From: S. Castles and M. J. Miller, The Age of Migration (3rd ed.), New York: Guilford Press, 2003.

Chapter 2

The Migratory Process and the Formation of Ethnic Minorities

International migration is hardly ever a simple individual action in which a person decides to move in search of better life-chances, pulls up his or her roots in the place of origin and quickly becomes assimilated in the new country. Much more often migration and settlement is a long-drawn-out process, which will be played out for the rest of the migrant's life, and affect subsequent generations too. (Migration can even transcend death: members of some migrant groups have been known to arrange for their bodies to be taken back for burial in their native soil: see Tribalat, 1995: 109–11.) Migration is a collective action, arising out of social change and affecting the whole society in both sending and receiving areas. Moreover, the experience of migration and of living in another country often leads to modification of the original plans, so that migrants' intentions at the time of departure are poor predictors of actual behaviour. Similarly, no government has ever set out to build an ethnically diverse society through immigration, yet labour recruitment policies often lead to the formation of ethnic minorities, with far-reaching consequences for social relations, public policies, national identity and international relations.

The aim of the chapter is to link two bodies of theory which are often dealt with separately: theories on migration and settlement, and theories on ethnic minorities and their position in society. This chapter provides a theoretical framework for understanding the more descriptive accounts of migration, settlement and minority formation in later chapters. However, the reader may prefer to read those first and come back to the theory later.

Explaining the migratory process

The concept of the *migratory process* sums up the complex sets of factors and interactions which lead to international migration and influence its course. Migration is a process which affects every dimension of social existence, and which develops its own complex dynamics.

Research on migration is therefore intrinsically interdisciplinary: sociology, political science, history, economics, geography, demography, psychology and law are all relevant (Brettell and Hollifield, 2000). These

disciplines look at different aspects of population mobility, and a full understanding requires contributions from all of them. Within each social-scientific discipline there is a variety of approaches, based on differences in theory and methods. For instance, researchers who base their work on quantitative analysis of large data-sets (such as censuses or representative surveys) will ask different questions and get different results from those who do qualitative studies of small groups. Those who examine the role of migrant labour within the world economy using historical and institutional approaches will again get different findings. All these methods have their place, as long as they lay no claim to be the only correct one. A detailed survey of migration theory is not possible here (see Massey *et al.*, 1993, 1994, 1998), but a useful distinction may be made between three of the main approaches used in contemporary debates: economic theory, the historical-structural approach and migration systems theory (Hugo, 1993: 7–12).

Economic theories of migration

The neo-classical economic perspective has its antecedents in the earliest systematic theory on migration: that of the nineteenth-century geographer Ravenstein, who formulated statistical laws of migration (Ravenstein, 1885, 1889). These were general statements unconnected with any actual migratory movement (Cohen, 1987: 34–5; Zolberg, Suhrke and Aguao, 1989: 403–5). This tradition remains alive in the work of many demographers, geographers and economists. Such 'general theories' emphasize tendencies of people to move from densely to sparsely populated areas, or from low- to high-income areas, or link migrations to fluctuations in the business cycle. These approaches are often known as 'push–pull' theories, because they perceive the causes of migration to lie in a combination of 'push factors', impelling people to leave the areas of origin, and 'pull factors', attracting them to certain receiving countries. 'Push factors' include demographic growth, low living standards, lack of economic opportunities and political repression, while 'pull factors' are demand for labour, availability of land, good economic opportunities and political freedoms.

This model is mainly found in neo-classical economics, although it has also been influential in sociology, social demography and other disciplines. It is individualistic and ahistorical. It emphasizes the individual decision to migrate, based on rational comparison of the relative costs and benefits of remaining in the area of origin or moving to various alternative destinations. Constraining factors, such as government restrictions on emigration or immigration, are mainly dealt with as distortions of the rational market,

which should be removed. Its central concept is 'human capital': people decide to invest in migration, in the same way as they might invest in education or vocational training, because it raises their human capital and brings potential future gains in earnings. People will migrate if the expected rate of return from higher wages in the destination country is greater than the costs incurred through migrating (Chiswick, 2000). Borjas puts forward the model of an immigration market:

> Neo-classical theory assumes that individuals maximize utility: individuals 'search' for the country of residence that maximizes their well-being ... The search is constrained by the individual's financial resources, by the immigration regulations imposed by competing host countries and by the emigration regulations of the source country. In the immigration market the various pieces of information are exchanged and the various options are compared. In a sense, competing host countries make 'migration offers' from which individuals compare and choose. The information gathered in this marketplace leads many individuals to conclude that it is 'profitable' to remain in their birthplace ... Conversely, other individuals conclude that they are better off in some other country. The immigration market nonrandomly sorts these individuals across host countries. (Borjas, 1989: 461)

Borjas claims that 'this approach leads to a clear – and empirically testable – categorization of the types of immigrant flows that arise in a world where individuals search for the "best" country' (Borjas, 1989: 461). On this basis, the mere existence of economic disparities between various areas should be sufficient to generate migrant flows. In the long run, such flows should help to equalize wages and conditions in underdeveloped and developed regions, leading towards economic equilibrium. Borjas has argued that this may lead to negative effects for immigration countries, notably the decline of average skill levels (Borjas, 1990). However, this finding is not uncontested within neo-classical research: Chiswick claims that migrants are positively self-selected, in the sense that the higher skilled are more likely to move because they obtain a higher return on their human capital investment in mobility. This has negative effects for countries of origin, by causing a 'brain drain' (Chiswick, 2000).

Empirical studies cast doubt on the value of neo-classical theory. It is rarely the poorest people from the least-developed countries who move to the richest countries; more frequently the migrants are people of intermediate social status from areas which are undergoing economic and social change. Similarly the push–pull model predicts movements from densely populated areas to more sparsely peopled regions, yet in fact countries of immigration like the Netherlands and Germany are among the world's

more densely populated. Finally the push–pull model cannot explain why a certain group of migrants goes to one country rather than another: for example, why have most Algerians migrated to France and not Germany, while the opposite applies to Turks?

Neo-classical migration theories have therefore been criticized as simplistic and incapable of explaining actual movements or predicting future ones (see Sassen, 1988; Boyd, 1989; Portes and Rumbaut, 1996: 271–8). It seems absurd to treat migrants as individual market-players who have full information on their options and freedom to make rational choices. Historians, anthropologists, sociologists and geographers have shown that migrants' behaviour is strongly influenced by historical experiences as well as by family and community dynamics (Portes and Böröcz, 1989). Moreover migrants have limited and often contradictory information, and are subject to a range of constraints (especially lack of power in the face of employers and governments). Migrants compensate through developing cultural capital (collective knowledge of their situation and strategies for dealing with it) and social capital (the social networks which organize migration and community formation processes).

It therefore seems essential to introduce a wider range of factors into economic research. One attempt to do this was 'dual labour market theory', which showed the importance of institutional factors as well as race and gender in bringing about labour market segmentation (Piore, 1979). The 'new economics of labour migration' approach emerged in the 1980s (Stark, 1991; Taylor, 1987). It argued that markets rarely function in the ideal way suggested by the neo-classicists. Migration needs to be explained not only by income differences between two countries, but also by such factors as the chance of secure employment, availability of investment capital, and the need to manage risk over long periods. For instance, as Massey *et al.* (1987) point out, Mexican farmers may migrate to the USA because, even though they have sufficient land, they lack the capital to make it productive. Similarly, the role of remittances in migration cannot be understood simply by studying the behaviour of migrants themselves. Rather it is necessary to examine the long-term effects of remittances on investment, work and social relationships right across the community (Taylor, 1999).

