

---

## Causal Asymmetries

---

Many debates in the social sciences revolve around issues of causal primacy and other asymmetries.<sup>1</sup> Often, proponents of contending positions agree that certain causes are relevant for explaining some phenomenon, but differ in their assessments of the relative importance of these causes (*quantitative asymmetry*) or in their understanding of the qualitatively different ways in which they enter into particular causal processes (*qualitative asymmetry*). Our aim in this chapter is to clarify the meaning of both quantitative and qualitative causal asymmetries and analyze their interconnection. We shall argue that sustainable causal primacy claims amount to assertions of one or another kind of *quantitative* asymmetry; claims for causal primacy that appeal to *qualitative* asymmetries either reduce to quantitative asymmetry claims or else are confused in ways that elude successful reconstruction.

These issues have been particularly salient in debates between Marxists and their critics. Marxists have generally argued that class (or the closely associated concepts of economic structure or forces of production) is the most important cause of many phenomena—from large-scale social changes to forms of the state, ideology and the oppression of women.<sup>2</sup> Critics of Marxism, including many “post-Marxists”, have

---

1. In the philosophical literature on causation, the expression “causal asymmetry” typically designates asymmetries between causes and their effects. We shall use the expression differently—to designate asymmetries among causes in multicausal systems.

2. There are contemporary Marxists who reject claims for the causal primacy of class or, perhaps more precisely, reject the very legitimacy of making such claims. In somewhat different ways, the idea that it is meaningless to assign weights to different causes has been advanced by Richard Wolfe and Stephen Resnick in *Knowledge and Class* (Chicago: University of Chicago Press, 1988), and by Barry Hindess and Paul Q. Hirst, *Marx's Capital and Capitalism Today* (London: Routledge & Kegan Paul, 1977).

argued that other causal processes, irreducible to class, are at least as important. Two examples illustrate the issues in contention.

Both Marxists and neo-Weberians agree that class structure and the institutional rules of the political game are causes of state policies in contemporary capitalist societies. They differ, however, in their claims about the role these causes play. Marxists generally hold that class structure is the most fundamental cause of state policies, while neo-Weberians claim either that political factors are more important, or that the relative importance of particular causes depends on historical circumstances. Typically, Marxists also argue that there is a qualitative asymmetry between class and political institutions in the shaping of state policies: class structure determines the *limits* of possible variations in state policies, while political institutions and practices only *select* policy outcomes within these limits. A parallel claim is usually absent from neo-Weberian accounts of state policies: class and political factors are simply treated as two relevant causes.<sup>3</sup>

Another example arises in debates between Marxists and feminists over explanations of the oppression of women. Feminists and Marxists generally agree that specific forms of oppression of women—from wife battering and sexual harassment to job discrimination, unequal divisions of housework and political exclusion—are affected both by causes rooted in class relations and by causes rooted in gender relations. However, most Marxists, following Engels, have insisted that class domination is the most important cause of the oppression of women, and even that the system of property relations determines the ways in which other causes operate. Thus, while many Marxists acknowledge the importance of patriarchal ideology in explaining the oppression of women, they argue that the consequences of these cultural factors depend upon the class structural context within which they operate.<sup>4</sup> Thus they argue for the causal primacy of class and for the qualitative asymmetry of class and gender. Non-Marxist feminists have usually insisted, in opposition, that gender-based mechanisms, distinct from and

3. A Marxist perspective on these problems is provided by Goran Therborn in *What Does the Ruling Class Do When It Rules?* (London: NLB, 1978). A neo-Weberian perspective can be found in Theda Skocpol, "Bringing the State Back in: False Leads and Promising Starts in Current Theories and Research", in Peter Evans, Dietrich Rueschmeyer and Theda Skocpol, eds, *Bringing the State Back In* (Cambridge: Cambridge University Press, 1985), pp. 3–37; and "Political Response to Capitalist Crisis: Neo-Marxist Theories of the State and the Case of the New Deal", *Politics & Society* 10:2 (1980). For a comparison of Marxist and neo-Weberian approaches which attempts to forge a Marxian synthesis, see Robert Alford and Roger Friedland, *The Powers of Theory* (Cambridge: Cambridge University Press, 1985).

4. See Johanna Brenner and Maria Ramas, "Rethinking Women's Oppression", *New Left Review* 144 (1984).

irreducible to class factors, are more important than class in explaining forms of women's oppression, and also that these mechanisms have effects that are independent of class structure.<sup>5</sup>

Remarkably, causal asymmetries have received little attention from philosophers of science. While there is a vast literature on the metaphysics of causation and the status of causes within scientific explanations, there has been hardly any discussion of what it means to assign differential importance to causes in multicausal systems or of the qualitatively different ways in which causes enter into causal processes.<sup>6</sup> If the intuitions that practicing social scientists had about these issues were clear and consistent, the dearth of philosophical analysis would be understandable. But of course, there is no consensus among social scientists. Throughout the social sciences, causal primacy claims abound—amidst general confusion about what causal primacy and other causal asymmetries involve. Such confusion is an impediment to theoretical advance.

### Preliminary Considerations

Before launching into an analysis of causal asymmetries, it will be useful to lay out some of the assumptions we shall make and to indicate the boundaries of our discussion.

### Causes and Explanations

We take it for granted that it is legitimate to attempt to construct causal explanations of particular phenomena by identifying the underlying

5. The purported mechanisms may be psychosexual (e.g. Nancy Chodorow, *Mothering*, Berkeley: University of California Press, 1978); biological (Mary O'Brien, *The Politics of Reproduction*, London: Routledge & Kegan Paul, 1983); cultural (Michèle Barrett, *Women's Oppression Today*, revised edition, London: Verso, 1989); or economic (Heidi Hartman, "The Unhappy Marriage of Marxism and Feminism", in Lydia Sargent, ed, *Women and Revolution*, Boston: South End Press, 1981).

6. For some discussion of causal primacy, see Richard Lewontin, "The Analysis of Variance and the Analysis of Causes", *American Journal of Human Genetics* 26 (1974) pp 400–11, Elliott Sober, "Apportioning Causal Responsibility", *Journal of Philosophy* 85 (1988), pp. 303–18, Richard Miller, *Fact and Method* (Princeton, NJ: Princeton University Press, 1987); and Clark Glymour *et al.*, *Discovering Causal Models in the Social Sciences* (Orlando: Academic Press, 1987). There is an extensive literature that attempts to partition the causal factors contributing to an event, for example, by distinguishing triggering causes and causal background conditions. See, for example, the essays in Tom Beauchamp, ed., *Philosophical Problems of Causation* (Encino and Belmont: Dickenson Publishing Co., 1974). This body of work, however, is not focused on the main subject of our analysis since it is not concerned to explicate a quantitative or comparative notion of "greater causal importance".

mechanisms that generate them. We thus adopt a "realist" view of scientific explanation. In our view, these mechanisms exist independently of our theories of them. As realists, we reject the stance, emblematic of "post-modernist" discourse theory, that science is simply one linguistic practice among others, in which the validity of claims is settled entirely within its own discursive practices.<sup>7</sup> We assume, in other words, that causes are real and that science aims at their discovery. Our task is to sort out the senses in which, in particular explanatory contexts, some (real) causes are more important than others, and the different ways in which these causes enter into causal processes.

We shall sometimes use *cause* and *explanation* interchangeably. Some readers may find this usage tendentious. But it is not. We are agnostic here in the continuing debate on the relationship between causality and explanation. Some philosophers have insisted that scientific explanations must always be causal; others have denied this view.<sup>8</sup> We sympathize with the former position, but the analysis that follows does not depend on its truth. It is enough, for our purpose, if it is conceded that *some* explanations are causal. These are the only explanations that concern us. When we talk about explanations of events, we mean *causal* explanations. Whether these are the only explanations proper to science, it is causal explanations that generate problems about causal primacy and causal asymmetries.

### Genuine vs. Spurious Causes

Distinguishing genuine causal relations from spurious correlations is a persistent problem throughout science. This problem is particularly acute in the social sciences, where reliable techniques for testing hypotheses are frequently unavailable. However, it will not be necessary for us to address this problem here. Our aim is only to make sense of the claim that one (genuine) cause is more important than another and that different (genuine) causes enter into causal processes in qualitatively distinct ways. Accordingly, we shall not attempt to defend the causal explanations we use as illustrations, but shall assume that real causal relations, not spurious correlations, are involved. If our examples are incorrect, others could be substituted.

7. For an influential discussion on the left of the view that causal relevance is constructed within discourse, see Ernesto Laclau and Chantal Mouffe in *Hegemony and Socialist Strategy: towards a radical democratic politics* (London: Verso, 1985).

8. For example, compare Wesley Salmon's *Scientific Explanation and the Causal Structure of the World* (Princeton, NJ: Princeton University Press, 1984) with Carl Hempel's *Aspects of Scientific Explanation* (New York: Free Press, 1965).

### Pragmatic vs. Explanatory Importance of Causes

It is sometimes suggested that primacy claims are only pragmatic. Thus, from a political point of view, the "most important cause" of poverty might be the cause that is most susceptible to political manipulation. Attributions of relative importance, therefore, could simply be designations by investigators of those causes that, for one reason or another, interest them more. We do not doubt that there is, in fact, an important pragmatic component to many actual primacy claims, nor even that some primacy claims are only expressions of investigators' interests. However, we do deny that causal primacy is *only* a pragmatic notion. In what follows, we shall not be concerned with the pragmatics of explanation at all. Our aim is to provide an account of primacy claims that represent or purport to represent objective asymmetries among real causal processes.

### Epistemological Primacy

In many empirical problems, information about some causes is more important than information about others in affecting the ability to predict outcomes. In this sense, some causes are more important epistemologically than others. If, for example, there are two separately necessary and jointly sufficient conditions for an outcome, and one of these conditions is almost always present while the other is present only half of the time, it is more useful—for predictive purposes—to find out whether the second condition obtains than whether the first condition does. Thus suppose that the necessary and sufficient conditions for the election of a socialist candidate are (a) the presence of a working-class majority in the electorate, and (b) the presence of a well-organized socialist party. Imagine too that the working class is a majority in 90 per cent of all constituencies but that socialist parties are well organized in only 50 per cent of elections. In this case, knowledge of the party variable would increase the ability to predict the outcome of the election more than knowledge of the class variable.

However, like pragmatic considerations generally, epistemological primacy is not causal primacy in the sense that interests us. The relative importance of a cause for producing outcomes is not the same thing as the relative importance of *knowledge* of that cause for predicting outcomes. In a causal process in which two necessary conditions are jointly sufficient, there is no sense in saying that one condition is causally more important than another.

### Individuating Causes and their Effects

A problem that constantly befuddles debates about the importance of different causes in multicausal processes is the correct designation of the object of explanation (the explanandum) and of the causes that explain it (the explanans).

Consider, first, the object of explanation. For disputes about causal primacy and other asymmetries to be resolvable, the disputants must of course agree about what they are trying to explain. In practice, however, many debates over causal primacy are confounded by subtle—or not so subtle—shifts in the explanatory problem. For example, in the debate between Marxists and neo-Weberians over the explanation of state policies in capitalist societies, the issues are frequently put in apparently opposed, abstract ways. However, when we turn to the empirical arguments offered on one or the other side, it often turns out that the neo-Weberians are trying to explain relatively fine-grained details of the timing and provisions of state policies, whereas the Marxists are trying to explain relatively coarse-grained properties that involve the consistency of the policies with the reproduction of capitalism.<sup>9</sup> Both parties could be correct about their respective explananda. There may still be disagreements about which of these explananda is more important to study—because of pragmatic considerations, reflecting the interests of investigators, or because certain explananda may themselves be important as causes in other explanations. But debates over what questions to ask should not be conflated with debates over the relative importance of causes in explanations of the same phenomena. While many causal primacy disputes only arise in consequence of confusions or unclarity in specifying explananda, the issues we shall explore occur in contexts where the explanandum is fixed.

Problems of demarcating the categories used in explanations are not

9. In their analysis of the origins of the modern welfare state, for example, Ann Orloff and Theda Skocpol ("Why Not Equal Protection? Explaining the Politics of Public Social Spending in Britain, 1900–1911, and the United States, 1880s–1920", *American Sociological Review*, 49:6 (1984), pp. 726–50) argue that Marxist accounts are unsatisfactory since class-centered causes do not explain the specific timing of the introduction of key pieces of welfare state policy. In explaining the introduction of national old age insurance programs in the United States, Britain and Canada, for example, they demonstrate that the timing across the three countries is heavily shaped by the specific institutional properties of the state. A Marxist reply to this analysis could be that while class factors may not explain why these policies were introduced in Britain before the First World War, in Canada in the 1920s and in the United States in the 1930s, class-centered mechanisms do explain why no capitalist state had such policies in the mid-nineteenth century while all developed capitalist societies had such programs by the mid-twentieth century. This reply, of course, constitutes a shift in explananda—from explaining variations across countries in the introduction of social programs to explaining the common temporal pattern among these countries.

restricted to the definition of the "dependent variable"; they are equally important, and difficult, in specifying "independent variables". The importance of a cause in the explanation of some phenomenon plainly depends, at least in part, on how the contributing causes are described.

