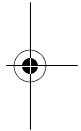
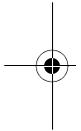


1

THEORY AND EXPLANATION

Consider the following facts. The richest 20 percent of Brazil's population earns about 30 times more than the poorest 20 percent. This level of inequality has continued virtually unchanged from the 1950s, when analysts began measuring income inequality in Brazil. In the Russian Federation, this ratio (of income shares of the richest and poorest 20 percent of the population) was around 12 in 1998, which is about twice as high as it had been just a decade earlier, before the collapse of the Soviet Union.¹ These are dramatic examples: one of a very high level of inequality that has remained unchanged for decades, the other of a rapid change from a low to a high level of inequality. How do we explain this diversity of experiences? There is no doubt that random events (such as Boris Yeltsin's speech atop a tank during the 1991 coup attempt in Moscow) play a large role, just as there is no doubt that the structure of inequality can be understood in terms of economic, social, and spatial structure, and inequality change can be understood in terms of changes to these underlying structures. I begin from the position that, despite the likelihood of random events that can change distribution patterns, explanation is possible.

However, the search for explanation is stymied by epistemological boundaries. Because distribution is one of the pillars of development² there is a large literature on inequality. Unfortunately, this literature, and therefore our understanding of inequality, is fragmented by academic discipline. The discourse on income distribution is dominated





2 • Fragments of Inequality

by economists. The discourse on classes, stratification, and power is in the domain of sociologists. And spatial inequality has traditionally been studied by geographers and regional scientists. The theoretical foundations and methodologies of these disciplinary approaches are often so far apart that it has been difficult to create a solid theoretical understanding of inequality as an outcome of economic, social, and spatial processes. This book attempts to fill some of these large gaps in interdisciplinary knowledge about the structure of inequality and processes of distributional change using a new theoretical approach combining elements of economic, sociological, and geographical theory.

This new approach is built on evolutionary foundations. I begin from the recognition that human behavior and action follow evolutionary principles (identified in fields as diverse as evolutionary psychology, sociobiology, and behavioral economics) and not some unproven assumptions about rationality and self interest. This means that individuals and groups are both important elements of social and spatial structure, and, therefore, both are relevant for theorizing structure and change in inequality. Second, we must acknowledge that evolution itself is not linear and continuous. The standard Darwinist-gradualist view of evolution has to be supplemented with the near certainty of discontinuity and nonlinearity, or what is called the punctuated equilibrium model of evolution. These ideas are detailed later in this chapter and permeate the explanatory part of the book (Chapter 4 onward).

These shifts—from linear to punctuated models, from individuals to groups, from abstract and monolithic to fragmented space—purposefully suggest a fundamental shift away from the dominant mode of inequality analysis, which, as is well known, is the economic approach. Because of the limiting assumptions used in mainstream economic theory, assumptions that infer substantially more equality (of knowledge and power) than exists in reality, economic approaches turn out to be limited in their explanatory power. Inequality, I conclude, is too important a subject to be left to economists.³ I retain the important contributions: the theories on human capital, the interaction of demand and supply of different forms of capital, but reject the narrow specifications of the rational, self-interested actor model. The goal is to shift the discourse away from economic to social theories of inequality.

QUESTIONS AND ANSWERS

The analysis focuses on income inequality. The idea of income inequality is easily understood, relatively easily measured, universal in its

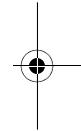
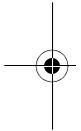


manifestation, and tangible, at some level, to everyone with social awareness. Data are collected to measure income inequality with increasing frequency and sophistication. Therefore, there is a concrete empirical basis from which one can begin an examination, and to which one can turn for support or falsification. I seek answers to the four fundamental questions on income inequality:

- What explains the level of income inequality in a given nation?
- Why do income inequality levels vary so greatly worldwide?
- What causes the level of income inequality to change?
- What explains the diversity of trends in income inequality change?