The neo-classical model tends to treat the role of the state as an aberration which disrupts the 'normal' functioning of the market. Borjas, for instance, suggests that the US government should 'deregulate the immigration market' by selling visas to the highest bidder (Borjas, 1990: 225–8). But examination of historical and contemporary migrations (see Chapters 3–7 below) shows that states (particularly receiving countries) play a major role in initiating, shaping and controlling movements. The most common reason to permit entry is the need for workers – with states sometimes taking on the role of labour recruiter on behalf of employers –

but demographic or humanitarian considerations may also be important. Immigration as part of nation building has played a major role in new world countries such as the USA, Canada, Argentina, Brazil and Australia. Policies on refugees and asylum seekers are major determinants of contemporary population movements.

Thus the idea of individual migrants who make free choices which not only 'maximize their well-being' but also lead to an 'equilibrium in the marketplace' (Borjas, 1989: 482) is so far from historical reality that it has little explanatory value. It seems better, as Zolberg suggests, to analyse labour migration 'as a movement of workers propelled by the dynamics of the transnational capitalist economy, which simultaneously determines both the "push" and the "pull" ' (Zolberg, Suhrke and Aguao, 1989: 407). This implies that migrations are collective phenomena, which should be examined as subsystems of an increasingly global economic and political system.

The historical-structural approach

An alternative explanation of international migration was provided from the 1970s by what came to be called the *historical-structural approach*. This had its intellectual roots in Marxist political economy and in world systems theory. This approach stressed the unequal distribution of economic and political power in the world economy. Migration was seen mainly as a way of mobilizing cheap labour for capital. It perpetuated uneven development, exploiting the resources of poor countries to make the rich even richer (Castles and Kosack, 1985; Cohen, 1987; Sassen, 1988). While the 'push–pull' theories tended to focus on mainly voluntary migrations of individuals, like that from Europe to the USA before 1914, historical-structural accounts looked at mass recruitment of labour by capital, whether for the factories of Germany, for the agribusiness of California or for infrastructure projects like Australia's Snowy Mountain Hydroelectric Scheme. The availability of labour was both a legacy of colonialism and the result of war and regional inequalities within Europe. For world systems theories, labour migration was one of the main ways in which links of domination were forged between the core economies of capitalism and its underdeveloped periphery. Migration was as important as military hegemony and control of world trade and investment in keeping the Third World dependent on the First.

But the historical-structural approach was in turn criticized by many migration scholars: if the logic of capital and the interests of Western states were so dominant, how could the frequent breakdown of migration policies be explained, such as the unplanned shift from labour migration to permanent settlement in certain countries? Both the neo-classical

perspective and the historical-structural approach seemed too one-sided to analyse adequately the great complexity of contemporary migrations. The neo-classical approach neglected historical causes of movements, and downplayed the role of the state, while the historical-functional approach often saw the interests of capital as all-determining, and paid inadequate attention to the motivations and actions of the individuals and groups involved.

Migration systems theory and the trend to a new interdisciplinary approach

Out of such critiques emerged a new approach, *migration systems theory*, which attempts to include a wide range of disciplines, and to cover all dimensions of the migration experience. A migration system is constituted by two or more countries which exchange migrants with each other. The tendency is to analyse regional migration systems, such as the South Pacific, West Africa or the Southern Cone of Latin America (Kritz *et al.*, 1992). However, distant regions may be interlinked, such as the migration system embracing the Caribbean, Western Europe and North America; or that linking North and West Africa with France. The migration systems approach means examining both ends of the flow and studying all the linkages between the places concerned. These linkages can be categorized as 'state-to-state relations and comparisons, mass culture connections and family and social networks' (Fawcett and Arnold, 1987: 456–7).

Migration systems theory suggests that migratory movements generally arise from the existence of prior links between sending and receiving countries based on colonization, political influence, trade, investment or cultural ties. Thus migration from Mexico to the USA originated in the southwestward expansion of the USA in the nineteenth century and the deliberate recruitment of Mexican workers by US employers in the twentieth century (Portes and Rumbaut, 1996: 272–6). The migration from the Dominican Republic to the USA was initiated by the US military occupation of the 1960s. Similarly, both the Korean and the Vietnamese migrations to America were the long-term consequence of US military involvement (Sassen, 1988: 6–9). The migrations from India, Pakistan and Bangladesh to Britain are linked to the British colonial presence on the Indian subcontinent. Similarly, Caribbean migrants have tended to move to their respective former colonial power: for example, from Jamaica to Britain, Martinique to France and Surinam to the Netherlands. The Algerian migration to France (and not to Germany) is explained by the French colonial presence in Algeria, while the Turkish presence in Germany is the result of direct labour recruitment by Germany in the 1960s and early 1970s.

The migration systems approach is part of a trend towards a more inclusive and interdisciplinary understanding, which is emerging as a new mainstream of migration theory – at least outside the domain of neo-classical orthodoxy. The basic principle is that any migratory movement can be seen as the result of interacting macro- and micro-structures. Macro-structures refer to large-scale institutional factors, while micro-structures embrace the networks, practices and beliefs of the migrants themselves. These two levels are linked by a number of intermediate mechanisms, which are often referred to as 'meso-structures'.

The macro-structures include the political economy of the world market, interstate relationships, and the laws, structures and practices established by the states of sending and receiving countries to control migration settlement. The evolution of production, distribution and exchange within an increasingly integrated world economy over the last five centuries has clearly been a major determinant of migrations. The role of international relations and of the states of both sending and receiving areas in organizing or facilitating movements is also significant (Dohse, 1981; Böhning, 1984; Cohen, 1987; Mitchell, 1989; Hollifield, 2000).

The micro-structures are the informal social networks developed by the migrants themselves, in order to cope with migration and settlement. Earlier literature used the concept of 'chain migration' in this context (Price, 1963: 108–10). Research on Mexican migrants in the 1970s showed that 90 per cent of those surveyed had obtained legal residence in the USA through family and employer connections (Portes and Bach, 1985). Today many authors emphasize the role of information and 'cultural capital' (knowledge of other countries, capabilities for organizing travel, finding work and adapting to a new environment) in starting and sustaining migratory movements. Informal networks include personal relationships, family and household patterns, friendship and community ties, and mutual help in economic and social matters. Such links provide vital resources for individuals and groups, and may be referred to as 'social capital' (Bourdieu and Wacquant, 1992: 119). Informal networks bind 'migrants and non-migrants together in a complex web of social roles and interpersonal relationships' (Boyd, 1989: 639).

The family and community are crucial in migration networks. Research on Asian migration has shown that migration decisions are usually made not by individuals but by families (Hugo, 1994). In situations of rapid change, a family may decide to send one or more members to work in another region or country, in order to maximize income and survival chances. In many cases, migration decisions are made by the elders (especially the men), and younger people and women are expected to obey patriarchal authority. The family may decide to send young women to the city or overseas, because the labour of the young men is less dispensable on the farm. Young women are also often seen as more reliable in sending

remittances. Such motivations correspond with increasing international demand for female labour as factory workers for precision assembly or as domestic servants, contributing to a growing feminization of migration.

Family linkages often provide both the financial and the cultural capital which make migration possible. Typically migratory chains are started by an external factor, such as recruitment or military service, or by an initial movement of young (usually male) pioneers. Once a movement is established, the migrants mainly follow 'beaten paths' (Stahl, 1993), and are helped by relatives and friends already in the area of immigration. Networks based on family or on common origin help provide shelter, work, assistance in coping with bureaucratic procedures and support in personal difficulties. These social networks make the migratory process safer and more manageable for the migrants and their families. Migratory movements, once started, become self-sustaining social processes.