Consider the explanation of fatal automobile accidents. It might be concluded that driver dysfunction is a more important cause of automobile accidents than are driving conditions (weather, road quality, etc.), according to some plausible understanding of "most important cause". But if "driver dysfunction" is decomposed into, say, drunkenness from beer consumption, drunkenness from hard liquor consumption, drunkenness from wine, incompetence, drowsiness caused by medications, etc., we might then find that something else—perhaps weather or road conditions—becomes the "most important" cause of fatal automobile accidents within this expanded, disaggregated list. Thus, it might be held that conclusions about causal primacy are artifacts of an essentially arbitrary decision about how to aggregate different causes under more abstract systems of classification.

One way out of this difficulty is to advance the claim that certain causes can be grouped together in non-arbitrary ways. One might argue, for instance, that the intoxicants that lead to drunkenness should all be grouped together since they all cause accidents through the same proximate accident-producing mechanism. Then drunkenness in its various forms would constitute a "natural kind" in the inventory of causes of accidents. To the extent that such natural kind groupings can be elaborated, claims about causal primacy become less vulnerable to arbitrary re-descriptions of the explanatory problem.<sup>10</sup>

In the social sciences, the problem of describing causes and effects "correctly"—in a way that enhances understanding of the causal structure of the world—is pervasive. Frequently, there are no solid, theoretical foundations for classifying concrete causes into more abstract, natural kind categories. Thus, social scientists often are forced to rely on common-sense determinations and ordinary intuitions to categorize causes and the phenomena they purport to explain. There is little else they could do.<sup>11</sup> However, there is no a priori reason why such common-sense descriptions should correspond to real causal structures; investigators might someday *discover* that some (or all) of their intuitions about natural kind divisions are wrong. Arguably, some extant work in social

10. When no unique set of natural kinds is available, it may be useful to see whether a claim of causal primacy is robust over changes in the underlying taxonomy. Invariance of this sort is one sign that the asymmetry is not an artifact of one's mode of description.

11. In the physical and biological sciences, well-confirmed theories provide investigators with a better purchase on natural kinds, but even here puzzles about how to individuate explanans and explananda can arise.

science, Marxist or otherwise, already points in this direction. Until such theoretical advances are firmly established, substantive debates about causal asymmetries will always be vulnerable to redescription of the categories deployed in debates.

While the problem of describing causes and their effects may plague many actual debates over causal asymmetry, we do not believe that this difficulty diminishes the relevance of trying to refine our understanding of the logic of these asymmetries. Getting the right descriptions is an aspect of the general problem of discerning real causal relations; understanding how those causes are quantitatively and qualitatively interconnected remains a distinct problem. It is to that problem that we now turn.

### Quantitative Asymmetry

We shall identify two kinds of quantitative asymmetry, each tantamount to a sense of *causal primacy: distribution-dependent causal primacy* and *causal potency*. The first asymmetry expresses the idea that some causes are “more important” than others in explaining particular phenomena; the second expresses the idea that some causes are “more powerful” than others in producing particular effects.

#### Distribution-dependent Causal Primacy

It is commonplace to identify relative causal importance with the *relative frequency* of a given cause. Thus it would be natural to say that smoking is a more important cause of lung cancer than plutonium exposure, if more instances of lung cancer are caused by smoking than by exposure to plutonium. This claim does not imply that smoking is a more powerful carcinogenic agent than plutonium. It is compatible with the distribution-dependent primacy of smoking over plutonium exposure that plutonium is, in some intuitive sense, more “dangerous” than smoking. Thus, we might think that a “small amount” of plutonium poses a greater risk than does a “small amount” of smoking and still insist that smoking is a more important cause of lung cancer.

The distribution-dependent importance of a cause is a function of two relations: its distribution in the population and its potency. Two carcinogens could be equally frequent causes of cancer (and thus equally important in the distribution-dependent sense) despite very different distributions, if the rarer property is sufficiently more carcinogenic. This idea can be represented using the following definitions:

$F(s_i)$  = the percentage of the population that smokes at exposure level  $i$

$F(p_j)$  = the percentage of the population that is exposed to plutonium at level  $j$

$P(s_i)$  = the probability of an individual getting cancer *because of* exposure to smoking at level  $i$

$P(p_j)$  = the probability of an individual getting cancer *because of* exposure to plutonium at level  $j$ <sup>12</sup>

The levels “ $i$ ” and “ $j$ ” in these expressions are specified in whatever units are chosen for the factor in question—say, cigarettes per day or grains of plutonium, or whatever.

The percentage of the population that will, on average, in fact get cancer due to a given causal property at a given level of exposure is given by:

$C(s_i) = F(s_i) \times P(s_i)$  = the percentage of the population that will get cancer due to smoking at exposure level  $i$

$C(p_j) = F(p_j) \times P(p_j)$  = the percentage of the population that will get cancer due to plutonium at exposure level  $j$

The total incidence of cancer due to a given causal factor in a given population, then, is:

$$C(s) = \sum_i C(s_i) = \sum_i [F(s_i) \times P(s_i)]$$

$$C(p) = \sum_j C(p_j) = \sum_j [F(p_j) \times P(p_j)],$$

12. The probability of a person getting cancer *because of exposure* at a given level is not the same as the conditional probability of getting cancer for people who are exposed at that level. Suppose that 10 per cent of people with no exposure to any known carcinogen get cancer and 25 per cent of people who are exposed to smoking at level  $i$  get cancer. Under these conditions one could argue that smoking at level  $i$  increases the probability of getting cancer by 15 per cent (relative to the probability of getting cancer without smoking). The relevant probability here is therefore the *difference* in *conditional* probabilities of getting cancer for people at given exposure levels compared to some baseline. We are thus assuming in this example that the increase in conditional probability of an individual getting cancer from exposure at a given level compared to not smoking at all is an appropriate way of measuring the *causal potency* of that level of exposure. This may not always be the case. In focusing on potency, we want to know *how much of a difference* a cause makes towards its effects. If in real situations exposure to a given carcinogenic agent is itself empirically associated with some cancer-suppressing cause—if, for example, people who smoked also happened to be less prone genetically to cancer than non-smokers—then the empirical differences in conditional probabilities would not effectively identify the causal potency of the specific exposure levels. We shall ignore such complications in this analysis. For a discussion of these issues see, Nancy Cartwright, “Causal Laws and Effective Strategies”, *Nous* 13 (1979), pp. 419–37.

where in each case the summations are carried out over all the exposure levels for each cause.

The claim that smoking is a more important frequency-dependent cause of lung cancer than exposure to plutonium then means just that  $C(s) > C(p)$ .<sup>13</sup> In this case, the lower cancer-inducing levels of smoking affect a sufficiently greater proportion of the population than do the higher cancer-inducing levels of plutonium exposure—to a degree such that the greater potency of plutonium is offset. Smoking is a more important cause of cancer than plutonium in consequence of the actual frequencies and potencies of these causal factors.

The object of explanation in this case is the *distribution* of lung cancer within a population, not the contraction of the disease by specific individuals. For any given individual with lung cancer, exposure to plutonium could be the most important cause of his or her illness. To determine the importance of causes in this sense, it would be necessary to assess the potency of each of the causes that actually affected the person's history.<sup>14</sup> However, our concern here is not with particular events, but with the distribution of events in populations. To claim primacy for smoking in this sense is to maintain that, given the relative potency and distribution of the carcinogenic agents in contention, smoking is the most important cause of lung cancer within the population.

Frequency and potency are distinct properties of causal mechanisms. What explains the frequency of a cause will generally not explain its potency. To explain the frequency of smoking in a population, it might be necessary to appeal to theories of cultural mores or to invoke the political power of the tobacco industry. The extent of plutonium exposure can perhaps be explained by the growth of nuclear energy production. Causal potency, on the other hand, would, in all likelihood, be explained by biochemical or physiological accounts of the role exposure to these agents plays in producing cancers of the lung and other organs.

13. If some people were exposed to both plutonium and smoking, the equations would become more complex, but so long as the effects of exposure from each carcinogen are independent of the other—so long as the effects are strictly additive, not multiplicative—then the simultaneous exposure to two carcinogens in no way changes the meaning of distribution-dependent causal primacy discussed here. If, on the other hand, the two carcinogens interact in ways such that their conjoined effects are greater than their additive effects, attempts at attributing causal primacy in a distribution-dependent sense could well break down.

14. Explaining how an individual contracted lung cancer is similar to assessing situations of "temporal asymmetry" among a set of causes (see p. 160). The central issue is to identify which, among a variety of causes that intersect the biographical trajectory of the individual, had the biggest effect on the probability of that person's getting cancer. This is a problem of relative causal potency, not distribution-dependent primacy.

In this example, the explanandum, lung cancer, is dichotomous; a person either has it or does not have it. In such dichotomous cases, claims about distribution-dependent causal primacy, therefore, concern the relative importance of different causes in affecting the conditional probabilities of the outcome. Nothing of importance hinges on this fact. Our account would not be substantially changed had the explanandum been a continuous variable. The only difference is that, with continuous variables, we would have to compare correlations, not probabilities.

For example, in sociological studies of income inequality, two factors widely deemed causally efficacious are education and "occupational status" (roughly a measure of the social standing of an occupation). Which is the "more important" cause of personal income?<sup>15</sup> As posed, this question is unanswerable because a unit of education and a unit of status are incommensurable. A one-year difference in education may make a bigger difference in income than, say, a one-point difference on some scale of occupational status; but so far we have no more reason to compare one year with one status point than with, say, 365.

To overcome this problem, sociologists typically compare correlations among variables rather than effects measured in terms of fixed units. The correlation between education and income reflects how much deviation from the average amount of education is associated with deviations from the overall average income. The resulting units are "standard deviations", measures of deviations from the mean within a particular distribution relative to the total dispersion of values in that distribution.<sup>16</sup> A correlation of education and income of 0.5 means that an individual who moves up one standard deviation from the mean on education (which in a typical "normal distribution" would mean having an education at about the 67th percentile of the distribution), would have on average 0.5 standard deviations more income.

15. In the present context we shall ignore complications arising from the fact that, in the real world, these causes are not strictly independent since education is itself a cause of occupational status.

16. More technically, a standard deviation is defined in the following way:

$$s = \sqrt{\frac{\sum_{i=1}^N (X_i - X)^2}{N}}$$

The expression within parentheses is the deviation of each individual observation,  $X_i$ , from the mean of  $X$  for all observations. The standard deviation, then, is calculated by squaring each of these individual deviations, adding up these squared terms over all observations and then dividing by the total number of observations,  $N$ . The greater the dispersion of the distribution around the mean value of  $X$ , the greater will be the standard deviation.

Suppose that the correlation of status and income is 0.6 and the correlation of education and income is 0.4. What the correlations in this case show is that variation in income is more closely tied to variation in status than to variation in education.

Does this difference in correlations justify the claim that status is a "more important" cause of income than is education? Correlation is not causation. This truism might suggest that, among real causes, different *degrees* of correlation need not indicate different *degrees* of causal *importance*. However, this inference is incorrect: if a causal problem is properly specified, then in the same sense that one is justified in seeing some causes as more important than others because they are more frequent, one is justified in seeing differences in correlations as differences in causal importance.

The formula for a correlation coefficient is given by the following equation:

$$r_{xz} = b_{zx}(s_x/s_z)$$

Where:

$r_{xz}$  = the correlation of x and z

$b_{zx}$  = the unstandardized slope relating z to x (i.e.  $b_{zx}$  tell you how many units of z increase for each unit of x, when both are measured in their "natural" units)

$s_x$  = the standard deviation of the distribution of x

$s_z$  = the standard deviation of the distribution of z

Now suppose we have two correlations with z,  $r_{xz}$  and  $r_{yz}$  (the former could be the correlation of education and income, the latter status and income). To hold that  $r_{xz} > r_{yz}$  is to say that:

$$b_{zx}(s_x/s_z) > b_{zy}(s_y/s_z)$$

which reduces to:

$$b_{zx} \cdot s_x > b_{zy} \cdot s_y$$

The claim that a cause with a higher correlation constitutes a "more important cause" than one with a lower correlation rests on the linkage between the distribution of this cause in the population and the strength of its effects on Z. As with causal frequency, the comparison of correlations depends on one term ( $b_{zx}$  and  $b_{zy}$ ) that reflects the strengths of the effects of X and of Y on Z, and one term ( $s_x$  and  $s_y$ ) that is strictly a

function of the distributional properties of the two variables.<sup>17</sup>

One might object to the idea that a greater correlation implies a more important cause on the grounds that  $b_{zx}$  could be greater in some meaningful sense than  $b_{zy}$  while the associated correlation is smaller. This situation would resemble the case where plutonium is held to be a more potent cause of cancer than smoking, at the same time that it is a less important cause of the actual distribution of cancer in a particular population. To say that "X is more important than Y as a cause of Z because X has a higher correlation coefficient" means that the *distribution* of X is a more important cause of the *distribution* of Z, *given* the linkage between the distribution and causal potency of X.<sup>18</sup>

### Causal Potency

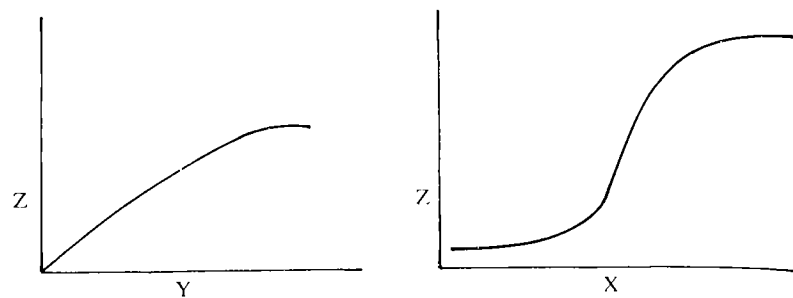
Causal potency plays a role in distribution-dependent causal primacy, since the relative importance of different causes depends upon the linkage between their prevalence and the strength of their effects. However, in making distribution-dependent causal primacy claims, it is not necessary actually to compare the potency of two causes. Causal potency can therefore be considered a distinct kind of causal primacy.

