Annu. Rev. Sociol. 2006. 32:271–97
doi: 10.1146/annurev.soc.32.061604.123127

First published online as a Review in Advance on February 24, 2006;
corrected Review in Advance published on May 18, 2006

CUMULATIVE ADVANTAGE AS A MECHANISM FOR INEQUALITY: A Review of Theoretical and Empirical Developments

Thomas A. DiPrete and Gregory M. Eirich
Department of Sociology, Columbia University, New York, NY 10027;
email: tad61@columbia.edu, gme2101@columbia.edu

Key Words inequality, life course, careers

■ **Abstract** Although originally developed by R.K. Merton to explain advancement in scientific careers, cumulative advantage is a general mechanism for inequality across any temporal process (e.g., life course, family generations) in which a favorable relative position becomes a resource that produces further relative gains. This review shows that the term cumulative advantage has developed multiple meanings in the sociological literature. We distinguish between these alternative forms, discuss mechanisms that have been proposed in the literature that may produce cumulative advantage, and review the empirical literature in the areas of education, careers, and related life course processes.

INTRODUCTION

The concept of cumulative advantage (CA) appears frequently in the social scientific literature, including literature on social mobility, poverty, race, crime, education, and human development. Traceable to the work of Merton (Merton 1973a,b, Merton 1988), the Matthew effect was defined as "the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark" (Merton 1973a, p. 446). In the related process of accumulative or cumulative advantage, exceptional performance early in the career of a young scientist attracts new resources, as well as rewards that facilitate continued high performance (see Zuckerman 1988 for a recent review of Merton's contribution to the development of CA theory). Cole & Cole (1973) generalized these insights to produce a dynamic theory of stratification in scientific careers (see also Cole 1970, Zuckerman 1977). Allison and colleagues (1982) systematized aspects of Merton's theory of CA in a mathematically precise fashion and showed that the predictions of certain versions of CA theory were consistent with observed characteristics of scientific careers.

Social scientific literature frequently invokes CA mechanisms as an explanation for inequality. Life course scholars see CA working in patterns of diverging cohort trajectories (e.g., O'Rand 1996, Ross & Wu 1996, or Mayer et al. 1999; see O'Rand 2002 and Dannefer 2003 for recent literature reviews). The organizations, markets, and networks literature has analogous terms for CA, such as first-mover advantage (Kerin et al. 1992), path-dependent increasing returns (Arthur 1994), and Gibrat's law of proportional effect (Sutton 1997). Sociobiologists refer to multiplier effects (Wilson 1975). Sociologists, social psychologists, and economists invoke CA processes with concepts of reputational effects (Gould 2002), positional goods (Hirsch 1976), halo effects (Greenwald & Banaji 1995), the scarring effects of unemployment (Ellwood 1982), and cumulative discrimination (Blank et al. 2004). Colloquial terms—"virtuous cycles," "vicious cycles," "the rich get richer," "the poor get poorer," "success breeds success"—are also common in the literature.

The central descriptive idea in the CA literature is that the advantage of one individual or group over another grows (i.e., accumulates) over time, which is often taken to mean that the inequality of this advantage grows over time. The advantage in question is typically a key resource or reward in the stratification process, for example, cognitive development, career position, income, wealth, or health. The use of CA as a description for growing inequality is just that, another term for describing a pattern of growing inequality. CA becomes part of an explanation for growing inequality when current levels of accumulation have a direct causal relationship on future levels of accumulation. A CA process is capable of magnifying small differences over time and makes it difficult for an individual or group that is behind at a point in time in educational development, income, or other measures to catch up. Ironically, despite the obvious theoretical and policy importance of CA models, and despite widespread references to their existence in the literature, the sustained development and testing of CA models has been more the exception than the rule.

As the use of CA has proliferated, its meaning has broadened, and this review is intended to encompass the major ways in which CA has been used in the stratification literature, beginning with early developments and then moving forward in time. We have found two major forms, which we illustrate by example here and describe more formally in the next section. The first—and one might say "strict"—form can be identified with Merton's theory of scientific careers, whereas the second form derives from early research on the status attainment model.

The strict form, which is found in mathematical models for exponential growth, contagion, and diffusion, is illustrated most simply by the process of wealth accumulation through the mechanism of compound interest. When an initial sum of money is placed into an account in which the interest is reinvested, the accumulation of new wealth depends upon the size of the existing principal. If the interest rate is 5% and is compounded yearly, then a $1,000 account produces a $50 accumulation, whereas a $100,000 account produces $5,000. If the interest rate is constant, the numerical difference between the two unequal starting amounts grows exponentially over time, from $99,000 at the start to $126,000 after five years, to $161,000 after 10 years, and so forth. The yearly increments to these amounts also grow exponentially. However, the ratio of the fortunes remains constant (at 100:1) regardless of how long the process continues. If instead the interest rate depends on the size of the principal (e.g., 7% interest on \$100,000, but only 5% interest on \$1,000), then both the difference and the ratio between the two fortunes would grow over time. The exponential growth from the compounding of interest can be contrasted to the arithmetic growth that characterizes the distance covered by a runner in a race. If the runner moves at a constant speed, his total distance continues to increase, but the increment to his total distance in any time period depends only on the speed and duration of the period, not on the total accumulated distance. In a strict CA process, future accumulation depends upon current accumulation.