Migration networks also provide the basis for processes of settlement and community formation in the immigration area. Migrant groups develop their own social and economic infrastructure: places of worship, associations, shops, cafés, professionals such as lawyers and doctors, and other services. This is linked to family reunion: as length of stay increases, the original migrants (whether workers or refugees) begin to bring in their spouses and children, or found new families. People start to see their life perspectives in the new country. This process is especially linked to the situation of migrants' children: once they go to school in the new country, learn the language, form peer group relationships and develop bicultural or transcultural identities, it becomes more and more difficult for the parents to return to their homelands.

The intermediate 'meso-structures' have been attracting increasing attention from researchers in recent years. Certain individuals, groups or institutions may take on the role of mediating between migrants and political or economic institutions. A 'migration industry' emerges, consisting of recruitment organizations, lawyers, agents, smugglers and other intermediaries (Harris, 1996: 132–6). Such people can be both helpers, and exploiters of migrants. Especially in situations of illegal migration or of oversupply of potential migrants, the exploitative role may predominate: many migrants have been swindled out of their savings and have found themselves marooned without work or resources in a strange country. The emergence of a migration industry with a strong interest in the continuation of migration has often confounded government efforts to control or stop movements.

Macro-, meso- and micro-structures are intertwined in the migratory process, and there are no clear dividing lines between them. No single cause is ever sufficient to explain why people decide to leave their country and settle in another. It is essential to try to understand all aspects of the migratory process, by asking questions such as the following:

1. What economic, social, demographic, environmental or political factors have changed so much that people feel a need to leave their area of origin?
2. What factors provide opportunities for migrants in the destination area?
3. How do social networks and other links develop between the two areas, providing prospective migrants with information, means of travel and the possibility of entry?
4. What legal, political, economic and social structures and practices exist or emerge to regulate migration and settlement?
5. How do migrants turn into settlers, and why does this lead to discrimination, conflict and racism in some cases, but to pluralist or multicultural societies in others?
6. What is the effect of settlement on the social structure, culture and national identity of the receiving societies?
7. How does emigration change the sending area?
8. To what extent do migrations lead to new linkages between sending and receiving societies?

Transnational theory

This last aspect – new linkages between societies based on migration – has attracted much attention in recent years, leading to the emergence of a new body of theory on 'transnationalism' and 'transnational communities'. One aspect of globalization is rapid improvement in technologies of transport and communication, making it increasingly easy for migrants to maintain close links with their areas of origin. These developments also facilitate the growth of circulatory or repeated mobility, in which people migrate regularly between a number of places where they have economic, social or cultural linkages. Debates on transnationalism were stimulated by the work of Basch *et al.* (1994), which argued that 'deterritorialized nation-states' were emerging, with potentially serious consequences for national identity and international politics. Portes defines transnational activities as

> those that take place on a recurrent basis across national borders and that require a regular and significant commitment of time by participants. Such activities may be conducted by relatively powerful actors, such as representatives of national governments and multinational corporations, or may be initiated by more modest individuals, such as immigrants and their home country kin and relations. These activities are not limited to economic enterprises, but include political, cultural and religious initiatives as well. (Portes, 1999: 464)

The notion of a transnational community puts the emphasis on human agency. In the context of globalization, transnationalism can extend

previous face-to-face communities based on kinship, neighbourhoods or workplaces into far-flung virtual communities, which communicate at a distance. Portes and his collaborators emphasize the significance of transnational business communities (whether of large-scale enterprises or of small ethnic entrepreneurs), but also note the importance of political and cultural communities. They distinguish between *transnationalism from above* – activities 'conducted by powerful institutional actors, such as multinational corporations and states' – and *transnationalism from below* – activities 'that are the result of grass-roots initiatives by immigrants and their home country counterparts' (Portes *et al.*, 1999: 221). Transnational communities can develop countervailing power to contest the power of corporations, governments and intergovernmental organizations. Indeed, informal linkages in the form of migration networks often undermine official migration policies which ignore the interests of migrants.

The term *transmigrant* may be used to identify people whose existence is shaped through participation in transnational communities based on migration (Glick-Schiller 1999: 203). Inflationary use of the term should be avoided: the majority of migrants still do not fit the pattern. Temporary labour migrants who sojourn abroad for a few years, send back remittances, communicate with their family at home and visit them occasionally are not transmigrants. Nor are permanent migrants who leave forever, and simply retain loose contact with their homeland. The key defining feature is that transnational activities are a central part of a person's life. Where this applies to a group of people, one can speak of a transnational community.

Transnational communities are not new, although the term is. The diaspora concept goes back to ancient times, and was used for peoples displaced or dispersed by force (e.g. the Jews; African slaves in the New World). It was also applied to certain trading groups such as Greeks in Western Asia and Africa, or the Arab traders who brought Islam to South-East Asia, as well as to labour migrants (Indians in the British Empire; Italians since the 1860s) (Cohen, 1997; Van Hear, 1998). The term diaspora often has strong emotional connotations, while the notion of a transnational community is more neutral. The new factor is the rapid proliferation of transnational communities under conditions of globalization (Vertovec, 1999: 447). Transnationalism is likely to go on growing, and transnational communities will become an increasingly important way to organize activities, relationships and identity for the growing number of people with affiliations in two or more countries.

From migration to settlement

Although each migratory movement has its specific historical patterns, it is possible to generalize on the social dynamics of the migratory process. It

is necessary, however, to differentiate between economically-motivated migration and forced migration. Most economic migrations start with young, economically-active people. They are often 'target-earners', who want to save enough in a higher-wage economy to improve conditions at home, by buying land, building a house, setting up a business, or paying for education or dowries. After a period in the receiving country, some of these 'primary migrants' return home, but others prolong their stay, or return and then remigrate. This may be because of relative success when migrants find living and working conditions in the new country better than in the homeland. But it may also be because of relative failure when migrants find it impossible to save enough to achieve their aims, necessitating a longer sojourn. As time goes on, many erstwhile temporary migrants send for spouses, or find partners in the new country. With the birth of children, settlement takes on a more permanent character.

It is this powerful internal dynamic of the migratory process that often confounds expectations of the participants and undermines the objectives of policy-makers in both sending and receiving countries. In many migrations, there is no initial intention of family reunion and permanent settlement. However, when governments try to stop flows – for instance, because of a decline in the demand for labour – they may find that the movement has become self-sustaining. What started off as a temporary labour flow is transformed into family reunion, undocumented migration or even asylum-seeker flows. This is a result of the maturing of the migratory movement and of the migrants themselves as they pass through the life cycle. It may also be because dependency on migrant workers in certain sectors has become a structural feature of the economy.

The failure of policy-makers and analysts to see international migration as a dynamic social process is at the root of many political and social problems. The source of this failure has often been a one-sided focus on economic models of migration, which mistakenly claim that migration is an individual response to market factors. This has led to the belief that migration can be turned on and off like a tap, by changing policy settings which influence the costs and benefits of mobility for migrants. Migration may continue due to social factors, even when the economic factors which initiated the movement have been completely transformed.

Such developments are well illustrated by the Western European experience of 'guestworker' type movements from the Mediterranean basin from 1945 to 1973. Other situations in which social factors have led to unexpected outcomes include migrations from former colonies to the UK, France and the Netherlands, and migration from Europe, Latin America and Asia to the USA, Australia and Canada (see Chapter 4). One lesson of the last half-century is that it is extremely difficult for countries with democratic rights and strong legal systems to prevent migration turning into settlement. The situation is somewhat different in labour-recruiting

countries which lack effective human rights guarantees, such as the Gulf states or some East and South-East Asian countries. The social dynamics of the migratory process do exist, but restrictions by the receiving governments may hinder family reunion and permanent settlement (Chapters 6 and 7).