The problem of assessing the relative potency of different causes is, of course, that causes come in radically heterogeneous units. There is, as noted, a powerful intuition that plutonium is a more potent cause of lung cancer than smoking. A tiny grain of plutonium lodged in the lungs will almost certainly produce cancer, whereas steady smoking for years

17. Note that while the standard deviation of Z— $s_z$ —affects the actual correlations, it does not enter into the comparison of the two correlations.

18. This account of correlation coefficients has no direct implications for the relative explanatory importance of *changes* in the distributions of X or Y for *changes* in the distribution of Z—because the causal potencies of X and Y may itself be caused in part by the distributions of X and Y. Thus, for education, the income-generating effect of, say, a high school degree almost certainly depends in part on the distribution of educational levels in the population: if the proportion of high school graduates increases, the causal efficacy of a high school degree would likely decline. Moreover, it is partly because of the income-producing power of education that people seek it. In this (dynamic) sense, the distribution of educational levels is in part explained by the causal potency of education on income. Wherever there is reason to suspect some dynamic interdependence of the causal potency of X and Y with their respective distributions, it is illegitimate to draw conclusions about the consequences of changes in the distributions of X and Y for changes in the distribution of Z from facts about the correlations of X and Y with Z.

This kind of problem, in the guise of frequency-dependent selection, is important in evolutionary biology. It occurs when the fitnesses of traits in a population depend on their frequencies. For example, the advantage an organism receives from protective coloration might be enhanced by the trait's rarity. Another example is studied in the theory of the evolution of sex ratios. In many instances a parent maximizes fitness by producing offspring of the minority sex. See Elliott Sober, *Nature of Selection* (Cambridge, MA: MIT Press, 1984) for discussion of these issues.

**Figure 7.1** Contrasting Functional Forms Linking a Cause and its Effect

merely makes lung cancer more likely. Still, there seems to be no theoretically well-motivated way to regard a “tiny” amount of plutonium as a comparable unit to “years” of smoking. Unless the units of two causes can be made commensurable, it is meaningless to compare their relative causal power.

The situation, however, is not quite hopeless. Whenever it is possible to map out the magnitude of effects for all relevant values of a cause—by establishing the overall functional relation between a cause and its effects—there are a variety of strategies that allow investigators to standardize the “units” of the cause for purposes of comparing causal potencies. Suppose that we want to compare the potency of two causes, X and Y, on Z, where these have the functional forms represented in Figure 7.1.<sup>19</sup> Each of these graphs is expressed in the natural units of each cause. Cause X has an S-shaped relation to Z: as values of X increase, Z increases quite slowly at first, then very rapidly, asymptotically reaching some maximum value. Cause Y has a more linear relation to Z over most values of Z. With these two functional forms, there are various ways one could compare the potency of the two causes. One could compare the maximum effects of the two causes—i.e. the magnitudes of Z at the point where further increases in the causes produce no change in Z. In this case, X would be maximally more potent than Y. Or, one could pick a midpoint in the function and ask whether around that point a 10 per cent increase in X produces a larger effect on Z than does a 10 per cent increase in Y. In that region of the values of the two causes, X appears to be more potent.<sup>20</sup> Such comparisons do not stan-

19. Again, we shall not explore the issue of possible interactions between X and Y in producing Z; for simplicity we shall only discuss the additive case. In many contexts, of course, the functional form of the relationship of X to Z will itself vary over values of Y.

20. As these functions illustrate, if the causal relationship between X and Z is non-linear, then the verdict about the causal potency of X compared to some other variable may depend upon the point at which such comparisons are made.

**Table 7.1** Hypothetical Relation between Race, Gender and Annual Earnings

		Men	Women	Row means
Race	White	\$20,000	\$12,000	\$16,000
	Black	\$16,000	\$8,000	\$12,000
Column means		\$18,000	\$10,000	\$14,000

Cell entries are mean incomes for race by gender categories (fictional data).

standardize the effects of the causes in terms of the factual distributions of these variables, but in terms of the causal structure that links them to their effects.<sup>21</sup>

There is one common kind of cause for which these functional relations can be straightforwardly mapped out—causes that assume an essentially binary (on/off) form. For example, if we wanted to know whether the racial division between blacks and whites or the gender division between men and women has a more powerful impact on individuals’ earnings in the United States, we can directly compare the causal potency of the two causes. Consider the hypothetical example presented in Table 7.1. In this table, the difference in income between men and women is \$8,000 while the difference between whites and blacks is \$4,000.<sup>22</sup> A straightforward interpretation of these results is that the gender dichotomy has greater causal potency for earnings than does the race dichotomy.

This conclusion depends upon the assumption that these dichotomies

21. In the example just discussed, it could be the case, for example, that in the real world the values of X are all located at the lower part of the possible values, so the steep part of the curve is never encountered, whereas the values of Y are located at all values.

22. In this concocted example, the difference between men and women is exactly the same among blacks as it is among whites (and equivalently, the racial differences are identical within each gender). In technical terms this means that there are no “interactive” effects between the two factors being considered; the effects are strictly additive. That is, the difference in income between white men and black women is equal to the simple sum of the gender difference and the race difference. It should also be noted that this table can be represented as a multiple regression equation using “dummy” variables (0–1 variables) to represent gender and race in which:

$$\text{Income} = \text{constant} + B_1\text{RACE} + B_2\text{GENDER}$$

where RACE = 1 for whites, and 0 for blacks, and GENDER = 1 for men, and 0 for women. In this example, the constant = \$8,000,  $B_1 = \$4,000$ , and  $B_2 = \$8,000$ . If there had been an interactive effect of race and gender—i.e. if the effects of gender were different in the two races—then this would appear as a multiplicative term in this equation (Gender  $\times$  Race).



adequately represent the metric for the two causes under investigation. Treating the two causes as dichotomies, in effect, implies that a "one unit" change in sex is equivalent to a "one unit" change in race. At first glance, the assumption that these causes can be adequately represented in binary form seems acceptable, particularly for sex, in so far as sexes are biologically dichotomous. The situation is less clear for race, to the extent that it is a social construction, not a biological given, that racial distinctions are binary. One can imagine a world in which what is causally important about "race" is skin color, and in which skin color is finely graded into different shades. Then the proper metric of race-causes would not be the simple dichotomy black/white. In consequence, questions could be raised about the legitimacy of comparing a "unit" of race (however understood) with a unit of sex.

On further scrutiny, similar considerations apply even to sex. Sex is biologically dichotomous, and this fact, in virtue of its salience, affects its socially efficacious causal properties. But "gender" (sex as a social category) need not mimic sex as a biological category. It is conceivable, for instance, that sex-causes operate through a relatively continuous metric—from highly masculine to highly feminine. The biological dichotomy might still be a good empirical *indicator* of these causal determinations. But, again, a "unit" of sex would no longer be commensurable with a "unit" of race.<sup>23</sup>

Whether it is appropriate to treat sex or race as dichotomous variables depends on how these causes operate in the world. If biological sex is causally efficacious for earnings only because of its link to masculinity/femininity or if race is causally efficacious by virtue of its linkage to gradations of skin color, it would misrepresent the causal powers of these variables to treat them as if they were dichotomous. This is not an issue that can be decided a priori, but only after evaluation of the relevant evidence.

The requirements for establishing either of the kinds of quantitative causal asymmetries we have identified—distribution-dependent causal primacy or causal potency—are arduous. To demonstrate distribution-dependent causal primacy it is necessary to know the distribution of the various causes in the relevant population, and to have a way to measure the magnitudes of their effects. To demonstrate relative causal potency it is necessary to be able to specify the functional form that links causes to their effects.

23. If race and/or gender were really continuous variables, then there would still be ways to render their units comparable. One could, for example, compare the effects for the extreme values for each variable. This strategy should not be confused with an analysis based on a simple dichotomy.

In practice, these requirements are often impossible to satisfy for many of the issues that have animated debates in the social sciences, especially among radical scholars. Generally, it is impossible to define the distribution of causes or their relative potency in explanations of such phenomena as large-scale social change, revolutions, the contradictions of state policies, and the transformation of gender relations.

There is a variety of ways that social scientists react to this situation. One response is to restrict investigations to those for which properly specified quantitative answers can be provided. Doing so effectively diminishes social science, forcing a focus only on easily measurable phenomena distributed in well-defined populations. This response is characteristic of much contemporary American sociology.<sup>24</sup> A second response is to continue to ask the broad, classical questions, but to avoid making primacy claims. The explanatory objective of social science would then be just to identify causes, without trying to specify their relative importance. Many radical social theorists today embrace this objective.

A third alternative is to shift attention away from quantitative to qualitative asymmetries. Often, qualitative asymmetries are taken as providing the basis for causal primacy claims. We shall argue, however, that differences in the way causes enter into multi-causal processes have nothing directly to do with relative causal importance. It is often of great explanatory interest to discern qualitative asymmetries. But it is a mistake to regard these differences as distinctions in causal importance.

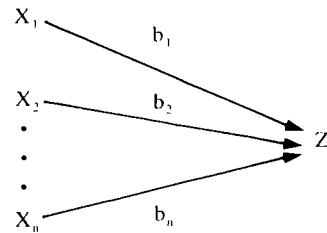
### Qualitative Asymmetry

In the simplest causal structure in which several causes operate together to produce some outcome, each cause generates its effects independently of the others. The result is the "sum" of the effects of each separate cause. This simple additive model is represented in Figure 7.2. When the outcome is an "event", the effects of each cause can be interpreted as its impact on the probability of the occurrence of the event. Each cause, then, contributes to this probability independently of the others.<sup>25</sup> When the outcome is a "variable", each cause contributes to

24. The narrowing of questions motivated by the requirements of measurement also tends to direct analysis towards data tagged to individuals. In general, it is easier to define the relevant populations and observe the relevant distributions of attributes when the variables attach to individuals rather than institutions, structures, societies or states.

25. The sum of the effect-probabilities of a set of additive causes can be greater than 100 per cent. This is particularly the case when there are multiple sufficient causes present for some event to occur.

Figure 7.2 Underlying Causal Model without Qualitative Causal Asymmetries



the magnitude of the variable. For most real-world processes, additive models are obviously implausible. Nevertheless, most empirical research in the social sciences assumes this kind of model. While this assumption may be useful in some contexts, many of the debates that animate social science, particularly in its more radical traditions, involve issues raised by more realistic accounts that countenance qualitative asymmetries.

The asymmetries we shall explore vary along two dimensions: first, whether they are *systemic* or *contingent*; and second, whether they are *synchronic* (involving the simultaneous operation of several causes) or *diachronic* (involving the temporal ordering of causes). Putting these dimensions together yields four general categories, as represented in Figure 7.3: contextual asymmetry, functional asymmetry, temporal asymmetry and dynamic asymmetry. We shall discuss each of these asymmetries in turn.

### Contextual Asymmetry

In many instances, some causes—which we shall term “contextual causes”—determine the conditions under which other causes generate their effects. Such contextual causes need not be integrated into any general system; contextual asymmetries, as in the examples that follow, can exist among contingent assemblages of causes. In addition, since both contextual causes and the causes they structure operate simultaneously, contextual asymmetry claims are synchronic.

We shall examine two putative types of contextual asymmetry. The first, based on the contrast between causal limits and selections, does represent a genuine qualitative asymmetry. However, this type of asymmetry has often been misconstrued as a kind of causal primacy. The second case we shall investigate focuses on interactive effects of two or more causes operating simultaneously. Contrary to first appearances, such interactions do not constitute an instance of even qualitative causal asymmetry, or so we shall argue.

Figure 7.3 Forms of Qualitative Causal Asymmetry

	Synchronic	Diachronic
Contingent	Contextual asymmetry	Temporal asymmetry
Systemic	Functional asymmetry	Dynamic asymmetry

### Limits and Selections

The concept of “structural limitation” has played an important role in contemporary Marxism. The idea is that in some explanatory problems it is possible to partition causes into two general categories: those that explain the *range of possible outcomes*, and those that explain *what actually occurs* within that range of possibilities. The first set of causes is said to impose “limits” on the outcome, whereas the second “selects” an outcome from within those limits. Thus in rational choice theory, it is common to distinguish the feasible set of alternatives individuals face from the choices they make from within that set. The former is explained by, among other things, the distribution of resources, the institutional rules of the game, and the relative power of actors; the latter by such subjective factors as individuals’ preferences. The causes implicated in setting limits are often said to be “more fundamental” than those that select outcomes from within those limits in so far as the former establish the conditions of possibility under which the latter operate.