I come to the answers using an approach that seems simple and obvious but has never been used. Let me begin with the obvious: the world is fragmented. It is fragmented into geographical or spatial units that differ in terms of the average life chances of their inhabitants. This gives rise to spatial inequality. The spatial units themselves are fragmented into social groups with unequal power and resources; group membership is a significant determinant of an individual's life chances. This gives rise to social inequality. I argue that income inequality in any given nation is a result of its particular combination of social inequality (which arises from social fragmentation or heterogeneity) and spatial inequality (which arises from spatial fragmentation or heterogeneity). Nations vary in their specific combinations of social and spatial fragmentation that are the outcomes of specific histories; hence, they vary in their levels of income inequality. Broadly, the more fragmentation there is, the higher is the level of income inequality. Distributional change takes place as a result of changes to social and spatial inequality. Changes in social inequality can be dramatic when there are distributional transitions, which are possible under conditions of revolution, invasion, and war; in general, though, changes in social inequality are more likely to be gradual. Changes in spatial inequality are always gradual. The diversity of trends in distributional change are explained by the fact that nations differ in the rates and directions of change in social and spatial inequality and in their specific histories, which may or may not include one or more incidents of fundamental distributional change. In general, in the last 50 to 100 years, social inequality levels have declined in most nations (especially in the more developed nations) while spatial inequality levels have increased.

To put it in another way: a nation can be thought of as a combination of a social system and a spatial system. The social system is made





4 • Fragments of Inequality

up of individuals who compete for scarce resources and for status or recognition. The individuals are also organized into groups that are at least class-based and usually also identity-based (using ascriptive markers of race, ethnicity, religion, or language). Systematic and durable differences exist between group average incomes. Groups with access to more productive resources and power have higher average incomes and superior life chances than groups without. Hence, every society is fragmented to some degree. A social system is historically formed through migrations, invasions, wars, revolutions, expulsions, and trade, all of which have brought people with distinct ascriptive identities into common geographical confines. The variations in these conditions or events are largely responsible for the level of heterogeneity in a social system. A spatial system is made up of urban and rural areas in the first instance, with further delineations between urban areas (big city or small city) and rural areas (valley and hill, coastal and inland). These spatial differentiations are also largely historically constructed as a result of imperialism, colonialism, and trade. The different spaces offer different average incomes and life chances for their residents. Hence, a national territory is also spatially fragmented to some degree. Combinations of these two fragmentations create the conditions of economic inequality in given nations.

The critical questions relate therefore to change—social change and spatial change—and this is where I focus. Specifically, I concentrate on two forms of distributional change: the quick distributional transition as a result of state transition and consequent fundamental social change, and the slow, gradual transformation as a result of spatial transformation. One cannot, however, get to these narratives without a clear understanding of the structures that exist and are changed, and the agents and the processes of change. As a result, the first half of the book is devoted to setting the table. I begin by presenting data on income distribution and distributional change, with next a discussion of economic theories of inequality followed by a discussion on why and how economic theory must be and can be supplemented with social and spatial theory to create a true picture of the structure of inequality. Finally, in the later chapters (5 and 6), I am able to focus on the question of change.

THREE GENERAL PRINCIPLES

The setting would not be complete without a discussion of the principles that are foundational for this work. Some of these will become

obvious from the brief arguments outlined in the following section. Nevertheless, it is necessary and useful to have a clear understanding of what it means to try to build theory using basic principles from the economic and social sciences. This distinction is important. At several points in the book, particularly in Chapters 3, 4 and 5, I refer to the idea that there are two fundamentally different ways of understanding human action and interaction: in simple terms, we can call these the economic perspective and the social perspective. I argue that it is not possible to build theories that make sense when these two perspectives are kept in separate boxes, rather, it is necessary to find ways to integrate them. Let us consider (briefly, since these ideas are spelled out in greater detail in other chapters) the principal building blocks of such an integrated theory.