The dynamics are different in the case of refugees and asylum seekers. They leave their countries because persecution, human rights abuse and generalized violence makes life there unsustainable. Most forced migrants remain in the neighbouring countries of first asylum – which are usually poor and often politically unstable themselves. Onward migration to countries which offer better economic and social opportunities is only possible for a small minority. However, there is evidence of selectivity: it is mainly those with financial resources, human capital (especially education) and social networks in destination countries who are able to migrate onwards (Zolberg and Benda, 2001). This onward migration is motivated both by the imperative of leaving a country of origin where life has become perilous, and by the hope of building a better life elsewhere. Attempts by policy-makers to make clear distinctions between economic and forced migrants are hampered by these 'mixed motivations'.

This has led to the notion of the 'migration-asylum nexus', which points to the complex links between the varying reasons for migration. Labour migrants, permanent settlers and refugees move under different conditions and legal regimes. Yet all these population movements are symptomatic of modernization and globalization. Colonialism, industrialization and integration into the world economy destroy traditional forms of production and social relations, and lead to the reshaping of nations and states. Underdevelopment, impoverishment, poor governance, endemic conflict and human rights abuse are closely linked. These conditions lead both to economically-motivated migration and to politically-motivated flight.

The formation of ethnic minorities

The long-term effects of immigration on society emerge in the later stages of the migratory process when migrants settle permanently and form distinct groups. Outcomes can be very different, depending on the actions of the receiving state and society. At one extreme, openness to settlement, granting of citizenship and gradual acceptance of cultural diversity may allow the formation of *ethnic communities*, which are seen as part of a multicultural society. At the other extreme, denial of the reality of settlement, refusal of citizenship and rights to settlers, and rejection of cultural diversity may lead to formation of *ethnic minorities*, whose presence is widely regarded as undesirable and divisive. Most countries of immigration have tended to lie somewhere between these two extremes.

Critics of immigration portray ethnic minorities as a threat to economic well-being, public order and national identity. Yet these ethnic minorities may in fact be the creation of the very people who fear them. Ethnic minorities may be defined as groups which

(a) have been assigned a subordinate position in society by dominant groups on the basis of socially-constructed markers of phenotype (that is, physical appearance or 'race'), origins or culture;
(b) have some degree of collective consciousness (or feeling of being a community) based on a belief in shared language, traditions, religion, history and experiences.

An ethnic minority is therefore a product of both 'other-definition' and of 'self-definition'. *Other-definition* means ascription of undesirable characteristics and assignment to inferior social positions by dominant groups. *Self-definition* refers to the consciousness of group members of belonging together on the basis of shared cultural and social characteristics. The relative strength of these processes varies. Some minorities are mainly constructed through processes of exclusion (which may be referred to as *racism*) by the majority. Others are mainly constituted on the basis of cultural and historical consciousness (or *ethnic identity*) among their members. The concept of the ethnic minority always implies some degree of marginalization or exclusion, leading to situations of actual or potential conflict. Ethnicity is rarely a theme of political significance when it is simply a matter of different group cultural practices.

Ethnicity

In popular usage, ethnicity is usually seen as an attribute of minority groups, but most social scientists argue that everybody has ethnicity, defined as a sense of group belonging, based on ideas of common origins, history, culture, experience and values (see Fishman, 1985: 4; Smith, 1986: 27). These ideas change only slowly, which gives ethnicity durability over generations and even centuries. But that does not mean that ethnic consciousness and culture within a group are homogeneous and static. Cohen and Bains argue that ethnicity, unlike race 'refers to a real process of historical individuation – namely the linguistic and cultural practices through which a sense of collective identity or "roots" is produced and transmitted from generation to generation, *and is changed in the process*' (Cohen and Bains, 1988: 24–5, emphasis in original).

The origins of ethnicity may be explained in various ways. Geertz, for example, sees ethnicity as a 'primordial attachment', which results 'from being born into a particular religious community, speaking a particular language, or even a dialect of a language and following particular social

practices. These congruities of blood, speech, custom and so on, are seen to have an ineffable, and at times, overpowering coerciveness in and of themselves' (Geertz, 1963, quoted from Rex, 1986: 26–7). In this approach, ethnicity is not a matter of choice; it is pre-social, almost instinctual, something one is born into.

By contrast, many anthropologists use a concept of 'situational' ethnicity. Members of a specific group decide to 'invoke' ethnicity, as a criterion for self-identification, in a situation where such identification is necessary or useful. This explains the variability of ethnic boundaries and changes in salience at different times. The markers chosen for the boundaries are also variable, generally emphasizing cultural characteristics, such as language, shared history, customs and religion, but sometimes including physical characteristics (Wallman, 1986: 229). In this view there is no essential difference between the drawing of boundaries on the basis of cultural difference or of phenotypical difference (popularly referred to as 'race'). The visible markers of a phenotype (skin colour, features, hair colour, and so on) correspond to what is popularly understood as 'race'. We avoid using the term 'race' as far as possible, since there is increasing agreement among biologists and social scientists that there are no measurable characteristics among human populations that allow classification into 'races'. Genetic variance within any one population is greater than alleged differences between different populations. 'Race' is thus a social construction produced by the process we refer to as racism.

Similarly, some sociologists see ethnic identification or mobilization as rational behaviour, designed to maximize the power of a group in a situation of market competition. Such theories have their roots in Max Weber's concept of 'social closure', whereby a status group establishes rules and practices to exclude others, in order to gain a competitive advantage (Weber, 1968: 342). For Weber (as for Marx), organization according to 'affective criteria' (such as religion, ethnic identification or communal consciousness) was in the long run likely to be superseded by organization according to economic interests (class) or bureaucratic rationality. Nonetheless, the instrumental use of these affiliations could be rational if it led to successful mobilization.

Other sociologists reject the concept of ethnicity altogether, seeing it as 'myth' or 'nostalgia', which cannot survive against the rational forces of economic and social integration in large-scale industrial societies (Steinberg, 1981). Yet it is hard to ignore the growing significance of ethnic mobilization, so that many attempts have been made to show the links between ethnicity and power. Studies of the 'ethnic revival' by the US sociologists Glazer and Moynihan (1975) and Bell (1975) emphasize the instrumental role of ethnic identification: phenotypical and cultural characteristics are used to strengthen group solidarity, in order to struggle more effectively for market advantages, or for increased allocation of

resources by the state. Bell sees ethnic mobilization as a substitute for the declining power of class identification in advanced industrial societies; the decision to organize on ethnic lines seems to be an almost arbitrary 'strategic choice'. This does not imply that markers, such as skin colour, language, religion, shared history and customs, are not real, but rather that the decision to use them to define an ethnic group is not predetermined.

Whether ethnicity is 'primordial', 'situational' or 'instrumental' need not concern us further here. The point is that ethnicity leads to identification with a specific group, but its visible markers – phenotype, language, culture, customs, religion, behaviour – may also be used as criteria for exclusion by other groups. Ethnicity only takes on social and political meaning when it is linked to processes of boundary drawing between dominant groups and minorities. Becoming an ethnic minority is not an automatic result of immigration, but rather the consequence of specific mechanisms of marginalization, which affect different groups in different ways.