This kind of explanatory configuration has played an important role in on-going debates about politics—for instance, in discussions of “non-decision-making” and in the analysis of “non-events”.<sup>26</sup> The theory of non-decision-making focuses on mechanisms that exclude potential alternatives from the political agenda. It purports to explain why certain alternatives are never raised in the public arena, or, if they are raised,

26. For early discussions of these issues see Peter Bachrach and Morton Baratz, “The Two Faces of Power”, *American Political Science Review*, 51 (December 1962), pp. 947–52. Stephen Lukes popularized these issues in his analysis of the three faces of power in *Power: a Radical View* (London: Macmillan, 1974) in which he distinguishes between the power to directly affect a person’s behavior (face 1) and the power to define the range of alternatives or set the agenda (face 2). Claus Offe has presented a particularly pointed discussion of these issues as they apply to the problem of the capitalist state in his influential essay, “Structural Problems of the Capitalist State: class rule and the political system. On the selectiveness of political institutions”, in Klaus von Beyme, ed., *German Political Studies*, vol. I (Beverly Hills and London: Sage, 1974).

why they tend to be marginalized. The mechanisms that explain these "non-events" are usually quite distinct from the mechanisms that explain how decisions are made from among alternatives actually in contention.

Marxists argue that capitalist property relations impose limits on state policies. Exceptional circumstances apart, only policies that are broadly compatible with private capital accumulation are possible. The range of allowable policies may be more or less extensive, but the feasible set is limited, in the final analysis, by the functional requirement of reproducing capitalism. Within the range of possible policies, specific choices will generally depend on causal processes distinct from those that reproduce capitalist property relations as such—the rules that govern elections, patterns of regional conflict, the strength of unions, private interest associations, political parties and other collective actors on the political stage, cultural and ideological resources of political actors, and so on. These factors may themselves be shaped by the functional requirements of the capitalist mode of production, but they also enjoy a certain degree of causal autonomy. Thus it could be argued that under conditions in which capital accumulation is heavily based on industrial production, environmental pollution policy is limited by the requirement that it not so threaten profits as to precipitate massive disinvestment and capital flight. However, within this constraint, a wide range of environmental policies may still be feasible, and there is no reason to believe that the policies actually adopted will always be those most favored by capitalists. The mobilizing capacity of environmental movements and the rules of electoral politics as they intersect the geographical distribution of environmental issues might explain which policy is adopted from within the allowable set.<sup>27</sup>

Where a distinction can be drawn between limits and selections, an exclusive focus on the selection process will result in incomplete explanations. A more profound explanation will investigate both limits and selections. However, it is also tempting to maintain that causes that generate limits are *ipso facto* "more fundamental" or "more important" than are causes that select specific outcomes from within those limits. Thus Marxists often argue that since class structural factors impose limits on possible state policies, they are the most important causes of state activity. This conclusion, we believe, is not warranted.

Examples abound in which selections are more "important" than limits, according to any likely understanding of "important". Imagine the following case: an individual chooses a pear from a basket of fruit.

27. For a discussion of the social forces which impose limits on the agenda of pollution control, see Matthew A. Crenson, *The Unpolitics of Air Pollution: a study of non-decision-making in the cities* (Baltimore: Johns Hopkins University Press, 1971).

Two causes are involved: the range of fruits available in the basket and the person's preferences for different kinds of fruit. Suppose that there are thirty kinds of fruit in the world and that twenty-five of them are included in the basket. Which is the more important cause of the individual's choice of a pear—the composition of the fruit basket or the individual's tastes in fruit? The answer is indeterminate given the information so far specified. It might be that, even if all thirty kinds of fruit had been available, the individual would still have chosen a pear. In this case, the structural limitation on the individual's choice is irrelevant. On the other hand, if the individual would have preferred one of the five excluded kinds of fruit, the limiting process would provide an important part of the explanation of the final outcome.

In general, there is no simple way to determine whether the reduction of possibilities represented by "limits" is larger or smaller than the reduction represented by "selections", and unless we can compare this narrowing of options, claims about the relative importance of limits and selections are ill founded. More generally, to make the claim that a process of limitation is a more important cause of an outcome than is a process of selection is to argue, in the senses we have already discussed, either for the relative causal potency of limits and selections or for the relative distribution-dependent causal importance of the two kinds of causes. Unless units can be assigned so that the effects of limits and selections can be compared, determinate answers are likely to be unavailable.<sup>28</sup>

Nevertheless, there is a powerful intuitive sense on the part of political radicals that the Marxist claim (or something very close to it) is correct: that the limits imposed by the nature of the property relations in a society more powerfully explain the policies of the state than the mechanisms that select particular policies from within those limits. What explains this intuition, we believe, is imprecision in specifying explananda. What Marxists want to account for are not quite state policies as such, but certain *excluded* state policies—namely, radical, pro-working-class policies. The claim, then, is that the central mechanisms that

28. While it will, in general, be problematic to compare the magnitudes of a process of limitation and a nested process of selection (i.e. one that occurs within a given set of limits), it may be possible to compare two processes of limitation. That is, suppose  $X_1$  and  $X_2$  both impose limits on possible values of  $Z$ , without one of these constituting a selection within the other. Then, if all of the values of  $Z$  permitted by  $X_2$  are included in the set permitted by  $X_1$ , but not vice versa, we could unambiguously say that  $X_1$  more powerfully limits  $Z$  than does  $X_2$ . However, in real-world explanatory contexts in which multiple causal limits operate, it is generally not the case that one causal limit strictly subsumes another.

explain why the state does *not* systematically empower and mobilize the working class are causes that shape the agenda of politics—the limits—not causes that select options within the given political agenda.

It would be difficult to exaggerate the importance in this context of specifying explananda precisely. The vulnerability of structural limits asymmetry arguments to the specification of the object of explanation is illustrated well in a recent analysis by Ann Orloff and Theda Skocpol of the emergence of social insurance policies in capitalist democracies.<sup>29</sup> They chose as their “dependent variable” the timing of the introduction of social insurance in Britain, Canada and the United States—1904 in Britain, 1922 in Canada and 1933 in the United States. How should this timing be explained? Why this order? Why was the United States such a laggard, not introducing social security until the massive disruption of the Depression? Why was social security introduced in Britain so early? To answer these questions, they examine the kinds of explanation they feel would be advanced by Marxists (class conflict would explain the timing), by industrialization theorists (the level of development would explain timing) and by social democratic political theorists (the strength of socialist parties and labor movements would explain the timing). After showing that none of these explanations is satisfactory, they offer their own explanation, which revolves around the particular institutional capacities of these three states and the historical legacies of prior policies with which political forces in these societies had to contend. For example, they argue that the experience of pervasive corruption associated with Civil War pensions in the United States contributed significantly to the long delay in the introduction of social security there.

This analysis appears to conflict with Marxist interpretations, but we think the conflict is illusory. Very generally, the Marxist analysis of capitalist development explains why no capitalist country had a social security program in 1850 and why all developed capitalist countries had some form of social security by 1950. The structure of capitalist property relations and the conditions for the reproduction of capital accumulation explain the basic limits of possibility on such redistributive state policies. Within those limits, however, a wide variety of historically contingent factors—contingent with respect to theories of capitalist development—explain specific types of variation, such as the timing of the initial introduction of social insurance programs. By choosing a relatively fine-grained aspect of social policy—the timing of its introduction rather than the fact that it was introduced within, say, a 100-year period—the structural limits on the process fall into the background, and

29. Ann Orloff and Theda Skocpol, “Why Not Equal Protection?”

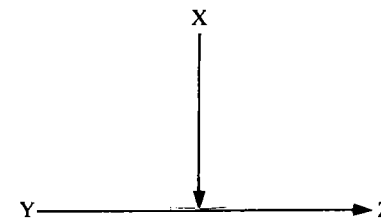
state-centered political mechanisms become more important.<sup>30</sup> Once alternative explananda are clearly distinguished from each other, and once it is granted that a two-stage process is incompletely analyzed by an exclusive focus on the second stage, it turns out that no issue of substance divides the opposing parties.

### *Interactive Causality*

There is a second kind of causal structure, which we shall call “interactive causality”, that appears to be a type of contextual asymmetry. As illustrated in Figure 7.4, cause X affects the causal relation between Y and Z. Of course, X may also directly affect Z in the simple additive way. The purported asymmetry derives from the fact that, in addition to its direct effects, X shapes the effects of Y on Z, while Y does not, symmetrically, shape the effects of X on Z. We shall argue that interactive causalities of this sort are not genuine contextual asymmetries.

Interactive causality arguments often appear in explanations of specific historical events, for example, when “precipitating events” are distinguished from “fundamental causes”. Thus in Figure 7.4, X might stand for the underlying social conditions that comprise the “fundamental cause” of some event; and Y could be viewed as a precipitating cause in the sense that it would not have the consequences that it in fact has, but for the presence of the fundamental causes with which it interacts. In explanations of the First World War, for example, the assassination of Archduke Ferdinand is often treated as a precipitating event

Figure 7.4 Interactive Asymmetry



30. This reconstruction of the argument by Orloff and Skocpol suggests a more general proposition: whenever one makes an argument about structural limits on some social process, it will be true that the more fine-grained the form of variation is that one is trying to explain within that process, the more likely it is that relatively contingent factors will play an important explanatory role. Imagine that instead of trying to explain the order of introduction of social insurance for these three countries, we found two countries that introduced their reforms in the same year, but in different months. Explaining the relative timing of introduction for those two countries, one would suspect, would be even less systematically determined than the timing explored in the Orloff-Skocpol research.

because it had the consequences it did only against the background of the geopolitical tensions generated by inter-imperialist rivalries in Europe. It is conceivable that, had the Archduke not been assassinated, the First World War might not have occurred. However, most explanations of the origins of the war assume that some other precipitating event would have occurred anyway, making the war virtually inevitable around the time it actually began. Most historians would therefore contend that the occurrence of *some precipitating event or other*, if not the actual assassination of Archduke Ferdinand, was a necessary condition for the outbreak of the First World War. Then the argument would be that the efficacy of this precipitating event, whatever it is, is asymmetrically related to the prevailing social and political conditions.

Marxists sometimes advance similar claims, especially when they inveigh against their critics. Thus Marxists frequently hold that the effects of many causal processes depend upon, say, the dominant mode of production or the structure of class relations or the balance of class forces. These claims depend upon the existence of interactive causal structures—linking distinctively Marxist explanatory factors with other causal mechanisms. An example of this kind of argument can be found in what has come to be known as the “Brenner Debate”.<sup>31</sup> Since the publication in 1976 of Robert Brenner’s article “Agrarian Class Structure and Economic Development in Pre-Industrial Europe”, the venerable Marxist debate over the transition from feudalism to capitalism has been joined by historians with a variety of theoretical orientations.<sup>32</sup> At the core of the debate are competing explanations for the differing trajectories of economic growth and, eventually, the emergence of capitalist development, among the different societies in early modern Europe. To explain these phenomena, Brenner argued that “economic rationality”, as it has been understood since Adam Smith, became possible and even inevitable in consequence of outcomes of historically specific class struggles, themselves consequent upon particular characteristics of the class structure of late feudalism in England and, in mitigated fashion, elsewhere in the West. Brenner also argued that now orthodox demographic and commercial explanations for economic

31. See T.H. Aston and C.H.E. Philpin, eds, *The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-Industrial Europe* (Cambridge: Cambridge University Press, 1985).

32. The article that launched the Brenner debate appeared in *Past and Present* 70 (February 1976). The earlier version of this debate was waged in expressly Marxian terms in the pages of *Science and Society*, in a series of articles sparked by Paul Sweezy’s criticisms of Maurice Dobb’s *Studies in the Development of Capitalism* (London: 1946, repr. 1963, 1972). The *Science and Society* debate has been republished, with additional material, as *The Transition from Feudalism to Capitalism* (New York and London: Verso, 1978).

growth are flawed in so far as they fail to acknowledge the centrality of class structure and conflict.

While historians may disagree about the facts of the matter, it is plain that the principal issues in contention in the Brenner Debate are of a conceptual character, and have to do with claims about causal interactions. From the outset, Brenner formulated his thesis in these terms:

It is the purpose of this article, to argue that . . . attempts at economic model-building are necessarily doomed from the start precisely because, most crudely stated, it is the structure of class relations, of class power, which will determine the manner and degree to which particular demographic and commercial changes will affect long-term trends in the distribution of income and economic growth—and not vice versa.<sup>33</sup>

In this statement, X in Figure 7.4 represents the “structure of class relations”, Y stands for “demographic and commercial changes”, and Z represents “long-term trends in the distribution of income and economic growth”. Brenner’s claim is that X determines the effects of Y on Z while Y does not determine the effects of X on Z.