First, we need to resolve a fundamental question in a non-judgmental way: What is the basic unit of a society? In the economic perspective, this is the individual. Individuals compete in a world of scarce resources to attain their primary goals of survival first and growth second, where survival and growth both have intergenerational dimensions. This means that people are primarily self-interested beings. (There are complications about the meaning of self-interest; these are taken up in Chapter 4.) The social perspective does not deny the importance of self-interest, but suggests that it is moderated by the interests of the groups to which the individual belongs. Hence, in the social perspective, the basic unit of a society is the group. Groups compete in a world of scarce resources to attain their primary goals of survival first and growth second, where survival and growth both have intergenerational dimensions. This means that groups are primarily self-interested entities (and basic group interests are established as norms, more on which soon). Virtually every statement we can make about individuals we can also make about groups: they compete, they are self-interested, they seek domination over others, they can engage in violence when threatened, they are interested in intergenerational continuity. It stands to reason, therefore, that individuals are motivated by both individual and group interests, and often the two cannot be identical. What happens when individual interests clash with group interests? Again, it stands to reason that the latter generally prevails, but, and this is a very important idea, it is possible for individuals to persuade groups to change or modify their interests. This, often, is the source of social change.

Therefore, to build sensible theory, we must recognize the coexistence of individual and group interests. In defining groups, however,



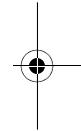
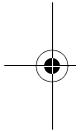
6 • Fragments of Inequality

we run into the problem of plurality, because individuals, more and more, are members of multiple groups. They are members of racial, ethnic, religious, and linguistic groups, and, without exception, they are members of economic groups or classes. This problem of plurality is usually partially resolved because the groups substantially overlap; that is, ethnicity, race, and class intersect. The problem can be tackled even more definitively if we recognize the existence of another type of group, one that is based on geography or location. At small geographical scales, groups are more easily delineated and it is possible to identify paired oppositions.

Individuals are organized into groups and groups are organized into territorial units. Therefore, we have another level of competition. Almost anything we can say about individuals and groups we can repeat about territorial units: they compete for scarce resources, they seek domination over others, etc. The world, I repeat, is fragmented. Individuals compete with other individuals, groups compete with other groups, and territories compete with other territories. Individuals often have to act, not in their individual interest, but in the interest of the group or the territory. On the positive side, it follows that individuals cooperate and have reciprocal relationships with members of their group and their territory. A good understanding of the social world must begin from recognizing this tripartite division of interests and the potential for inter-unit conflict and intra-unit cooperation and competition.

Second, economic and social interactions are characterized by “increasing returns” and “norms.” Increasing returns are also called “cumulative causation” processes (they are not identical phenomena, as the latter includes the problem of vicious cycles) or “positive feedbacks” in which the “payoff to taking an action (increases with) the number of people taking the same action ... or the payoff to engaging in a collective action depends on the number of participants” (Bowles 2004, 12). Norms include two overlapping classes of features: *ideologies*, which are sets of beliefs and desires, and *habits*, which are routine and standardized responses to a variety of situations. The existence of these features is usually well known by agents when they take decisions; as a result, they often turn into self-fulfilling prophecies or self-reinforcing actions. Let me explain:

Formal economic models of production are based on the assumption of constant returns to scale and diminishing returns to capital. These two assumptions drive (among other things) the preoccupation with convergence and equilibrium that are the mainstays of economic



thought. Later we will see, for instance, that income convergence between territorial units is supposed to be the long-term outcome. Yet common sense suggests that many critical aspects of economic and social life are the way they are because of increasing returns, or at least the existence of the general belief that there are increasing returns. Cities, for instance, would not exist without scale economies and increasing returns. If it were equally easy for an entrepreneur to locate a factory or office anywhere, why would she locate it where rents are high? Either there really are benefits to locating where many other similar enterprises exist (that is, there really are increasing returns to the density of interactions), or it is a habit, a decision taken without much conscious thought because one presumes that others who have taken similar decisions must have given it much thought. Hence, increasing returns and habits result in the growth of cities, and consequently, there is spatial heterogeneity. Increasing returns are not limited to geography, but are common in other significant areas. As far as income distribution is concerned, the most important of these are intergenerational increasing returns, which result in the concentration of property ownership and widening disparities in human capital acquisition.

The existence of norms (ideologies and habits) simultaneously constrains and simplifies decision-making. Such norms also seriously damage the “rationality” and “perfect information” assumptions of mainstream economic theory. Individuals, groups, institutions, and states all have ideologies and habits. They constrain decision-making because when the number of options considered is limited, it is possible that better options than the ones chosen are not even considered. At the same time, since fewer options are considered, the cost of making decisions is minimized. One can think of norm-based vs. rational behavior as analogous to the thought processes of chess champion Garry Kasparov vs. the chess computer Deep Blue. From almost any given situation in a chess game there are millions of possible moves and paths. Kasparov ignores almost all of them because he thinks they lead to nowhere good; sometimes he misses the best possible move. Deep Blue, on the other hand, has to compute the outcome of each path, including the obviously pointless ones, before making a move.