Racism

Racism towards certain groups is to be found in virtually all immigration countries. Racism may be defined as the process whereby social groups categorize other groups as different or inferior, on the basis of phenotypical or cultural markers. This process involves the use of economic, social or political power, and generally has the purpose of legitimating exploitation or exclusion of the group so defined.

Racism means making (and acting upon) predictions about people's character, abilities or behaviour on the basis of socially constructed markers of difference. The power of the dominant group is sustained by developing structures (such as laws, policies and administrative practices) that exclude or discriminate against the dominated group. This aspect of racism is generally known as institutional or structural racism. Racist attitudes and discriminatory behaviour on the part of members of the dominant group are referred to as informal racism. Many social scientists now use the term 'racialization' to refer to public discourses which imply that a range of social or political problems are a 'natural' consequence of certain ascribed physical or cultural characteristics of minority groups. Racialization can be used to apply to the social construction of a specific group as a problem, or in the wider sense of the 'racialization of politics' or the 'racialization of urban space'.

In some countries, notably Germany and France, there is reluctance to speak of racism. Euphemisms such as 'hostility to foreigners', 'ethnocentrism' or 'xenophobia' are used. But the debate over the label seems sterile: it is more important to understand the phenomenon and its causes. Racism operates in different ways according to the specific history of a society and

the interests of the dominant group. In many cases, supposed biological differences are not the only markers: culture, religion, language or other factors are taken as indicative of phenotypical differences. For instance, anti-Muslim racism in Europe is based on cultural symbols which, however, are linked to phenotypical markers (such as Arab or African features).

The historical explanation for racism in Western Europe and in post-colonial settler societies (like Australia) lies in traditions, ideologies and cultural practices, which have developed through ethnic conflicts associated with nation building and colonial expansion (compare Miles, 1989). The reasons for the recent increase in racism lie in fundamental economic and social changes which question the optimistic view of progress embodied in Western thought. Since the early 1970s, economic restructuring and increasing international cultural interchange have been experienced by many sections of the population as a direct threat to their livelihood, social conditions and identity. Since these changes have coincided with the arrival of new ethnic minorities, the tendency has been to perceive the newcomers as the cause of the threatening changes: an interpretation eagerly encouraged by the extreme right, but also by many mainstream politicians.

Moreover, the very changes which threaten disadvantaged sections of the population have also weakened the labour movement and working-class cultures, which might otherwise have provided some measure of protection. The decline of working-class parties and trade unions and the erosion of local communicative networks have created the social space for racism to become more virulent (Wieviorka, 1995; Vasta and Castles, 1996). (We lay no claim to originality with regard to this definition and discussion of racism. It is oriented towards current sociological debates, which have generated a large body of literature. See, for example, CCCS (1982); Rex and Mason (1986); Cohen and Bains (1988); Miles (1989); Wieviorka (1991, 1992); Solomos (1993); Goldberg and Solomos (2002). There is no unanimity among social scientists about the correct definition and explanations of racism, but we have no space for a more detailed discussion of these matters here.)

Ethnicity, class, gender and life cycle

Racial and ethnic divisions are only one aspect of social differentiation. Others include social class, gender and position in the life cycle. None of these distinctions is reducible to any other, yet they constantly cross-cut and interact, affecting life chances, lifestyles, culture and social consciousness. Immigrant groups and ethnic minorities are just as heterogeneous as the rest of the population. The migrant is a gendered subject, embedded in a wide range of social relationships.

In the early stages of post-1945 international labour mobility, the vital nexus appeared to be that between migration and class. Migration was analysed in terms of the interests of various sectors of labour and capital (Castles and Kosack, 1985) or of the incorporation of different types of workers into segmented labour markets (Piore, 1979). International migration continues to be an important factor helping to shape labour market patterns and class relations (see Chapter 8). However, there has been a growing awareness of the crucial links between class, ethnicity and gender.

Even in the early stages, the role of women in maintaining families and reproducing workers in the country of origin was crucial to the economic benefits of labour migration. Moreover, a large proportion of migrant workers were female. As Phizacklea (1983: 5) pointed out, it was particularly easy to ascribe inferiority to women migrant workers, just because their primary roles in patriarchal societies were defined as wife and mother, dependent on a male breadwinner. They could therefore be paid lower wages and controlled more easily than men. Since the 1970s, restructuring and unemployment have made full employment more the exception than the rule for some minorities. Very high rates of unemployment among ethnic minority youth have meant that 'they are not the unemployed, but the never employed' (Sivanandan, 1982: 49). Members of ethnic minorities have experienced racism from some white workers and therefore find it hard to define their political consciousness in class terms.

Migrant women's work experience often remains distinct from that of men. They tend to be overrepresented in the least desirable occupations, such as repetitive factory work and lower-skilled positions in the personal and community services sectors. However, there has been some mobility into white-collar jobs in recent years, partly as a result of the decline of manufacturing. Professional employment is often linked to traditional caring roles. Minority women have experienced casualization of employment and increasing unemployment (which often does not appear in the statistics due to their status as 'dependants').

Complex patterns of division of labour on ethnic and gender lines have developed (Waldinger *et al.*, 1990). In a study of the fashion industry in European countries, Phizacklea (1990: 72–93) argued that this industry was able to survive, despite the new global division of labour, through the development of 'subcontracting webs': large retail companies were able to put pressure for lower prices on small firms controlled by male ethnic entrepreneurs, whose market position was constrained by racial discrimination. These in turn were able to use both patriarchal power relations and the vulnerable legal position of women immigrants to enforce extremely low wages and poor working conditions in sweatshops and outwork. Collins *et al.* (1995: 180–1) present a similar picture of the links between racialization and gender in ethnic small business in Australia.

Racism, sexism and class domination are three specific forms of 'social normalization and exclusion' which are intrinsic to capitalism and modernity, and which have developed in close relationship to each other (Balibar, 1991: 49). Racism and sexism both involve predicting social behaviour on the basis of allegedly fixed biological or cultural characteristics. According to Essed, racism and sexism 'narrowly intertwine and combine under certain conditions into one, hybrid phenomenon. Therefore it is useful to speak of *gendered racism* to refer to the racist oppression of Black women as structured by racist and ethnicist perceptions of gender roles' (Essed, 1991: 31, emphasis in original).

Anthias and Yuval-Davis (1989) analyse links between gender relations and the construction of the nation and the ethnic community. Women are not only the biological reproducers of an ethnic group, but also the 'cultural carriers' who have the key role in passing on the language and cultural symbols to the young (see also Vasta, 1990, 1992). In nationalist discourses women serve as the symbolic embodiment of national unity and distinctiveness. They nurture and support the (male) warrior-citizens. In defeat and suffering, the nation is portrayed as a woman in danger. Such symbolism legitimates the political inferiority of women: they embody the nation, while the men represent it politically and militarily (Lutz *et al.*, 1995).

The role of gender in ethnic closure is evident in immigration rules which still often treat men as the principal immigrants while women and children are mere 'dependants'. Britain has used gender-specific measures to limit the growth of the black population. In the 1970s, women from the Indian subcontinent coming to join husbands or fiancés were subjected to 'virginity tests' at Heathrow Airport. The authorities also sought to prevent Afro-Caribbean and Asian women from bringing in husbands, on the grounds that the 'natural place of residence' of the family was the abode of the husband (Klug, 1989: 27–9). In many countries, women who enter as dependants do not have an entitlement to residence in their own right and may face deportation if they get divorced.