We believe that the translation of these kinds of causal interactions into claims about qualitative asymmetries among causes is illegitimate. In the case of precipitating causes, one could as well say that the precipitating cause explains why the underlying conditions produced the effects they did when they did, as vice versa. Without that precipitating cause, the underlying conditions would not have had the consequences they did. If the underlying conditions appear more fundamental, it is because what they address is important to most investigators’ interests, while the explananda for which the precipitating event is indispensable are usually of little concern. Thus, in almost any imaginable explanatory program, it is more important to know, say, why a war of global dimensions broke out roughly when it did than to know why hostilities began precisely on August 14, 1914. Similarly, in the Brenner debate, one could argue that the endemic demographic cycles of feudal society explain why variations in class structure affected long-term growth the way they did. In both cases, the arrows in Figure 7.4 can be switched, with Y intersecting the arrow between X and Z, *without indicating any change in how these causes actually work*.

The essential symmetry in an interactive causal process is reflected in the mathematical equation generally used to represent such interactions.

33. “Agrarian Class Structure and Economic Development in Pre-Industrial Europe”, *Past and Present* 70 (February 1976), p. 11.

Take the Brenner theory of long-term economic growth in the transition from feudalism to capitalism. Very roughly, this process can be represented by the following equation:

$$\text{Growth} = B_1(\text{CLASS}) + B_2(\text{DEMOGRAPHY}) + B_3(\text{CLASS} \times \text{DEMOGRAPHY})$$

The coefficients  $B_1$ ,  $B_2$  and  $B_3$  designate the effects on growth of the factors contained in the parentheses. The interactive effect is represented by the multiplicative term. In Brenner's formulation of his thesis, quoted above, this expression is rewritten as follows:

$$\text{Growth} = B_1(\text{CLASS}) + [B_2 + B_3(\text{CLASS})](\text{DEMOGRAPHY})$$

The total effect of demography on growth— $[B_2 + B_3(\text{CLASS})]$ —thus embodies a term reflecting the class structure; or, as Brenner would have it, the manner in which demography affects growth is determined by class. The problem, of course, is that the equation can be rewritten symmetrically as:

$$\text{Growth} = [B_1 + B_3(\text{DEMOGRAPHY})](\text{CLASS}) + B_2(\text{DEMOGRAPHY})$$

In other words, there is nothing in the formal structure of the interaction that supports a claim for qualitative asymmetry.

Nevertheless, in both the mainstream historian's account of the origins of the First World War and in Brenner's analysis of the emergence of capitalism, there do seem to be real asymmetries among causes. In each instance, however, the asymmetry is quantitative, not qualitative.

To describe a cause as a precipitating event is to say that the probability of the outcome it helped produce was already very high before the precipitating cause occurred. The general social conditions, in our example, raised the probability of world war from a low to a high level, while the Archduke's assassination constituted the "final straw" that pushed the probability up to 1.<sup>34</sup> This is a special case of what we shall call *temporal asymmetry*.<sup>35</sup> It involves an assessment of the relative

34. The claim that the social conditions raised the probability to a high level implies a comparison with some appropriate counterfactual of what the probability would have been under alternative social conditions.

35. See pp. 160–5 below.

causal potency of the precipitating cause compared to the social context within which that precipitating cause occurred.

In Brenner's case, the asymmetry between demography and class structure is a special case of distribution-dependent primacy. Demographic patterns did not vary sufficiently across the various zones of Europe to explain variations in transitions, whereas class factors of the sort Brenner analyzes did. What this means is that although the demographic factors (interacting symmetrically with class) play an important role in explaining the transition within every society in which the transition occurred, it does not explain *variations* in transition across Europe. Variations must be explained primarily by class factors. Brenner's statement of his core thesis should therefore be revised as follows (changes appear in italics):

Attempts at economic model-building are necessarily doomed from the start precisely because, most crudely stated, it is *variation in* the structure of class relations, of class power, which will determine the manner and degree to which particular, *relatively universal*, demographic and commercial changes will affect *variations* in long-term trends in the distribution of income and economic growth.

This is an empirical claim about the way in which different causes with different patterns of distribution—one variable across Europe, one relatively invariant—intersected. It is not a claim about qualitative asymmetries.

### Functional Asymmetry

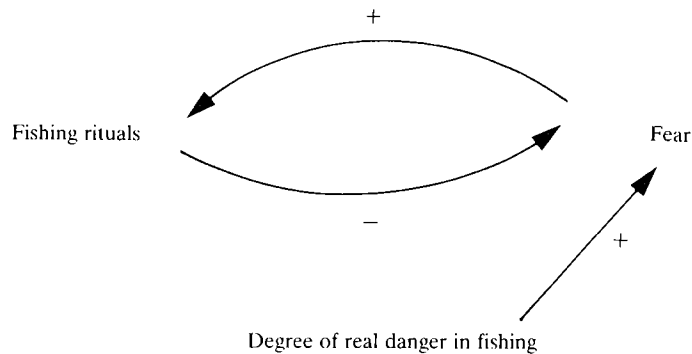
Functional asymmetries exist whenever causes are systemically joined as parts of a functionally integrated system. Like contextual asymmetries, functional asymmetries are synchronic.

Consider the famous example of functional explanation in Malinowski's study of fishing rituals among the Trobriand Islanders.<sup>36</sup> Malinowski observed that elaborate fishing rituals only occurred for deep-sea fishing, not for lagoon fishing. Deep-sea fishing was dangerous; lagoon fishing was not. Malinowski concluded that the rituals served the function of reducing the fear caused by the danger, and that the existence and persistence of the fishing rituals could be explained by this function. The structure of his explanation is represented in Figure 7.5.

In this model, fear functionally explains rituals: as fear rises there is

36. The use of this example to illustrate functional explanation, and the accompanying diagrammatic representation, come from Arthur Stinchcombe, *Constructing Social Theories* (New York: Harcourt, Brace, Jovanovich and World, 1968).

Figure 7.5 Model of Functional Explanation



pressure for the social production of rituals. As rituals increase in response to this pressure (through an unspecified search and selection mechanism), fear is reduced. An equilibrium occurs when the level of ritual effectively neutralizes the levels of fear necessary to produce more ritual. So long as the exogenous fear-producing mechanism (the level of danger associated with deep-sea fishing) remains the same, the level of ritual will therefore continue. A variety of feedback mechanisms could regulate such a functional system involving different mixes of conscious search for solutions to fear, trial and error, and social analogues to natural selection.<sup>37</sup> Whatever mechanisms regulate the system, rituals persist because of their functional relation to fear.

In this model, the explanatory asymmetry has a distinctive structure: the two terms—fear and ritual—have reciprocal effects on each other. Fear increases rituals; rituals decrease fear. The asymmetry comes from the functional interconnection of these terms: fear functionally explains rituals but rituals do not functionally explain fear. Thus one would not say that it is the function of fear to produce rituals while it is the function of rituals to reduce fear.<sup>38</sup>

37. For a discussion of the problem of specifying causal mechanisms within functional explanations, see G.A. Cohen, *Karl Marx's Theory of History: a Defense* (Princeton, NJ: Princeton University Press, 1978); Jon Elster, "Marxism, Functionalism and Game Theory: the case for methodological individualism", *Theory and Society* 11:4 (1982), pp. 497-512; Philippe Van Parijs *Evolutionary Explanation in the Social Sciences: an emerging paradigm* (Totowa, NJ: Rowan & Littlefield, 1981).

38. It is a truth about the world that the functional explanation in this example runs in one direction only. Imagine the following modification of the story: in the community there are professional ritual producers whose material interests depend upon the proliferation of rituals. They have learned that people are likely to attend rituals more consistently when those rituals are directed towards fearful situations. They therefore design rituals that dramatize the fear in ways that enhance the fear associated with the activity in the absence of the ritual. There would then be a functional symmetry: fear would function to produce rituals while rituals would function to reduce fear.

There is a second way of representing functional explanations, which we discussed in Chapter 3, that perspicuously reveals this asymmetry. G.A. Cohen has argued that functional explanations depend upon "dispositional facts" that pre-exist the establishment of functional relations. Thus, before the development of rituals, it was a dispositional fact of the culture that rituals would be fear reducing. This dispositional fact can then be taken to explain the emergence of the rituals:

- (1) *dispositional fact*: [Ritual → reduced fear]
- (2) *functional explanation*: [Ritual → reduced fear] → Ritual

While it also is a dispositional fact of the society that fear produces ritual, this second dispositional fact does not explain fear. Thus:

- (3) *dispositional fact*: [Fear → increased ritual]
- (4) *false functional explanation*: [Fear → increased ritual] → fear.

The fact that (2) is true while (4) is false establishes the functional asymmetry between fear and ritual.

Although we agree with Cohen that one can *represent* a functional explanation in terms of dispositional facts and their effects, we deny that this representation *defines* what a functional claim means. A causal connection can take the form depicted in (2) without entailing a functional claim. Consider a man who is obese; suppose that his obesity causes him not to exercise. The fact that he does not exercise has the consequence that he remains obese. The causal structure can then be represented as follows:

- (5) [Obesity → No Exercise] → Obesity.

If a statement of form (2) defined what functional claims mean, it would follow, implausibly, that the function of obesity is to prevent exercise.<sup>39</sup> We thus reject Cohen's analysis of functional claims and we do not offer a proposal of our own. It will suffice, for present purposes, to hold just that functional asymmetries exist whenever the persistence of some social practice is explained by its beneficial effects on something else, and these causes are joined by an appropriate feedback loop.

39. This counterexample is due to Christopher Boorse, "Wright on Functions", *Philosophical Review* 85 (1976), pp. 70-86. Boorse uses it to criticize an analysis proposed by Larry Wright in "Functions", *Philosophical Review* 82 (1973), pp. 139-68. Wright's account of functional explanation is essentially what Cohen endorses. Wright's and Boorse's papers are reprinted in Elliott Sober, ed., *Conceptual Issues in Evolutionary Biology* (Cambridge, MA: MIT Press, 1984).

Arguments for functional asymmetry have played an important role in recent Marxist debates. Thus, Cohen's reconstruction of historical materialism is based on a series of arguments about the functional asymmetry of technological development in explaining social change. On Cohen's account, Marx explains the nature of the social relations of production in a given society by the functions they fulfill for the development of the forces of production. In any society there are reciprocal effects of the forces of production on the relations of production and of the relations on the forces. A given technology can reinforce or undermine a set of production relations; the production relations can stimulate or retard (fetter) the development of the forces of production. Given such reciprocal effects, how can there be an asymmetry between forces and relations of production? Cohen argues that in classical historical materialism, this claim is based on a functional asymmetry within this pattern of reciprocal causation: the social relations of production are the way they are because of their beneficial effects on the development of the forces of production and not vice versa.<sup>40</sup> As noted in Chapter 2, this structure can be schematized as follows:

Level of PF → [PR → development of PF] → PR.

where PF = productive forces, and PR = production relations. In other words, the level of development of the forces of production explains which kinds of production relations will further enhance the development of the forces of production, and this (dispositional) fact explains which production relations actually pertain. There is then a qualitative explanatory functional asymmetry between the productive forces and the production relations.

These kinds of functional interactions do constitute a type of qualitative asymmetry in causal arguments. By themselves, however, they do not provide a basis for assigning differential explanatory importance to one or the other terms in the functional relation. In the functional interconnection of fear and rituals described by Malinowski, it makes no sense to say that fear is a "more important" cause of the level of ritual than ritual is of the level of fear. Similarly, in Cohen's reconstruction of historical materialism, it cannot be concluded, *simply by virtue of their functional interconnection*, that the forces of production are a more important cause of the relations of production than the relations of production are of the forces of production.

40. Cohen makes a parallel argument for the functional asymmetry of the economic structure with respect to the "superstructure": the superstructure takes the form it does and persists because of its effects on reproducing the economic structure (the "base").

Yet Cohen does argue for what he terms the "Primacy Thesis", the claim that the forces of production determine the relations of production to a greater extent than the relations determine the forces.<sup>41</sup> As argued in Chapter 5, we are sympathetic to the idea the Primacy Thesis expresses: that there exists an endogenous dynamic process in human history that causes some economic structures to be on the historical agenda and others not. But despite what Cohen suggests, this claim, if it is true, is not a consequence of the functional argument he has provided. It follows instead from two possible claims about the relative causal potency of different causes.

First, recall that, for historical materialism, the forces of production have a tendency to develop which ultimately leads them to generate instabilities in the relations of production (i.e. when the relations fetter the forces). Recall too that superstructural forms tend to *reproduce* existing production relations. We thus have two causes each affecting the relations: the forces destabilize the relations (when fettering occurs), the superstructures stabilize the relations. The critical thesis that allows for an epochal trajectory of social forms is that the first of these causal forces is more potent than the second. The Primacy Thesis, properly reconstructed, does not claim explanatory priority for the forces over the relations in consequence of how they are related functionally. It is not even a claim about the greater explanatory importance of the forces relative to the relations. It is strictly a causal potency claim: in general, the forces of production are more powerful than the superstructure *with respect to their effects on the relations of production*.