This analogy is useful but incomplete because it makes no reference to the relationship between norms and power. Norms are inscribed with the relations of power in a society—among genders, classes, social groups, and between the state and its subjects or citizens. Norms prescribe and proscribe behaviors that uphold the social order. They are the most direct expressions of group interests, and the interesting

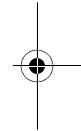
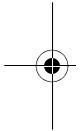


8 • Fragments of Inequality

aspect is that this happens without constant collective action. To understand the structure of income distribution in a society it is vitally important that we know about its norms. To understand the processes of distributional change, it is necessary to know how norms change.

Third, we must have a clear understanding of change in general. Specifically, we must incorporate the possibility of discontinuities in evolutionary change. Theories in mainstream economics are built upon the understanding that social and economic systems resemble models in physics—more accurately, Newtonian physics. Objects (agents) follow physical laws (economic principles) that tend to keep a universe of objects (the economic system) in equilibrium. Any perturbation (shock) to this universe sets forth forces (economic actions) that restore equilibrium. This physical metaphor makes many logical problems analytically tractable, but it bears little resemblance to the universe of real people. Many social and economic theorists prefer to use biological and evolutionary metaphors. In these views, equilibrium does not and cannot exist because organisms (agents) continually seek to perpetuate (grow) their genes that have the possibility, but not certainty, of undergoing spontaneous mutation (change). In a world of genetic competition, only those organisms survive whose genetic mutations give them the capacity to better adapt to their environments. The Darwinian view of evolution is that it is a gradual process, incremental genetic changes being compiled over millions of years to produce the still evolving organic forms we see today. The post-Darwinian view is that there are discontinuities (punctuations) or sudden sharp transitions in evolution rather than (or in addition to) continual minute adaptations. Stephen Jay Gould, the primary proponent of the punctuated equilibrium thesis, draws his metaphors from the social science of history (Gould 2002). History, he says, is marked by quick transitions separated by long equilibria; so is evolution.

There are obvious problems in basing analysis on metaphor. Robert Nisbet (1969, 6) points out that: “to build rigorous propositions of scientific analysis upon ... metaphor, mistaking attributes of analogy for attributes of reality, can be ... profoundly limiting and distorting.” Nisbet also argues that change is not natural or normal, lacks direction, is not immanent or endogenous, is non-cumulative and non-patterned, and is neither necessary nor completely explainable. This is a serious attack on both the evolutionary and mechanical views of social change. Nisbet was a political conservative. He would probably be shocked at his intellectual cohort today, made up as it is of post-structuralist, post-modernist, post-grand narrative scholars, but as far



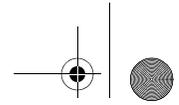
as I am concerned, his critique has excellent intellectual credentials.⁴ Individual actions are indeed unpredictable. Social groupings are often shifting and their identities often amorphous. Grand theories are over-inclusive and not very good for prediction. Nonetheless, when there are patterns and trends (as there are in income distributions) good social scientists must begin from the position that there are general explanations. This book is, in the tradition of liberal modernism, a search for explanations for change beginning from the principle that some changes are regular and some unpredictable.

THE SPECIFIC ARGUMENTS

The general principles discussed above form the foundation on which are constructed the detailed arguments of this book. Let us now consider the specific arguments that are detailed in the coming chapters.

Inequality levels and trends over the last five to seven decades are diverse and variable but identifiable. Income inequality levels vary from very high in many Latin American nations to moderate in most developing and south Asian nations to very low in most northern European and several central European nations. Inequality trends are also variable. In evolutionary terms, the trends show evidence of both *Darwinian gradualism* and *punctuated equilibria*. That is, there is evidence that income inequality levels usually change very slowly, often not at all; these are instances of gradualism. There are also instances where income inequality levels change very substantially in very short periods; these are punctuations. Every nation has the potential to undergo both forms of change. Most nations, including obvious examples like Russia and China and less obvious ones like the U.S. and India, have undergone both; some others, such as Scandinavian nations, have undergone only the gradualist form of change in the modern period. These realities, along with an important list of cautions on income distribution data comparability over time and space, are established in Chapter 2 (Patterns and Trends).