The stages of the life cycle – childhood, youth, maturity, middle age, old age – are also important determinants of economic and social positions, culture and consciousness. There is often a gulf between the experiences of the migrant generation and those of their children, who have grown up and gone to school in the new country. Ethnic minority youth become aware of the contradiction between the prevailing ideologies of equal opportunity and the reality of discrimination and racism in their daily lives. This can lead to the emergence of counter-cultures and political radicalization. In turn, ethnic minority youth are perceived as a 'social time bomb' or a threat to public order, which has to be contained through social control institutions such as the police, schools and welfare bureaucracies (see Chapter 10).

Culture, identity and community

In the context of globalization, culture, identity and community often serve as a focus of resistance to centralizing and homogenizing forces (Castells, 1997). These have become central themes in debates on the new ethnic minorities. First, as already outlined, cultural difference serves as a marker for ethnic boundaries. Second, ethnic cultures play a central role in community formation: when ethnic groups cluster together, they establish their own neighbourhoods, marked by distinctive use of private and public spaces. Third, ethnic neighbourhoods are perceived by some members of the majority group as confirmation of their fears of a 'foreign takeover'. Ethnic communities are seen as a threat to the dominant culture and national identity. Fourth, dominant groups may see migrant cultures as primordial, static and regressive. Linguistic and cultural maintenance is taken as proof of inability to come to terms with an advanced industrial society. Those who do not assimilate 'have only themselves to blame' for their marginalized position.

For ethnic minorities, culture plays a key role as a source of identity and as a focus for resistance to exclusion and discrimination. Reference to the culture of origin helps people maintain self-esteem in a situation where their capabilities and experience are undermined. But a static, primordial culture cannot fulfil this task, for it does not provide orientation in a hostile environment. The dynamic nature of culture lies in its capacity to link a group's history and traditions with the actual situation in the migratory process. Migrant or minority cultures are constantly recreated on the basis of the needs and experience of the group and its interaction with the actual social environment (Schierup and Ålund, 1987; Vasta *et al.*, 1992). An apparent regression, for instance to religious fundamentalism, may be precisely the result of a form of modernization which has been experienced as discriminatory, exploitative and destructive of identity.

It is therefore necessary to understand the development of ethnic cultures, the stabilization of personal and group identities, and the formation of ethnic communities as facets of a single process. This process is not self-contained: it depends on constant interaction with the state and the various institutions and groups in the country of immigration, as well as with the society of the country of origin. Immigrants and their descendants do not have a static, closed and homogeneous ethnic identity, but instead dynamic *multiple identities*, influenced by a variety of cultural, social and other factors.

The concept of *national culture and identity* has become highly questionable. Increasing global economic and cultural integration is leading to a simultaneous homogenization and fragmentation of culture. As multinational companies take over and repackage the artefacts of local cultures it becomes possible to consume all types of cultural products

everywhere, but at the same time these lose their meaning as symbols of group identity. National or ethnic cultures shed their distinctiveness and become just another celebration of the cultural dominance of the international industrial apparatus. Hence the constant search for new sub-cultures, styles and sources of identity, particularly on the part of youth.

Gilroy sees the focus of this recreation of culture in the social movements of local communities, as well as in youth sub-cultures. He argues that legacies of anti-colonial struggles have been reshaped in Britain in the reproduction of classes and 'races' which become youth culture:

> The institutions they create: temples, churches, clubs, cafes and blues dances confound any Eurocentric idea of where the line dividing politics and culture should fall. The distinction between public and private spheres cuts across the life of their households and communities in a similar manner. Traditional solidarity mediates and adapts the institutions of the British political system against which it is defined. (Gilroy, 1987: 37)

Culture is becoming increasingly politicized in all countries of immigration. As ideas of racial superiority lose their ideological strength, exclusionary practices against minorities increasingly focus on issues of cultural difference. At the same time, the politics of minority resistance crystallize more and more around cultural symbols. Yet these symbols are only partially based on imported forms of ethnicity. Their main power as definers of community and identity comes from the incorporation of new experiences of ethnic minority groups in the immigration country.

State and nation

Large-scale migrations and growing diversity may have important effects on political institutions and national identity. In the contemporary world, nation-states (of which there are some 200) are the predominant form of political organization. They derive their legitimacy from the claim of representing the aspirations of their people (or citizens). This implies two further claims: that there is an underlying cultural consensus which allows agreement on the values or interests of the people, and that there is a democratic process for the will of the citizens to be expressed. Such claims are often empty slogans, for most countries are marked by heterogeneity, based on ethnicity, class and other cleavages. Only a minority of countries consistently use democratic mechanisms to resolve value and interest conflicts. Nonetheless, the democratic nation-state has become a global norm.

Immigration of culturally diverse people presents nation-states with a dilemma: incorporation of the newcomers as citizens may undermine myths of cultural homogeneity; but failure to incorporate them may lead

to divided societies, marked by severe inequality and conflict. This problem arises from the character of the nation-state, as it developed in Western Europe and North America in the context of modernization, industrialization and colonialism. Pre-modern states based their authority on the absolute power of a monarch over a specific territory. Within this area, all people were subjects of the monarch (rather than citizens). There was no concept of a national culture which transcended the gulf between aristocratic rulers and peasants. The modern nation-state, by contrast, implies a close link between cultural belonging and political identity (Castles and Davidson, 2000).

A *state*, according to Seton-Watson (1977: 1), 'is a legal and political organization, with the power to require obedience and loyalty from its citizens'. The state regulates political, economic and social relations in a bounded territory. Most modern nation-states are formally defined by a constitution and laws, according to which all power derives from the people (or nation). It is therefore vital to define who belongs to the people. Membership is marked by the status of citizenship, which lays down rights and duties. Non-citizens are excluded from at least some of these. Citizenship is the essential link between state and nation, and obtaining citizenship is of central importance for newcomers to a country.

Seton-Watson describes a *nation* as 'a community of people, whose members are bound together by a sense of solidarity, a common culture, a national consciousness' (Seton-Watson, 1977: 1). Such essentially subjective phenomena are difficult to measure. Moreover, it is not clear how a nation differs from an ethnic group, which is defined in a very similar way (see above). Anderson provides an answer with his definition of the nation: 'it is an imagined political community – and imagined as both inherently limited and sovereign' (Anderson, 1983: 15). This concept points to the political character of the nation and its links with a specific territory: an ethnic group that attains sovereignty over a bounded territory becomes a nation and establishes a nation-state. As Smith (1991: 14) puts it: 'A nation can ... be defined as a named human population sharing an historic territory, common myths and historical memories, a mass, public culture, a common economy and common legal rights and duties for all members.'

Anderson (1983) regards the nation-state as a modern phenomenon, whose birthdate is that of the US Constitution of 1787. Gellner (1983) argues that nations could not exist in pre-modern societies, owing to the cultural gap between elites and peasants, while modern industrial societies require cultural homogeneity to function, and therefore generate the ideologies needed to create nations. However, both Seton-Watson (1977) and Smith (1986) argue that the nation is of much greater antiquity, going back to the ancient civilizations of East Asia, the Middle East and Europe. All these authors seem to agree that the nation is essentially a belief system, based on collective cultural ties and sentiments. These

convey a sense of identity and belonging, which may be referred to as national consciousness.

Specific to the modern nation-state is the linking of national consciousness with the principle of democracy: every person classified as a member of the national community has an equal right to participate in the formulation of the political will. This linking of nationality and citizenship is deeply contradictory. In liberal theory, all citizens are meant to be free and equal persons who are treated as homogeneous within the political sphere. This requires a separation between a person's political rights and obligations, and their membership of specific groups, based on ethnicity, religion, social class or regional location. The political sphere is one of universalism, which means abstraction from cultural particularity and difference. Difference is to be restricted to the 'non-public identity' (Rawls, 1985: 232–41).