A second form of CA often found in the sociological literature can be traced to Blau & Duncan's (1967) *The American Occupational Structure*. Their analyses of the 1962 "Occupational Changes in a Generation" (OCG) data showed that African American males suffered from cumulative disadvantage relative to white males. Blau & Duncan's concept of cumulative disadvantage, which emphasized group differences rather than inequality within a group or within an entire population, is—as we show below—not necessarily strict CA in the Mertonian sense. Instead, cumulative advantage or disadvantage in the Blau-Duncan sense referred to persisting direct and interaction effects of a status variable, where the interaction effects implied group differences in the returns to socioeconomic resources. Being black was a cumulative disadvantage because race had both direct and indirect effects on outcomes at different stages in the life course and because highly educated blacks received lower status returns than did highly educated whites (an interaction effect referred to as perverse equality). The Blau-Duncan approach can be generalized to variables that are conceptualized as exposures to a treatment over some (possibly long) duration, such as growing up in a poor versus a rich family, growing up in a single-parent versus a two-parent family, growing up in a poor versus an affluent neighborhood, or being assigned to a low versus a high academic track in school.

An additional difference between what we have termed strict CA and the Blau-Duncan approach to CA concerns growth in inequality over time. Strict CA—as illustrated by the example of wealth accumulation—produces growing differences between individuals and growing mean differences between groups over time. However, as Allison et al. (1982) discussed, it does not produce growing inequality in scale-invariant measures of inequality such as the coefficient of variation, unless the rate of accumulation also varies within the population. The Blau-Duncan form of CA does not in fact always produce growing mean differences over time, and therefore it does not necessarily lead to growing inequality by measures such as the coefficient of variation.[1]

[1] Suppose that the outcomes are expressed in terms of standardized variables, that R stands for race, that Y_1 and Y_2 are status outcomes at two points in the life course, and that the Blau-Duncan CA process is expressed through a combination of persisting direct and indirect effects. In the simplest model, $Y_1 = \beta_{11} R + \epsilon_1$, and $Y_2 = \beta_{21} R + \beta_{22} Y_1 + \epsilon_2$. Whether standardized racial group differences grow between time 1 and 2 depends upon whether $\beta_{21} + \beta_{22}\beta_{11}$ is larger or smaller than β_{11}.

Both Merton's and Blau and Duncan's research on CA is somewhat dated, but their approaches continue to dominate this literature. At the same time, the distinction between strict CA and the Blau-Duncan form is insufficient to fully understand the important distinctions in existing models. The next section elaborates these models more formally, and the following sections deal with mechanisms and empirical examples.

FORMAL CONSIDERATIONS

Evidence for CA processes can take several forms. One class of evidence is distributional in character. CA-like processes produce right-skewed distributions on the outcome variable of interest. Lotka (1926) argued that scientific productivity follows a distribution that, in its generalized form (Huber 2002), is

$$p_n = p_1 \frac{1}{n^k},$$

where p_n is the proportion of scientists with n publications, p_1 is the number of scientists with one publication, and k is a constant. A related form is the Yule process (Simon 1955), which produces a distribution that can be approximated by the formula

$$p_n = \frac{1}{n(n+1)}.$$

The power-law function, which is also referred to as Zipf's law and as a Pareto distribution, is a continuous distribution that is closely related to the discrete Yule distribution

$$p(y) = Cy^{-\alpha},$$

where y is the outcome, $p(y)$ is the value of the probability density function at y, and alpha is the exponent of the power law. Price (1965, 1976) argued that the number of citations to scientific papers follows a Yule process and produces a power law, whereas Fox & Kochanowski (2004) found that commercial success of recording artists follows a generalized Lotka distribution, although the success distributions varied by gender and race. As Newman (2005) recently showed, the distribution of citations, the wealth of rich people, and the sales of books all follow power laws. Newman further points out that strict CA (which we formalize below) is one mechanism that produces a power-law distribution, but other mechanisms can also produce power-law distributions (see Mitzenmacher 2003 for a discussion of the related literature based on the lognormal distribution). It follows that aggregate distributional evidence is not sufficient proof that a CA mechanism is generating the outcomes in question. Furthermore, these functional forms typically describe the extreme values—i.e., the tail—of a distribution produced by a CA process than they do the less extreme values. Finally, distributional evidence explains little about the social process producing CA, and so additional evidence is desirable.