There is a second and parallel way of reconstructing the primacy thesis. The development of the forces of production is driven by two causes in historical materialism: by human nature (conjoined with scarcity) and by the relations of production. When these two causes contradict each other—when human nature pushes for increases in the forces of production and the relations fetter such development—human nature ultimately prevails. It is a more potent cause. This leads to changes in the relations (because human nature's causal potency is also stronger than the superstructures).

In both of these reformulations, Cohen's statement of the Primacy Thesis—the forces determine the relations to a greater extent than vice versa—is at best elliptical. The real issue is not that in a system of reciprocal causation, X explains Y to a greater extent than Y explains X. Rather, what we have is a structure in which relative causal potency claims are being made either about two causes of the development of the forces of production (human nature and the relations of production) or

41. See Chapter 2 above, and Cohen, *KMTH*, p. 134.



about two causes of the relations of production (the forces of production and superstructures). This primacy claim itself requires a functional analysis no more than do corresponding claims about the magnitudes of forces in physics. Cohen's functional proposal, we conclude, while interesting in its own right, is quite separate from the thesis of causal primacy that he advances.<sup>42</sup>

### Temporal Asymmetry

We have, to this point, examined qualitatively different ways causes operate at the same point in time. It will often be the case that evidence for such synchronic asymmetries will depend on examining changes over time. For example, to support functional asymmetry claims, it may be necessary to observe the development of a system to a point where functional requirements change. Nevertheless, the asymmetry claim itself is not about the trajectory of development of the system, but about the configurations of causes within the system.<sup>43</sup> We now turn to types of qualitative causal asymmetry—temporal asymmetry and dynamic asymmetry—that are essentially diachronic in character.

Temporal asymmetries obtain between causes in virtue of their location in a temporally ordered sequence culminating in some effect. Temporal asymmetries are characteristic of historical explanations in which an event is explained as the end-point of a chain of events. More generally, they are found wherever a temporal ordering of causes is central to an explanation. Thus, the conventional sociological account of occupational attainment, in which individuals' occupations are held to be directly caused by their education and social background, and in which social background is also a cause of education, involves temporal asymmetry claims.

As a simple statement of temporal ordering, temporal asymmetries

42. The independence of the two claims becomes clear if we shift the functional story from social development to living organisms. It is perfectly legitimate to produce a range of functional explanations for the interconnections of the parts of a living organism. This does not imply, however, that when a "contradiction" occurs between various structures and their functions that the structures change to restore functionality; the organism may simply die. When arteries become clogged so that the heart no longer functions to maintain the circulation of the blood, there is nothing that ensures a transformation of the arteries to allow for further survival of the organism. The primacy of the forces of production, therefore, rests on the dynamic claim that they have a tendency to develop in history (the development thesis) joined with a (weakly defended) claim that this tendency is causally more potent than the reproductive tendencies of superstructures on the relations of production.

43. In one sense all causal arguments are "diachronic" in so far as effects temporally follow causes. To speak about a mechanism generating an effect is to talk about a change that occurs in time. We use "synchronic" to describe facts about the relation among simultaneous causes and "diachronic" to describe facts about sequences of causes or sequences of effects of the same cause.

are pervasive and unproblematic in social scientific explanations. What is illegitimate is the identification of such orderings with claims for causal primacy. Thus, in accounts of occupational attainment, it would be a mistake to conclude from the fact that social background is a determinant of both education and current occupation, that it is a more important or more fundamental determinant of current occupation than is education.<sup>44</sup> Earlier causes are not *ipso facto* more important than later ones; they may also be weaker than the causes that follow them temporally.

Perhaps the most frequently encountered argument that moves from temporal ordering to a claim about causal primacy occurs in analyses that assign explanatory importance to the "origins" of particular processes. Sometimes it is held that in a chain of events  $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4$  (where the subscripts are temporally ordered), the first cause is inherently more important simply by virtue of being first. But since, in social science settings, it almost never happens that the links in causal chains are deterministic, there is no reason why earlier causes should be viewed as more important than later causes in either the distribution-dependent or the causal potency sense.<sup>45</sup>

The temptation to regard earlier events as more important than later ones in historical explanations derives from a failure to consider *what might have happened* given the empirical facts present at each stage. Take the simple sequence  $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4$ , and suppose that in the

44. If social background were not just a determinant of education, but *completely* determined education, then it could be argued that it is indeed a more important cause of occupation than education is. It is worth noting, however, that *even if* X completely determines Y, and Y completely determines Z, it is still not assured that X is a *more important* cause of Z than Y is. By transitivity, X does completely determine Z; but Y also completely determines Z. The fact that X determines Z via Y is, of course, crucial for the explanation, but Y is just as important a fact in this causal process as X is.

45. Cf. Richard F. Miller, *Fact and Method: Explanation, Confirmation and Reality in the Natural and Social Sciences* (Princeton, NJ: Princeton University Press, 1987, pp. 98ff). Miller appears to suppose a highly deterministic relation within temporally ordered sets of causes when he distinguishes shallow causes from deeper causes in his analysis of causal depth-as-priority. Miller examines examples in which an immediate, proximate cause of an event—for example, the role of middle-class mobilization in the rise of the Nazis—is viewed as "shallow" if it can be demonstrated that some other cause—in this case, the power and interests of the bourgeoisie in Germany—is, in Miller's words, both "intimately" connected with the outcome and explains the proximate cause itself. The restriction of "intimacy" of connection with the explanandum is introduced to avoid silly infinite regresses, in which distant causes are always given priority over proximate ones. Intimacy, however, really seems to be a criterion of the degree of determination involved in the causal processes under consideration. If a temporally distant cause determined an outcome with a very high level of probability, then it, too, would count as the "deeper" explanation of the outcome than the proximate causes. In any case, such a degree of determinism in a sequence of events is virtually never present in the social sciences. Thus, in contrast with Miller, we believe that the translation of temporal priority into explanatory priority requires an independent argument.

world  $X_1$ ,  $X_2$ ,  $X_3$ , and  $X_4$  actually did occur. Let us also suppose that once  $X_2$  occurs,  $X_4$  has a very high probability of occurring. It could still be the case that given  $X_1$ , the most likely future would be something other than  $X_2 \rightarrow X_3 \rightarrow X_4$ . But  $X_2$  actually occurred. In such a sequence it would be wrong to claim that of all the events in the chain,  $X_1$  had the greatest causal importance. Because many historians refuse to consider counterfactual trajectories of events, they therefore treat the "origins" of *de facto* sequences as most important by default.

In some historical explanations it may be plausible to assign causal primacy to causes that can be identified as "origins" of some subsequent trajectory. For example, the particular conditions and conflicts that lead to the founding of a state are sometimes thought to be more important than any subsequent events or conditions in explaining current institutional arrangements. Foundational struggles generate constitutions and other institutional norms and practices that regulate basic political structures and procedures for as long as the state exists. We might say that what goes on at the historical origin of states generates a set of restrictive structural limits within which subsequent causal processes operate as selections. We have already argued that, in general, it is impossible to establish that limits are more important than selections. Nevertheless, when the explanatory objective is to account for the exclusion of specific kinds of historical alternatives, it may well be possible to identify initial processes of institutionalization as "decisive moments" within which important exclusions are created.<sup>46</sup>

Many historical analyses involve "path dependent explanations"—explanations in which, in the extreme case, there is a single path to some result.<sup>47</sup> In such explanations, there is a well-defined meaning to the notion of an "origin": it is that point at which the historical trajectory is set in motion.<sup>48</sup> In as much as getting on the path is a necessary (though

46. Arthur Stinchcombe's well-known analysis of the importance of the historical timing of the founding of particular industrial sectors in explaining their current organizational form would be an example. When an industry first becomes consolidated, Stinchcombe argues, it adopts particular institutional forms from the forms historically available at the time. Once these are firmly in place, they are exceedingly difficult to change—i.e. they exclude other possibilities effectively—even if they become suboptimal over time. See Arthur L. Stinchcombe, "Social Structure and Organizations", in James March, ed., *Handbook of Organizations* (Chicago: Rand McNally, 1965), pp. 142–93.

47. Path-dependent explanations are particularly striking when there is a unique path to some outcome, but more generally such arguments require only that there be switchpoints in which certain "destinations" are ruled out and others ruled in. There may still be more than one route to a given destination.

48. Historical arguments need not encounter an infinite regress of causes, always tracing the explanation of an historical trajectory back to an earlier determinant. Certain steps in an historical chain of causes and effects definitively close off some paths of development and open up others. Such "switchpoints" can be treated as "origins" of subsequent trajectories.

perhaps not sufficient) condition for producing the outcome, there is some justification for treating origins as particularly important.

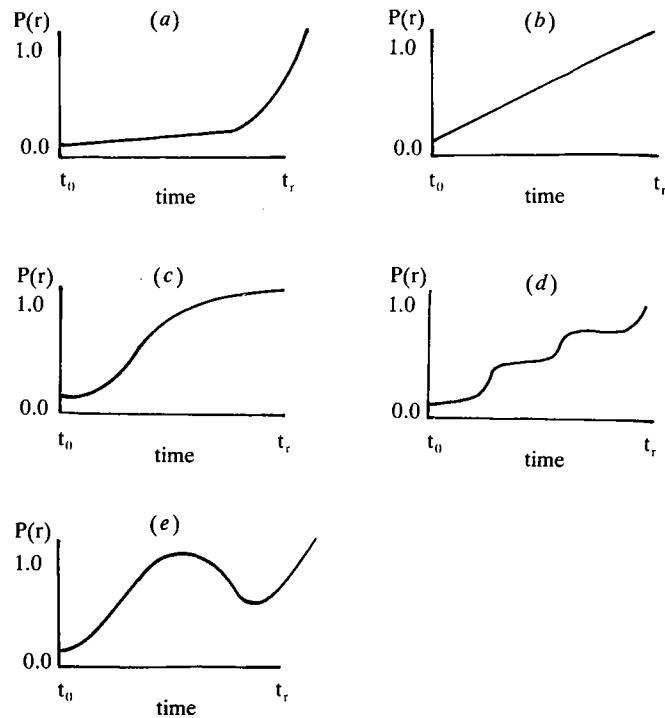
However, temporal asymmetry arguments, even when they refer to singular causal chains, do not imply that the origins of trajectories are more important than the causes that follow them. Nor do they imply that the distant past is more important than the recent past in explaining particular phenomena.

Suppose that we are trying to explain the occurrence of a revolution. Figure 7.6 illustrates a variety of temporal patterns of causes that could lead to the revolution. The vertical axis represents the probability that the revolution will occur. Since the revolution actually did occur, at the moment it began,  $t_r$ , the probability was 1. The question then becomes how this probability developed historically. Let  $t_0$  be some point in the past where we begin our investigation. The probability at  $t_0$  is the probability that a revolution will occur at time  $t_r$ , given the existing social conditions at time  $t_0$ .<sup>49</sup> Of course, it is extremely unlikely that precise values can be assigned to these evolving probabilities. Nevertheless, patterns of the sort portrayed in this figure are implicit in many historical explanations.

In Figure 7.6(a) the probability of a revolution remained very low until just before the actual revolution occurred. This situation might exist when a series of events, without deep social structural roots, came together in just the right temporal order—e.g. an unexpectedly prolonged stalemate in a war combined with the return from exile of a brilliant and charismatic leader. In Figure 7.6(b), there is no identifiable episode in which the probability of an eventual revolution rapidly increases. In this case, no cause or cluster of causes can be assigned causal primacy. Although models (a) and (b) are implicit in the work of atheoretical historians who are hostile to sociological theory, both none the less embody theories of how the process in question developed. To assign later events heavier weights, or to assign events equal weight, requires just as much a theoretical understanding of causal processes as assigning heavier weights to earlier events.

Figure 7.6(c) represents a theoretical stance opposed to 7.6(a): early

49. The precise shape of these temporally ordered probability curves will depend upon how the *explanandum* is defined. The *explanandum* could be the occurrence of the exact revolution that actually occurred (e.g. the American Revolution beginning on July 4, 1776), or it could be the occurrence of a revolution of a given type in a given country in a given period. If the event to be explained is defined in a fine-grained way, then proximate causes often will affect the probability of the event to a much greater extent than if the event is described in a coarse-grained way. Prior to early 1776 the probability of a revolution being launched on July 4, 1776 was undoubtedly far below 1, though the probability of a liberal bourgeois revolution of national independence occurring within the next few years might have already been very close to 1.

**Figure 7.6** Historicist Models of the Probabilities of Revolution

$P(r)$  = Probability of a revolution at time  $t$ ,  
 $t_0$  = time zero, starting point for the analysis  
 $t_r$  = time at which the revolution actually occurs

on, a set of conditions was created that rapidly increased the probability of a revolution. There was then an extended period in which the country was "ripe" for revolutionary upheaval; all that was needed was some spark to trigger the event. Figure 7.6(d) represents a trajectory in which there were a number of episodes in which the probability of revolution rapidly increased. In this picture, the middle episode is the "most important" in the sense of increasing the probabilities to the greatest extent (from 0.2 to 0.8). Figure 7.6(e) suggests a more complex historical pattern: after an early period in which the conditions for revolution rapidly increased there was a period in which those probabilities declined. Perhaps political reforms lowered the probability; then a military *coup* launched by a backward-looking ruling class blocked further reforms, restoring the earlier trajectory.

