show that richer countries are more likely to be democratic, more educated countries are more likely to be democratic, and more unequal countries are generally less democratic. We emphasize the basic correlations in the data and do not take a strong view on causal relationships. Chapter 3 also discusses the large literature in political science and sociology on the creation and consolidation of democracy, and we explain how our research contributes to this work.

Part 2 surveys existing models of collective decision making in democracies and nondemocracies. In Chapter 4, we focus on democracies and provide a simple analysis of basic issues in the study of collective choice, electoral politics, and competition, which is useful in later parts of the book. We also introduce some basic models of two-group distributional conflict, paying special attention to the relationship between inequality and redistribution, the implications of different political identities, and the factors that determine the distribution of power in

democracy. In Chapter 4, we also propose a reduced-form model of the distribution of power in a democracy. The appendix at the end of the book develops a series of models that provide microfoundations for this reduced form. In Chapter 5, we analyze nondemocracy with particular attention to the collective-action problem and the issue of commitment.

Part 3 provides our approach to democratization. In Chapter 6, we introduce our basic model of democratization. This chapter formalizes many of the issues already mentioned in this introductory chapter, giving us ways to think about the role of political power and the role of political institutions in allocating future political power. It illustrates how democratization creates a credible commitment to future redistribution by transferring political power to the majority in society. It also shows how democratization may be a response by the elite in the face of a credible threat of revolution by the majority. We see the possibility of an inverted U-shaped relationship between inter-group inequality and democracy in this chapter. Chapter 7 then develops our basic model of coups against democracy and studies the circumstances under which democracy, once created, consolidates.

Part 4 discusses a number of important extensions to this basic framework and some applications. In Chapter 8, we analyze how the presence of a large and affluent middle class affects the balance of the distributional conflict between the elites and the citizens in ways that can help create and consolidate democracy. Chapter 9 introduces factor endowments and markets to endogenize the distribution of income and discusses the impact of the structure of the economy on the creation and consolidation of democracy. In this chapter, we also conjecture about mechanisms that might account for political development – that is, the question of why and whether countries transition to democracy as they become richer, and the potential reasons for the relationship between income and democracy. Chapter 10 extends our model to allow for international trade and mobility of factors of production among countries and studies how globalization alters and adds to the results we have derived until this point.

Part 5 discusses the future of democracy and concludes the book.