A second form of evidence, which is dynamic but still aggregate in character, is growing inequality in the outcome over time (Allison et al. 1982, Dannefer 1987, O'Rand 1996). However, like the first form, this type of evidence is also imperfect both because a variety of CA-like processes can produce growing inequality over time and because some forms of strict CA do not necessarily produce growing inequality over time (Allison et al. 1982).

A third form of evidence comes from microlevel dynamic data, which allow the estimation and testing of parametric models. Allison and colleagues (1982) proposed modeling the rate of accumulation of publications and citations as a contagious Poisson process. The contagious Poisson model is distinguished from the standard Poisson model by the assumption that the intensity parameter λ depends on prior successes. Allison and coworkers specified λ as

$$\lambda_{it} = \alpha_i + \gamma_i Y_{it-1}, \tag{1.}$$

where α_i describes baseline differences between individuals, $\gamma_i > 0$, and Y_{it} is the cumulated number of accomplishments of individual i at time t. The fact that the rate of future events is a positive function of previous events creates a CA process.

The discrete model of Allison et al. (1982) is similar to the well-known differential equation for exponential growth (Yule process), namely

$$\frac{dY_{it}}{dt} = \gamma Y_{it},$$

which implies that

$$Y_{it} = Y_{i0}e^{\gamma t}, \tag{2.}$$

and which in discrete time is expressible as a difference equation:

$$Y_{it} - Y_{i,t-1} = \gamma Y_{i,t-1}. \tag{3.}$$

As applied to the process of citations of scientific papers, Equation 3 states that the expected number of new citations to a paper is proportional to the current number of citations. In the simple form, Equation 2, the model unrealistically implies unbounded growth, but it can easily be modified, as was proposed in 1838 by Verhulst to produce the equation for logistic (i.e., S-curve) growth, namely

$$Y_{it} - Y_{i,t-1} = \gamma Y_{i,t-1}\left(1 - \frac{Y_{i,t-1}}{Y^*}\right), \tag{4.}$$

where for simplicity we assume equally spaced observations (thus obviating the need for a Δt factor on the right side of the equation). Verhulst's equation presumed that CA stops because Y approaches Y^*, the so-called carrying capacity of the system, as would occur in a process of population growth. Instead,

$$\left(1 - \frac{Y_{t-1}}{Y^*}\right)$$

could be replaced by

$$\left(1 - \frac{t}{T^*}\right),$$

which causes the CA process to slow as time t reaches some critical value T^*, as may occur in a life course process. For small period-to-period changes,

$$\frac{Y_{it} - Y_{i,t-1}}{Y_{i,t-1}} \approx \ln(Y_{it}) - \ln(Y_{i,t-1}). \qquad 5.$$

We can therefore propose a model for ln (Y_{it}) in which time dependence is expressed as a quadratic function of a variable z (which can stand for age, time, labor-market experience, or other factors), where the quadratic term creates an alternative "brake" to that provided by the S-curve formulation above. Adding a stochastic disturbance ϵ to this equation yields

$$\ln(Y_{it}) = \gamma_1 z_{it} + \gamma_2 z_{it}^2 + \beta' X_{it} + \epsilon_{it} \qquad 6.$$

or

$$\Delta \ln(Y_{it}) = (\gamma_1 + 2\gamma_2 z_{i,t-1})\Delta z_{it} + \beta' \Delta X_{it} + \Delta \epsilon_{it}.$$

In Equation 6, which has the form of a standard earnings function, cumulative growth in Y (i.e., constant proportional growth if γ_2 is zero) is driven by increases in z. The variables in X affect the level of ln (Y), and changes in X produce changes in Y, but neither X nor changes in X produces cumulative change in Y unless the variables in X themselves are accumulating over time. Although the error terms and changes in the X variables perturb the baseline level of Y, the cumulative growth is determined purely by the magnitudes of γ_1 and γ_2. In Equation 6, one person's advantage over another who has the same value of z can be attributed either to differences in current X, to differences in current ϵ, or to a higher return to z or X (i.e., different coefficient values). Equation 6 describes a process of CA, but this simple form is different from the more complex form of CA where one's present outcome level is directly related to past outcome levels even after current X and ϵ are controlled, for example,

$$Y_{it} = (1 + \gamma)Y_{i,t-1} + \beta' X_{it} + \epsilon_{it}. \qquad 7.$$

This equation expresses a form of path-dependent CA in which the entire history of the determinants of Y (including both the history of whatever variables are in X and the history of the random shocks in ϵ) affects the growth of Y. The CA process continues so long as $\gamma > 0$, which implies that the impact of prior events grows over time.[2] That is, exogenous chance events have long-term consequences,