It is clear from these diagrams that the translation of temporal asymmetries in historical explanations into causal primacy claims need not be limited to the explanation of events or phenomena that actually occurred. One could try to explain the likelihoods of revolutions in different countries in the past where revolutions did *not* occur. Or one could equally construct explanations of this form in which the explanandum is the probability of a revolution (or anything else) in the future. Different societies have different probabilities today of having revolutions by the year 2000. These probabilities have evolved through different historical trajectories. An historical analysis of these trajectories should aim to identify the links in the chain in which these probabilities changed most dramatically.<sup>50</sup>

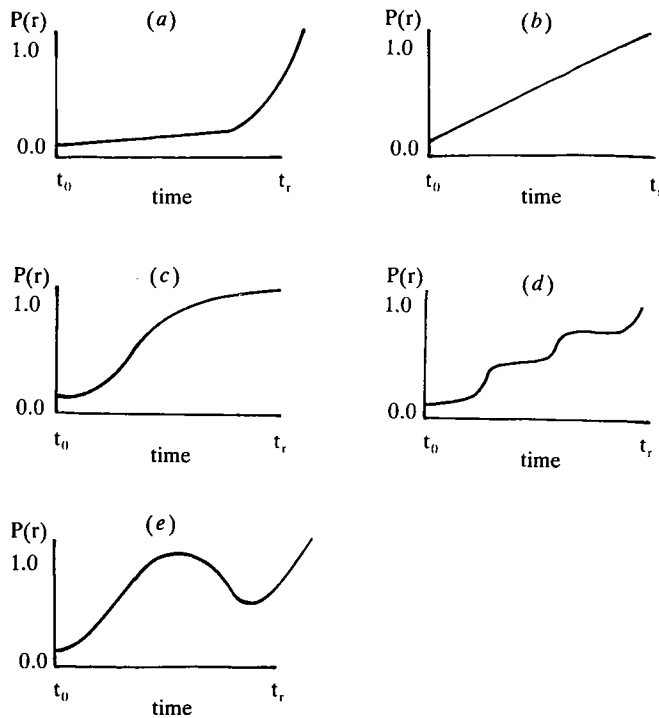
Temporal patterns of probabilities such as those illustrated in Figure 7.6 would have to be derived from substantive arguments about the mechanisms that actually affect probabilities. To argue that the probability of a revolution increased rapidly in the middle of the curve in Figure 7.6(c) requires a theory of social conditions and events conducive to revolutionary change, in addition to the observation that these conditions existed in this period. Such arguments, in turn, depend upon a developed understanding of the relative potency of the causes that enter the trajectory.

### Dynamic Asymmetry

As already noted, systemic causes may be distinguished from merely contingent ones. Dynamic asymmetries are systemic in the sense that they operate within particular systems, pushing these systems along a trajectory of social change. Dynamic causes are therefore to be distinguished both from endogenous causes that do not impart any developmental tendency, and also from exogenous causes. In the historical explanations modeled in Figure 7.6, the trajectories of event-probabilities were characterized as consequences of sequences of causal chains. But there was no supposition of any general mechanism that governed the overall sequence itself. The argument was not, therefore, a dynamic argument in our sense. The trajectories depicted were simply consequences of contingent occurrences. For dynamic asymmetries to exist, there must be some underlying mechanism generating a trajectory of development.

In the history of the social sciences, there have been many attempts at

50. Ellery Eells (*Probabilistic Causality*, Cambridge: Cambridge University Press, 1991) uses the trajectory of an event's probability to characterize whether the event occurred *because of* or *in spite of* some putative cause.

**Figure 7.6** Historicist Models of the Probabilities of Revolution

$P(r)$  = Probability of a revolution at time  $t$ ,  
 $t_0$  = time zero, starting point for the analysis  
 $t_r$  = time at which the revolution actually occurs

on, a set of conditions was created that rapidly increased the probability of a revolution. There was then an extended period in which the country was "ripe" for revolutionary upheaval; all that was needed was some spark to trigger the event. Figure 7.6(d) represents a trajectory in which there were a number of episodes in which the probability of revolution rapidly increased. In this picture, the middle episode is the "most important" in the sense of increasing the probabilities to the greatest extent (from 0.2 to 0.8). Figure 7.6(e) suggests a more complex historical pattern: after an early period in which the conditions for revolution rapidly increased there was a period in which those probabilities declined. Perhaps political reforms lowered the probability; then a military *coup* launched by a backward-looking ruling class blocked further reforms, restoring the earlier trajectory.

It is clear from these diagrams that the translation of temporal asymmetries in historical explanations into causal primacy claims need not be limited to the explanation of events or phenomena that actually occurred. One could try to explain the likelihoods of revolutions in different countries in the past where revolutions did *not* occur. Or one could equally construct explanations of this form in which the explanandum is the probability of a revolution (or anything else) in the future. Different societies have different probabilities today of having revolutions by the year 2000. These probabilities have evolved through different historical trajectories. An historical analysis of these trajectories should aim to identify the links in the chain in which these probabilities changed most dramatically.<sup>50</sup>

Temporal patterns of probabilities such as those illustrated in Figure 7.6 would have to be derived from substantive arguments about the mechanisms that actually affect probabilities. To argue that the probability of a revolution increased rapidly in the middle of the curve in Figure 7.6(c) requires a theory of social conditions and events conducive to revolutionary change, in addition to the observation that these conditions existed in this period. Such arguments, in turn, depend upon a developed understanding of the relative potency of the causes that enter the trajectory.

### Dynamic Asymmetry

As already noted, systemic causes may be distinguished from merely contingent ones. Dynamic asymmetries are systemic in the sense that they operate within particular systems, pushing these systems along a trajectory of social change. Dynamic causes are therefore to be distinguished both from endogenous causes that do not impart any developmental tendency, and also from exogenous causes. In the historical explanations modeled in Figure 7.6, the trajectories of event-probabilities were characterized as consequences of sequences of causal chains. But there was no supposition of any general mechanism that governed the overall sequence itself. The argument was not, therefore, a dynamic argument in our sense. The trajectories depicted were simply consequences of contingent occurrences. For dynamic asymmetries to exist, there must be some underlying mechanism generating a trajectory of development.

In the history of the social sciences, there have been many attempts at

50. Ellery Eells (*Probabilistic Causality*, Cambridge: Cambridge University Press, 1991) uses the trajectory of an event's probability to characterize whether the event occurred *because of* or *in spite of* some putative cause.

describing causal processes thought to impart a developmental trajectory to social phenomena: the "iron law of oligarchy" describes a set of mechanisms, embedded in the internal social system of organizations, that leads popular democratic institutions to develop centralized hierarchical institutional forms; the theory of structural differentiation elaborated by Parsons and others tries to explain the long historical pattern of increasing institutional specialization in distinct social "functions" in whole societies; and, perhaps most famously, historical materialism tries to theorize the overall trajectory of human history through an account of the dynamic relationship between forces and relations of production. In each of these cases, the trajectory of social change is not seen as the cumulative effect of *contingently* connected causes. It is determined by underlying, dynamic processes.

Claims for dynamic asymmetry generally occur in explanations that combine *dynamic endogenous* causes with *contingent exogenous* causes or in explanations that postulate a number of endogenous causes, not all implying the same dynamic trajectory. We shall examine dynamic asymmetries in each of these contexts, and then consider the relationship between dynamic asymmetry claims and causal primacy.

#### *Dynamic-systemic Causes vs. Contingent Causes*

Suppose we want to explain why animals die. One strategy might be to identify a range of contingent, proximate causes of individual deaths—diseases, accidents, predators, etc. Another approach would be to examine the genetic determinants, if any, of mortality. In so far as animals are "programmed to die", it might be thought that genetic determinants are endogenous dynamic causes of death, even if in a given population, malnutrition causes most actual deaths.<sup>51</sup> In such a case, malnutrition would be the primary cause of death (in the distribution-dependent sense), yet there would nevertheless be genetic mechanisms, dynamic causes, that make death inevitable.

Traditional historical materialist arguments contain a variety of claims that combine dynamic and contingent causes. Consider Marx's account of the systemic contradictions that are supposed to ensure the eventual demise of capitalism. In Marx's view, the process of capital accumulation contains a fundamental and devastating contradiction: each individual capitalist, in seeking to maximize profits, makes innovations that substitute capital for labor. The cumulative effect of these individuals'

51. The claim that lifespans are positively programmed genetically is controversial. However, even if lifespans are the result of physiological deterioration due to the failure of the genetic program to repair the organism adequately, rather than the result of a self-destruct mechanism, the causes of aging and death would still be endogenous.

choices raises "the organic composition of capital", the ratio of physical capital to living labor. Since, in his view, only direct, living labor produces value, and since profits derive from *surplus* value, the rising organic composition of capital has the effect, other things being equal, of lowering the average *rate* of profit over time. Given that the rate of profit of actual firms is distributed around this average rate, as the average declines, an increasing number of firms will experience net losses. This situation will lead, in turn, to increasing bankruptcies and eventually to a general economic crisis. Since in such crises capitalists can obtain capital cheaply (by buying up capital equipment from bankrupt capitalists), the conditions of profitability tend to be restored during the trough of a crisis, leading to a renewed period of profitable accumulation. However, because of the long-term tendency for the organic composition of capital to rise, the peak rate of profit in the new cycle will be lower than in previous ones. In the long term, therefore, economic recoveries will be less robust, with weaker booms and more prolonged busts.

If this scenario were correct, eventually capitalism would become an unreproducible social system. In time, the average rate of profit would be so low that capitalism would collapse. Marx, however, did not believe that capitalism would, in fact, last long enough for this final collapse to occur. The popular masses would topple the system before it reached that point. The declining rate of profit—or, more strictly, the deteriorating economic prospects it produces—helps explain why masses of people come to oppose capitalism and to favor socialism—before capitalism becomes unable to reproduce itself for strictly economic reasons. In different countries, for a variety of *contingent* political, cultural and historical reasons, the necessary ideological and political conditions for the overthrow of capitalism might be easier or harder to achieve. But, in the classical Marxist view, the contradictions of capitalism are so profound that eventually workers should be able to transform the system, even under relatively disadvantageous conditions. Like genetically driven aging culminating in death, capitalism is a social system with a built-in self-destruction mechanism. It might therefore seem that this mechanism most fundamentally explains capitalism's inevitable decline and fall, even if the explanation for the demise of actual capitalisms depends more on contingent political and cultural factors.<sup>52</sup>

A partition of causes into systemic and contingent is, however, highly

52. Needless to say, few Marxists today would accept this strong dynamic asymmetry argument. Few now believe that the contradictions of capitalism will eventually render capitalism unreproducible or socialism inevitable.

dependent upon how the "system" in which a cause is embedded is described. For example, in the case of death, one could define the system ecologically rather than organismically. The system in which an individual death is explained might then include disease agents and nutrition sources, as well as the genetic endowments of individuals. Within this larger system, diseases and nutrition would not be contingent mechanisms, but fully endogenous ones. It is only when the system or "unit of analysis" is defined as the organism that these causes become contingent and exogenous. The systemic/contingent distinction is thus always relative to descriptions of system-properties.

This fact seems to suggest that the claim that dynamic asymmetries are systemic is essentially arbitrary. This conclusion would indeed follow if there were no criteria for assessing the adequacy of claims that particular sets of causes constitute "systems"—in other words, if there were no sustainable theories of systemically interconnected social processes. Theories show what comprises a system and what does not. To say that A and B are parts of a system, and that C is not, is to say that A and B are causally interconnected in a systematic and reproducible manner, and that C is not connected with A and B in this way. In so far as social scientists succeed in their explanatory objectives, positing and defending a system-description of a set of causes is not arbitrary. It is a theoretical achievement. In the final analysis, progress in social science grounds the systemic/contingent distinction.

One of the long-standing debates in the social sciences involves the question of the extent to which societies can be analyzed as social systems. In so far as social structural configurations are reproduced over time through a set of interconnected causes, and in so far as changes in one institutional site of a society have ramifications for other institutional arrangements, it seems clear that societies are indeed causally integrated systems, albeit of a relatively loosely coupled and open variety.<sup>53</sup> What is more problematic is whether societies should be treated as functionally integrated systems. In any event, what matters in the present context is that unless some degree of systematicity is posited for interconnections among causal mechanisms in social scientific

53. Characterizing social systems as "loosely coupled" implies that *systemness* should be viewed as a variable property of causal problems rather than as an all-or-nothing property. While it may be true that "everything is connected to everything else", the explanatory ramifications of such interconnectedness can vary from a situation in which a change in one element generates commensurate changes in all other elements of a system to a situation in which changes in one element have negligible effects on other elements. A parallel point about the concept of individuality in biology is developed by E. Sober, in "Organisms, Individuals, and Units of Selection", in F. Tauber, ed., *Organism and the Origin of Self* (Princeton, NJ: Princeton University Press, forthcoming).

explanations, it is impossible to distinguish contingent from systemic causes—and therefore dynamic endogenous causes from exogenous ones.