In Chapter 3 (Economic Theory and Income Distribution) I present a summary of economic theories on inequality and suggest that on the whole they provide incomplete and narrow explanations. *Functional theories* of distribution (also known as factor theories), where the determinants of inequality are the demand-and-supply conditions for the factors of production (land, labor, and capital), are abstract and unrelated to political and social reality; in fact, this approach is discounted even by economists who specialize in income distribution. *Theories of personal*

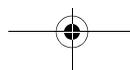
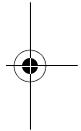
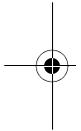


10 • Fragments of Inequality

income distribution, which are based on conceptualizing how luck, inheritance, ability, education, age, and policy influence the earning capabilities of individuals, are highly contested. However, these theories, especially human capital theories, are foundational for understanding how individuals are rewarded in market societies. *Theories of distribution change*, which identify variables such as growth, trade, technological change, etc., as the key reasons for change, are deterministic or particularistic and unable to provide general explanation. The Kuznets hypothesis (which suggests an inverted-U relationship between growth and inequality) and its obverse, the endogenous growth theory that initial equality leads to better growth performance, are both seen to be too linear, simplistic, and non-historical; more to the point, they are empirically unverified. The trade-based (Stolper-Samuelson) and technological change-based arguments are better theorized but are nonetheless simplistic abstractions of social and political reality. Many of the empirical tests of these theses are marred by the use of regression models, which are based on a linear evolutionary understanding of change.

The right way to analyze the structure of income distribution must begin from the fundamentals of human behavior. The rationality and self-interest assumptions of economic theory have to be reformulated with new theories of knowledge and action derived from behavioral economics, sociobiology, evolutionary psychology, social psychology, and more traditional sociological theory. This reformulation suggests that people act as individuals as well as members of groups. Their knowledge is perceptual and socially embedded and their actions habitual and largely conformist. Hence, social theory on institutions, groups, and power relations embedded in specific spatial (which are also historical) contexts, must be combined with economic theories of demand and supply of capital and skills to create a sound explanatory framework. These positions can be summarized in four arguments:

1. Individual knowledge and action rules are informed and constrained by cultural and institutional contexts.
2. Self-interest is often subsumed by group-interest, whereby societies are characterized by in-group cooperation and out-group competition.
3. The concentration of power, in the market or in non-market institutions, influences the distribution of resources and income.
4. The appropriate geographical scale for understanding income generation and distribution processes is the local scale.



These arguments are fleshed out in Chapter 4 (Social Theory and Income Distribution) and lead to the conclusion that the level of inequality in a society is a function of social fragmentation (expressed through inter-group bias and within-group solidarity) and spatial fragmentation (significant and systematic differences between geographical units). Income inequality can, as a result, be disaggregated into specific combinations of *social inequality* (that is, inequality within, but more importantly, between groups) at a local level and *spatial inequality* (that is, inequality between sub-national territorial units). At the end of Chapter 4, I set up and discuss a numerical illustration with different combinations of within-group, between-group, and between-region inequality, and show that social inequality (between-group inequality) at the local scale, and between-region inequality are the two primary components of variation in society-wide or national inequality.

Next, in Chapter 5 (Punctuated Equilibria and Social Inequality), I turn to questions on the composition of and changes to social inequality. This, I argue, is best accomplished by understanding the laws and norms that influence distribution. Institutions make laws and society has norms that create the conditions that determine the returns to inheritance and human capital, that is, the specific market rules of specific societies. Laws (or institutional settings) are changeable, and when there are significant state transitions, laws can and often do change. Norms are more durable. The most important norms that sustain durable social inequalities are marriage norms (homogamy, which dictates that marriages take place between status equals) and inheritance norms. In Chapter 5 I use these principles to construct a theoretical framework in which transition agents, who can be endogenous (insiders) or exogenous (outsiders), change the laws and norms of a society by mutation, adaptation, or invasion, based on an ideology of surplus expropriation, redistribution, or reinvestment. Insiders are reformers who change norms and laws at the margin; outsiders are revolutionaries who radically transform inter-group relations and laws. Hence, change agents and their ideologies are at the center of distributional change arising from social change.