This conflicts with the reality of nation-state formation, however, in which being a citizen depends on membership in a certain national community, usually based on the dominant ethnic group of the territory concerned. Thus a citizen is always also a member of a nation, a national. Nationalist ideologies demand that ethnic group, nation and state should be facets of the same community and have the same boundaries – every ethnic group should constitute itself as a nation and should have its own state, with all the appropriate trappings: flag, army, Olympic team and postage stamps. In fact, such congruence has rarely been achieved: nationalism has always been an ideology trying to achieve such a condition, rather than an actual state of affairs.

The construction of nation-states has involved the spatial extension of state power, and the territorial incorporation of hitherto distinct ethnic groups. These may or may not coalesce into a single nation over time. Attempts to consolidate the nation-state can mean exclusion, assimilation or even genocide for minority groups. It is possible to keep relatively small groups in situations of permanent subjugation and exclusion from the 'imagined community'. This has applied, for instance, to Jews and gypsies in various European countries, to indigenous peoples in settler colonies and to the descendants of slaves and contract workers in some areas of European colonization. Political domination and cultural exclusion is much more difficult if the subjugated nation retains a territorial base, like the Scots, Welsh and Irish in the UK, or the Basques in Spain.

The experience of 'historical minorities' has helped to mould structures and attitudes, which affect the conditions for new immigrant groups. The pervasive fear of 'ghettos' or 'ethnic enclaves' indicates that minorities seem most threatening when they concentrate in distinct areas. For nationalists, an ethnic group is a potential nation which does not (yet) control any territory, or have its own state. Most modern states have made conscious efforts to achieve cultural and political integration of minorities.

Mechanisms include citizenship itself, centralized political institutions, the propagation of national languages, universal education systems and creation of national institutions like the army, or an established church (Schnapper, 1991, 1994). The problem is similar in character everywhere, whether the minorities are 'old' or 'new': how can a nation be defined, if not in terms of a shared (and single) ethnic identity? How are core values and behavioural norms to be laid down, if there is a plurality of cultures and traditions?

Coping with diversity has become even more difficult in the era of globalization. In the nation-states of the nineteenth and early twentieth centuries, politics, the economy, social relations and culture were all organized within the same boundaries. Even movements for change, such as the labour movement or left-wing parties, based their strategies on the nation-state. Globalization has destabilized this model. The dynamics of economic life now transcend borders, and have become increasingly uncontrollable for national governments. De-industrialization of the older industrial nations has led to profound social changes. The nation-state is still the basic unit for defence, public order and welfare, but its room for autonomous action is severely reduced. No government can pursue policies which ignore the imperatives of global markets. The nexus between power and national boundaries is declining.

Citizenship

The states of immigration countries have had to devise a range of policies and institutions to respond to the problems of increased ethnic diversity (see Aleinikoff and Klusmeyer, 2000, 2001). These relate to certain central issues: defining who is a citizen, how newcomers can become citizens and what citizenship means. In principle the nation-state only permits a single membership, but immigrants and their descendants have a relationship to more than one state. They may be citizens of two states, or they may be a citizen of one state but live in another. These situations may lead to 'divided loyalties' and undermine the cultural homogeneity which is the nationalist ideal. Thus large-scale settlement inevitably leads to a debate on citizenship.

Citizenship designates the equality of rights of all citizens within a political community, as well as a corresponding set of institutions guaranteeing these rights (Bauböck, 1991: 28). However, formal equality rarely leads to equality in practice. For instance, citizenship has always meant something different for men than for women, because the concept of the citizen has been premised on the male family-father, who represents his woman and children (Anthias and Yuval-Davis, 1989). The citizen has generally been defined in terms of the cultures, values and interests of the majority ethnic group. Finally, the citizen has usually been explicitly or

implicitly conceived in class terms, so that gaining real participatory rights for members of the working class has been one of the central historical tasks of the labour movement. The history of citizenship has therefore been one of conflicts over the real content of the category in terms of civil, political and social rights (Marshall, 1964).

The first concern for immigrants, however, is not the exact content of citizenship, but how they can obtain it, in order to achieve a legal status formally equal to that of other residents. Access has varied considerably in different countries, depending on the prevailing concept of the nation. We can distinguish the following ideal types of citizenship:

1. The *imperial model*: definition of belonging to the nation in terms of being a subject of the same power or ruler. This notion pre-dates the French and American revolutions. It allowed the integration of the various peoples of multi-ethnic empires (the British, the Austro-Hungarian, the Ottoman). This model remained formally in operation in the UK until the Nationality Act of 1981, which created a modern type of citizenship for the first time. It also had some validity for the former Soviet Union. The concept almost always has an ideological character, in that it helps to veil the actual dominance of a particular ethnic group or nationality over the other subject peoples.
2. The *folk or ethnic model*: definition of belonging to the nation in terms of ethnicity (common descent, language and culture), which means exclusion of minorities from citizenship and from the community of the nation. (Germany came close to this model until the introduction of new citizenship rules in 2000.)
3. The *republican model*: definition of the nation as a political community, based on a constitution, laws and citizenship, with the possibility of admitting newcomers to the community, providing they adhere to the political rules and are willing to adopt the national culture. This assimilationist approach dates back to the French and American revolutions. France is the most obvious current example.
4. The *multicultural model*: definition of the nation as a political community, based on a constitution, laws and citizenship, with the possibility of admitting newcomers, who may maintain cultural difference and form ethnic communities providing they adhere to the political rules. This pluralist or multicultural approach became dominant in the 1970s and 1980s in Australia, Canada and Sweden, and was also influential in other Western countries. However, there was a move away from multiculturalism in many places in the 1990s.

All these ideal types have one factor in common: they are premised on citizens who belong to just one nation-state. Migrant settlement is seen as a process of transferring primary loyalty from the state of origin to the new state of residence. This process, which may be long-drawn-out and even span generations, is symbolically marked by naturalization and acquisition of citizenship of the new state. Transnational theory (see above) argues that this no longer applies for growing groups of migrants who form transnational communities and maintain strong cross-border affiliations – possibly over generations. This is seen as a challenge to traditional models of national identity. Thus an additional ideal type of citizenship seems to be emerging:

5. The *transnational model*: social and cultural identities which transcend national boundaries, leading to multiple and differentiated forms of belonging. Transnationalism could have important consequences for democratic institutions and political belonging in future. This corresponds with the fact that, through globalization, a great deal of political and economic power is shifting to transnational corporations and international agencies which are not currently open to democratic control (Castles and Davidson, 2000). The survival of democracy may depend on finding ways of including people with multiple identities in a range of political communities. It also means ensuring citizen participation in new locations of power, whether supra- or sub-national, public or private.

The applicability of these models to specific countries will be discussed in more detail in Chapter 10. In fact, the models are neither universally accepted nor static even within a single country (Bauböck and Rundell, 1998: 1273). Moreover, the distinction between citizens and non-citizens is becoming less clear-cut. Immigrants who have been legally resident in a country for many years can often obtain a special status, tantamount to 'quasi-citizenship'. This may confer such rights as secure residence status; rights to work, seek employment and run a business; entitlements to social security benefits and health services; access to education and training; and limited political rights, such as the rights of association and of assembly. In some countries, long-term foreign residents have voting rights in local elections. Such arrangements create a new legal status, which is more than that of a foreigner, but less than that of a citizen. Hammar (1990: 15–23) has suggested the term *denizen* for people 'who are foreign citizens with a legal and permanent resident status'. This applies to millions of long-term foreign residents in Western Europe, many of whom were actually born in their countries of residence.