Many social scientists are skeptical about the existence of dynamic mechanisms. Anthony Giddens, for example, has insisted that history has no overall determinate trajectory of development—that there is nothing equivalent to an organism's genetic mechanism that explains a society's trajectory. According to Giddens, social transformations can only be retroactively explained by the actual sequence of causal processes that happened to be temporally juxtaposed.<sup>54</sup> Since history has no systematic explanation, it does not make sense to talk about asymmetries between endogenous dynamic factors and exogenous contingent ones. But even if one rejects the possibility of a dynamic account of the overall trajectory of human history, it still is possible that more limited dynamic arguments of this sort can be formulated with respect to specific historical epochs or specific institutions. One could argue, as many Marxists do today, that historical materialism as a general theory of history is unsustainable, but that the Marxist theory of capitalist development is essentially sound.<sup>55</sup>

#### *Dynamic-systemic Causes vs. Structural-systemic Causes*

The interaction of dynamic- and non-dynamic-systemic causes plays an important role in a number of on-going discussions among radical social theorists—particularly debates about the state and about gender.

First, consider the state. If we look *statically* at a capitalist society and examine the social structural relationship between its political and economic institutions, it is hard to justify the Marxist idea that the state is "superstructural" on its economic "base" or even to identify any inherent asymmetry between them—structural, contextual or even functional. It is true that the capitalist structure of the economy imposes limits on what states can do, particularly since the state is dependent upon capitalist production for its revenues. But it is equally clear, given the state's role in superintending capitalist economies, that the institutional structure and policies of the state impose limits on what capitalist firms can do. The economy imposes structural limits on the state, but the state imposes structural limits on the economy. One might argue that the limits of capital on state policies are more powerful than the limits of the state on capitalist practices. But, as we have seen, it is unlikely that this intuition can be made sufficiently precise to support a genuine causal primacy claim.

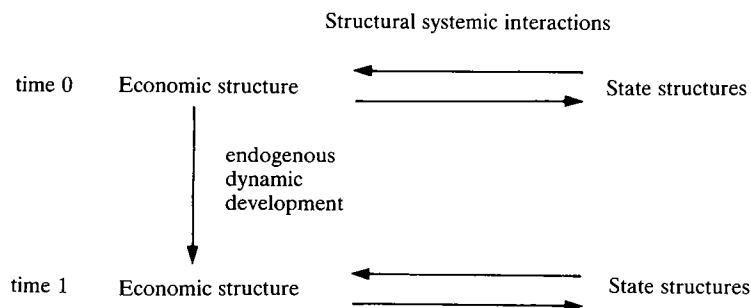
54. See Chapter 4.

55. Cf. Chapter 5.

However, there is a case for a causal asymmetry when dynamic and synchronic factors interact. Suppose that the nature and condition of the economy limit state policies, while state policies limit economic forms and initiatives. Suppose too, as Marxists believe, that capitalist economies contain an engine of social change, rooted in capitalist exploitation, competition and accumulation that explains the central tendencies of capitalist development from merchant capitalism to competitive industrial capitalism to multinational global capitalism. Suppose, finally, that there is no systematic engine of change internal to state institutions. While the state grows and develops in particular ways, its course is driven by causes external to it. If this characterization of the dynamic forces operating between the state and economy is correct, then even if it were true that state institutions and economic institutions constrain each other in a systemic way (thereby constituting two interconnected elements in a social system), there would still be a dynamic asymmetry: the trajectory of development of the state and the economy would be driven by dynamic causes operating in the economy, but not by dynamic causes endogenous to the state. This causal pattern is illustrated in Figure 7.7.

A similar explanatory structure pertains to the relation between class and gender. Synchronically, class relations and gender relations each impose limits on many social practices. While in specific cases it may be possible to argue that one or the other is in some sense "more important", it is unlikely that there is a plausible general case to be made for the explanatory primacy of class or gender in purely synchronic terms. But dynamically the situation is different. There appears to be nothing analogous to capital accumulation or the "dialectic of forces and relations of production" that pushes gender relations along a determinate path of development. Attempts at identifying such processes, for example Mary O'Brien's analysis of the dialectic of forces and relations

Figure 7.7 Interaction of Dynamic and Structural Systemic Causes



of biological reproduction, have not so far succeeded.<sup>56</sup> However, if even weak historical materialism is right, class does have an endogenous dynamic. In this sense, even if we cannot say in general that class is more important than gender or vice versa with respect to some range of explananda, we can identify a dynamic asymmetry between them.

Two conceivable rejoinders to these conclusions should be noted. First, atheoretical historians and social scientists would deny that class relations have any endogenous dynamic even under capitalism. They would maintain instead that gross changes are just cumulative effect of finer-grained transformations, with no mechanism governing the overall trajectory.<sup>57</sup> Other social scientists might argue, instead, that what appears to be a class dynamic is, in fact, something else: perhaps a technological dynamic. They might then claim that technological change drives the system independently of class, state or gender relations. These factors could, of course, still have synchronic effects on technology; but the dynamic properties of the system would rest on a technological imperative.

It is also conceivable that, in addition to a class dynamic, there are dynamic mechanisms that govern forms of gender relations or forms of the state. As already remarked, to date arguments of this sort have not been very persuasive in the gender case. But the situation is perhaps more promising for the state. Much has been made, in recent years, of the fact that states are embedded in competitive geopolitical state systems. Competition in these systems could function like capitalist market competition as a mechanism driving states to accumulate resources and power. Perhaps this is what Giddens has in mind when he maintains that states accumulate "authoritative resources" in a trajectory of increasing "space-time distancing".<sup>58</sup> It should also be noted that states are not exactly unitary entities. They are comprised of a myriad of agencies, bureaucracies, branches of government and parties, and also of individual politicians and bureaucrats. Competition and conflict among these corporate and individual actors could drive state institutions along a particular path of development. Thus efforts by individual bureaucracies and agencies to increase their budgets could push the state as a whole along a certain developmental trajectory. If arguments of this kind turn out to be sound, there might not be any dynamic asymmetry between class and the state after all.

56. See Mary O'Brien, *The Politics of Reproduction*.

57. See Chapter 4.

58. See Chapter 4. "Space-time distancing" refers to the geographical and temporal distances over which intentional action can be planned. See A. Giddens, *A Contemporary Critique of Historical Materialism* (Berkeley: University of California Press, 1981).

*Dynamic Asymmetry vs. Causal Primacy*

There is a temptation with dynamic asymmetries, as with the other qualitative asymmetries we have considered, to regard the asymmetry itself as a basis for a causal primacy claim. However, what we have found to be the case for other qualitative asymmetries holds here too: there is no reason to consider dynamic endogenous processes more important than contingent causes or synchronic systemic causes *simply because they are dynamic and endogenous*.

When it seems that dynamic causes are more important than other causes, it is often only because, consciously or not, the explanandum has been construed in a tendentious way. Consider our example of the death of animals. One might seek to explain why, for a given type of animal, a particular distribution of lifespans occurs. Alternatively, one might wish to explain why particular animals cannot continue to exist for longer than some specified time. Genetic upper bounds on the number of cell divisions might be the most important cause of the latter phenomenon, but not important at all in the former case. Even if we wanted to explain not lifespans but the overall trajectory of an animal's development—an explanandum for which genetic mechanisms are often very important—there would still be aspects of the trajectory that might be more heavily determined by contingent factors than by the genetic growth dynamic. *Variations* across human populations in the timing of the development of secondary sex characteristics, for example, may be more influenced by variations in nutrition levels than by variations in endogenous dynamic causes. The fact that the dynamic cause propels growth in all human beings does not imply that *variations* in the dynamic cause explain *variations* in growth across the human population. Similarly, the endogenous dynamic of capitalist economies might be the fundamental determinant of the broad contours of their trajectory of development and of the high probability of the eventual collapse of capitalism, without explaining very much about finer-grained aspects of capitalist development or the actual demise of any particular capitalist society.

Shift of explananda are particularly evident in discussions of the relation between class and gender. If we want to explain the massive expansion of women's labor force participation in the post-Second World War period in all capitalist countries, the most plausible explanation is likely to revolve around the dynamics of capitalist development in these societies and the ways in which this dynamic has dramatically changed labor force requirements. This dynamic, however, probably has little to do with explaining why women were the principal source of the untapped available labor supply, or why there are variations in the labor force participation rates of women across capitalist societies at roughly equal levels of economic development. Perhaps these variations are best

explained by differences in patriarchal ideologies linked to religious institutions and kinship structures. This suggestion is entirely compatible with the idea that the principal cause of the fact that women's labor force participation is expanding in advanced capitalist countries is a dynamic internal to the capitalist mode of production.

Even if class is very important in the explanation of many phenomena having to do with gender relations, it does not follow that it is important in the explanation of all explananda somehow connected with gender. The same conclusion holds, of course, for gender-based mechanisms. In general, the fact that some explanandum can be explained in a certain way does not imply anything about how closely related explananda are best explained.

### Conclusion

Three principal lessons may be drawn from this analysis of causal asymmetries.

First, causal primacy claims, if correct, should be recast as *quantitative* asymmetry claims. It is therefore unlikely, in most explanatory contexts, that causal primacy claims can be sustained with precision. In order to validate claims about quantitative asymmetries, one either must establish the relative importance of different causes within an empirical distribution of causes or else devise a strategy for comparing the potencies of causes. The latter task is especially difficult because of the incommensurability of the units in which causes are calibrated.

Second, it is illegitimate to infer explanatory importance directly from the fact of qualitative asymmetry: limits are not inherently more important than selections; functional asymmetries do not imply causal primacy; there is no general reason to regard earlier links in causal chains as more important than later ones; and endogenous dynamic causes are not in themselves more important than contingent or non-dynamic systemic causes. In every case, if one cause actually is more important than another, it is only because it is quantitatively more important, regardless of their qualitative relationship.

Finally, it is crucial that investigators have a clear sense of the explanatory problem under consideration before they attempt to resolve disputes about quantitative or qualitative causal asymmetries. Ambiguities in these matters afflict social scientific practice generally. In particular, we are convinced that many of the controversies among radical social theorists—including debates about the relative importance of class and gender, or class and the state—are confounded by shifts in explananda. If these disputes are to be successfully adjudicated—indeed if they



are genuine disputes at all—it is not enough that they address the same *topic*. Contending explanations must literally have the same explanandum.

Our conclusions suggest that for many of the most interesting problems in social science, it is unlikely that fine-grained assessments of the relative importance of different causes can be made. While it may be relatively easy to distinguish between important and minor causes, it will generally be very difficult to make nuanced judgments about the relative importance of causes within these broad categories. Thus, if both the institutional properties of the state and the nature of class relations systematically shape state policy formation, it is unlikely that either can be said to be more important *in general*. Assessments of relative importance *among important causes*, even if they can be made rigorously in particular instances, are likely to be so affected by the precise characterization of the explananda and by the range of variation allowed for different causes, that generalizations are likely to be vulnerable to small changes in the specification of the problem.

Nevertheless, many of the most durable debates in the social sciences—and many contemporary debates among radical social theorists—do revolve around problems of causal asymmetry and, especially, causal primacy. In traditional Marxism, at least in its Hegelian versions, claims for the global primacy of class were embedded in a conception of society as an integrated “totality”. If societies were tightly integrated systems, organized through some singular “essence” that stamped all the “parts” of the whole with a specific function and character, and if that essence were identified with class relations, then there indeed would be grounds for according class explanatory primacy *in general*. Class primacy within this sub-tradition of Marxism therefore amounted to a claim for class *reductionism*: whatever was explanatory in particular instances was ultimately a form of appearance of class. Virtually no Marxist today would accept this reductionist perspective. Societies are understood to contain a variety of irreducibly distinct causal mechanisms. While there are asymmetries among causes, including asymmetries that justify causal primacy claims, there is no principle that warrants the conclusion that class considerations always comprise the primary determinants of social phenomena.

Once the presumption that class is the singular determinant of the social totality is abandoned and a range of distinct, causally efficacious mechanisms is admitted, the sweeping, global claims to causal primacy characteristic of much of the Marxist tradition are unsustainable. We think it is very likely that class considerations are of great explanatory importance for many well-defined explananda. But we also believe that it is unlikely that this conclusion—or anything like it—could possibly hold for “society” in general.

It is wise, therefore, to shift discussion away from causal primacy to causal importance, and to focus on the systematic impact of a given set of causes. We shall suggest in Chapter 8 that the mechanisms identified in class analysis have considerable importance across a wide range of explanatory problems; and therefore that Marxist class analysis sustains a case for the causal *pervasiveness* of class, though not for its global primacy. Claims for causal pervasiveness, to be plausible, must be grounded in a specific explanatory agenda. We therefore must ask whether Marxism in fact has a coherent explanatory program and whether the mechanisms traditionally identified with Marxist explanations really do play an important explanatory role in Marxism's explanatory projects. To ask these questions is to investigate “the Marxist agenda”. This is the topic of the next chapter.