Punctuated equilibrium models of inequality change can be explained using this theoretical framework, where the equilibria are “initial conditions” and the punctuations are “institutional transitions.” I will show that there are many examples of distributional punctuations. Some are progressive, others are regressive. For instance, several distributional transitions in the early to mid twentieth century have been progressive, all at the end of the century have been regressive. Case



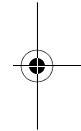
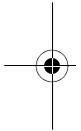
12 • Fragments of Inequality

studies of colonialism and nationalism from Latin America (Brazil and Mexico) and Asia (India and Indonesia), socialism and post-socialism (in the Soviet Union / Russia, China, Poland), non-socialist revolutions in Bolivia and Iran, postwar reconstructions in Taiwan, South Korea, and Japan, etc., repeatedly show that outsiders have wrought serious and rapid distributional change whereas insiders have worked toward gradual change. This theory is also used to show how social inequalities are reconstructed in many significant institutional transitions, and how pluralistic democracies may make such transitions impossible; hence gradualism may, over time, become the only model of distributional change.

I take up a detailed discussion of spatial fragmentation in Chapter 6 (Gradualism and Spatial Inequality). Spatial inequalities are created through trade and market exchange mediated by power, and “cumulative causation” and “increasing returns” are the key features of spatial divergence. State ideology (expropriation, redistribution, or reinvestment) is able to significantly influence the spatial distribution of economic activity, and market processes usually reinforce the tendency toward geographical clustering of more productive activity. Largely because state ideologies are episodic, changes in spatial inequality are also episodic. Nevertheless, the last two hundred years have featured significant geographical divergence, especially in developing nations. This tendency toward income concentration in leading regions or interregional divergence is manifested at all spatial scales, and is especially pronounced when smaller geographical units are compared.

Changes in spatial inequality are relatively slow. Certainly they do not have the speed of fundamental social change, but they are probably faster than gradual social change. Hence, changes in spatial inequality represent the major form of *gradualist models* of inequality change. Case studies from Brazil and Mexico, India and Indonesia, Russia and China, Japan, the U.S., and Europe illustrate the arguments. A remarkable aspect of the current period of globalization (post-1980) is the universal tendency toward increasing spatial inequality, including in developed nations where, during the middle third of the last century, spatial inequalities had declined substantially.

In the final chapter (Where We Stand) I consider the consequences of these arguments. I show that distributional change, the phenomenon that is at the heart of this analysis, depends on the joint outcome of changes to social inequality and spatial inequality. In the gradualist mode, when both change in the same direction, income inequality changes in that direction. That is, when both social and spatial inequality





decline, income inequality declines, when both increase, income inequality increases. In the middle third of the twentieth century, social and spatial inequality declined in the more developed nations (and later in China); their income inequality levels declined as a result. The most common combination, however, pairs declining social inequality with rising spatial inequality. This leads either to distributional stasis or gradual, very slow change in one direction.

The most pressing distributional question of our time—how does globalization affect income inequality?—can be understood in terms of how globalization affects social and spatial inequality. The effect of globalization on social inequality is indeterminate. Certainly there is little scope for punctuations or fundamental change under democracies. On the other hand, the long-term prospects for social change should be progressive, especially in situations where the historically dominant groups are also numerical minorities. However, spatial inequalities have increased and will continue to increase almost everywhere and at every scale. I suggest that increasingly segregated urban structures and regional variance arising from differential engagement with globalization will increase spatial inequalities at all geographical scales, from the neighborhood to the international system. Hence, we can expect globalization to lead to less social fragmentation and more spatial fragmentation, possibly with the net effect of gradual increases in income inequality in most nations. The answers are not clear cut, but there are reasons to believe that movement toward income equalization within individual nations may not be possible in the foreseeable future.