A further element in the emergence of quasi-citizenship is the development of international human rights standards, as laid down by bodies like the UN, the International Labour Organization (ILO) and the World Trade Organization (WTO). A whole range of civil and social rights are legally guaranteed for citizens and non-citizens alike in the states which adopt these international norms (Soysal, 1994). However, the legal protection provided by international conventions can be deficient when

states do not incorporate the norms into their national law, despite ratifying the conventions.

The EU provides the furthest-going example for transnational citizenship. The 1991 Maastricht Treaty established the legal notion of Citizenship of the European Union, which embraced the following individual rights:

- freedom of movement and residence in the territory of member states;
- the right to vote and to stand for office in local elections and European Parliament elections in the state of residence;
- the right to diplomatic protection by diplomats of any EU state in a third country;
- the right to petition the European Parliament and the possibility to appeal to an ombudsman (Martiniello, 1994: 31).

However, EU citizens living in another member state do not have the right to vote in elections for the national parliament of that state. People dependent on social security do not have a right to settle in another member country; and access to public employment is still generally restricted to nationals (Martiniello, 1994: 41). For the time being, it seems more appropriate to treat EU citizenship as a case of quasi-citizenship. The limited character is made even clearer by the fact that an 'EU passport' is legally still a passport of one of the member countries. So far, EU citizenship has done nothing for the majority of immigrants, who come from outside the EU. However, the process of European integration is continuing: the 1997 Treaty of Amsterdam (Article 63) established community competence in the areas of migration and asylum, and principles for a common policy were laid down by the European Council meeting at Tampere in 1999. The new policy – planned to come into force in 2004 – may mean common entry criteria for immigrants and refugees, and freedom of movement within the EU for legally-resident third country nationals.

The long-term question is whether democratic states can successfully operate with a population differentiated into full citizens, quasi-citizens and foreigners. The central principle of the democratic state is that all members of civil society should be incorporated into the political community. That means granting full citizenship to all permanent residents. Migrations are likely to continue and there will be increasing numbers of people with affiliations to more than one society. Dual or multiple citizenship will become increasingly common. In fact, nearly all immigration countries have changed their citizenship rules over the last 40–50 years – sometimes several times. More and more countries accept dual citizenship (at least to some extent). A major focus of reform is the introduction of measures to integrate the second generation into the political community through birthright citizenship or easier naturalization (see Aleinikoff and Klusmeyer, 2000; Castles and Davidson, 2000: Chapter 4). The consequence is that the meaning of citizenship is likely to change, and that the exclusive link to one nation-state will become more tenuous. This could lead to some form of 'transnational citizenship', as Bauböck (1994) suggests. But that in turn raises the question of how states will regulate immigration if citizenship becomes more universal.

Conclusion

This chapter has been concerned with some of the theoretical explanations of migration and ethnic minority formation. One central argument is that migration and settlement are closely related to other economic, political and cultural linkages being formed between different countries in an accelerating process of globalization. International migration – in all its different forms – must be seen as an integral part of contemporary world developments. It is likely to grow in volume in the years ahead, because of the strong pressures for continuing global integration.

A second argument is that the migratory process has certain internal dynamics based on the social networks which are at its core. These internal dynamics can lead to developments not initially intended either by the migrants themselves or by the states concerned. The most common outcome of a migratory movement, whatever its initial character, is settlement of a significant proportion of the migrants, and formation of ethnic communities or minorities in the new country. Thus the emergence of societies which are more ethnically and culturally diverse must be seen as an inevitable result of initial decisions to recruit foreign workers, or to permit immigration.

A third argument is that increasing numbers of international migrants do not simply move from one society to another, but maintain recurring and significant links in two or more places. They form transnational communities which live across borders. This trend is facilitated by globalization, both through improvements in transport and communications technology, and through diffusion of global cultural values. Transnational communities currently embrace only a minority of migrants, but may in the long run have enormous consequences for social identity and political institutions.

The fourth argument concerns the nature of ethnic minorities and the process by which they are formed. Most minorities are formed by a combination of other-definition and self-definition. Other-definition refers to various forms of exclusion and discrimination (or racism). Self-definition has a dual character. It includes assertion and recreation of ethnic identity, centred upon pre-migration cultural symbols and practices. It also includes political mobilization against exclusion and discrimination, using

cultural symbols and practices in an instrumental way. When settlement and ethnic minority formation take place at times of economic and social crisis, they can become highly politicized. Issues of culture, identity and community can take on great significance, not only for immigrants, but also for the receiving society as a whole.

The fifth argument focuses on the significance of immigration for the nation-state. It seems likely that increasing ethnic diversity will contribute to changes in central political institutions, such as citizenship, and may affect the very nature of the nation-state.

These theoretical conclusions help to explain the growing political salience of issues connected with migration and ethnic minorities. The migratory movements of the last 50 years have led to irreversible changes in many countries. Continuing migrations will cause new transformations, both in the societies already affected and in further countries now entering the international migration arena. The more descriptive accounts which follow will provide a basis for further discussion of these ideas. Chapters 3–7 are mainly concerned with the early stages of the migratory process, showing how initial movements give rise to migratory chains and long-term settlement. Chapters 8–11 are concerned mainly with the later stages of the migratory process. They discuss the ways in which settlement and minority formation affect the economies, societies and political systems of immigration countries.

Guide to further reading

Amongst the many recent works on globalization, the following are useful as introductions: Castells (1996, 1997, 1998), Held *et al.* (1999), Bauman (1998) and Cohen and Kennedy (2000). Two recent works provide overviews of international migration theory: Massey *et al.* (1998) presents a systematic discussion and critique (based on two earlier articles: Massey *et al.* 1993 and 1994), while Brettell and Hollifield (2000) contains chapters addressing the contributions of some of the main social scientific disciplines to the study of migration. Boyle *et al.* (1998) is a good introductory text written by geographers. An earlier, but still valuable, compendium on migration theories is to be found in a special issue of *International Migration Review* (1989, 23:3). Kritz *et al.* (1992) is an excellent collection on migration systems theory. Phizacklea (1983), Morokvasic (1984) and Lutz *et al.* (1995) have edited useful collections on the relationship between migration and gender. Sassen (1988) gives an original perspective on the political economy of migration, while Borjas (1990) presents the neo-classical view.

Goldberg and Solomos (2002) is a comprehensive collection of essays on various aspects of racial and ethnic studies. Rex and Mason (1986) provides detailed expositions of theoretical approaches to race and ethnic relations. Mosse (1985), Cohen and Bains (1988), Miles (1989), Balibar and Wallerstein (1991), Essed (1991) and Wieviorka (1995) are good on racism. Anderson (1983), Gellner (1983) and Ignatieff (1994) provide stimulating analyses of nationalism, while Smith (1986, 1991) discusses the relationship between ethnicity and nation. Analyses of the relationship between migration and citizenship are to be found in Bauböck (1991, 1994), Bauböck and Rundell (1998), Aleinikoff and Klusmeyer (2000, 2001) and Castles and Davidson (2000). Gutmann (1994), Schnapper (1994), Soysal (1994) and Kymlicka (1995) present various perspectives on the same theme. DeWind (1997) is a collection of articles on the changing character of immigrant incorporation in the USA. Good introductions to the emerging field of transnational communities include Basch *et al.* (1994), Cohen (1997), Portes *et al.* (1999), Vertovec (1999) and Faist (2000). Van Hear (1998) discusses transnational theory from the perspective of refugee movements. Zolberg and Benda (2001) is very useful for understanding the links between economic migration and refugee movements.