[2]To clarify, suppose that there are no explicit X variables, so Y is purely a function of the exogenous shocks in ϵ. By rewriting Y at each time in terms of past values of Y, Equation 7 implies that $Y_0 = \epsilon_0$, $Y_1 = (1 + \gamma)\,\epsilon_0 + \epsilon_1$, $Y_2 = (1 + \gamma)^2\epsilon_0 + (1 + \gamma)\,\epsilon_1 + \epsilon_2$, and so

which is not true in the simple form of CA. The idea that exogenous chance events have long-term consequences is an essential characteristic of Merton's CA process and distinguishes it from the simple form of CA, where exogenous chance events have only short-term consequences and where individual quality differences are the major explanation for inequality.

Equation 7 suggests that lagged Y has a direct causal effect on current Y, but the true causal process in a path-dependent CA process is often more complex than this. In Merton's model, for example, previous scientific success does not directly cause current scientific success. Rather, he argued that scientific success increases one's level of resources, and it is the heightened level of resources that, in combination with scientific skill, leads to future productivity. If R is the resources made available on the basis of prior success, and if these resources cause increased productivity, then the true model takes the following form:

$$Y_{it} - Y_{i,t-1} = \alpha R_{it} + \beta' X_{it} + v_{it} \tag{8.}$$

where

$$R_{it} = \delta Y_{i,t-1} + w_{it}, \tag{9.}$$

so that

$$\begin{aligned} Y_{it} - Y_{it-1} &= \alpha(\delta Y_{i,t-1} + w_{it}) + \beta' X_{it} + v_{it} \\ &= \alpha\delta Y_{i,t-1} + \beta' X_{it} + \alpha w_{it} + v_{it} \\ &= \gamma Y_{i,t-1} + \beta' X_{it} + \epsilon_{it}, \end{aligned}$$

where $\alpha\delta = \gamma$ and $\alpha\, w_{\mathrm{it}} + v_{\mathrm{it}} = \epsilon_{\mathrm{it}}$, and where v and w are stochastic disturbances. It is important to note that Equation 8 is the same model as Equation 7 but has been elaborated to show the mechanism that drives the CA process. Such feedback loops can operate between wealth and income, depression and income, or other sets of reward and resource variables.

Finally, Equation 7 can be further generalized to allow for direct effects of the entire history of shocks, for example,

$$Y_{it} = \sum_{j=0}^{t-1} \gamma_{tj} Y_{ij} + \beta' X_{it} + \epsilon_{it}, \tag{10.}$$

where the magnitude of each γ coefficient depends both on the specific point in history of the outcome that it weights and on the gap between that point in history and current time. The pattern of these γ coefficients determines the importance of past outcomes on current outcomes. It allows, for example, for spectacular performance early in one's career (which causes an individual to be labeled a rising

forth. This expression shows that the entire history of stochastic shocks affects the current value of Y (and if X were in the equation, the entire history of X would also affect the current value of Y). When $\gamma < 0$, the effects of the past shocks die off into the future. But when $\gamma > 0$, the effects of past shocks increase as time progresses.

star) to produce continuing favored access to resources, even when performance in subsequent years is only average. This model of full path dependence can be further elaborated as in Equation 8, so that the feedback goes through resources or some other variable.

Faia (1975) and Allison et al. (1982) noted that although inequality in the population rose with larger γ in Equation 1, the level of inequality in the population (measured as the coefficient of variation, i.e., the standard deviation divided by the mean) did not grow over time if γ took the same value for everyone. This theoretical result was inconsistent with existing evidence about the structure of scientific careers. Only when γ varied at the individual level did Equation 1 result in increasing inequality over time. Allison and colleagues argued that γ might vary across scientists because " 'good' papers are more likely to bring further advantage than 'bad' papers" (Allison et al. 1982, p. 622). Other sources of heterogeneity in γ are possibly derived from unchanging variables such as race or gender, as well as exposure to treatments such as living in a poor neighborhood, living in a single-parent family, being in a high or low academic track in school, or attaining a particular education level. Heterogeneity in γ could also derive from the distribution of relative performance or standing, for example, when the career gains from scientific accomplishments are magnified by the judgment that these accomplishments are in general superior to the accomplishments of other scientists. This possibility is an essential element of tournament models of career mobility (Rosenbaum 1979), which assert that career success is determined by relative performance in a series of promotion contests, rather than through returns to an absolute level of some resource (e.g., one's stock of human capital).

Heterogeneity of outcomes on the basis of status or position is a core issue in sociological research. The dependence of outcomes on status can take many forms, only some of which correspond to a strict CA process. With the status labeled as s, the most relevant forms for our purposes include

$$Y_{it} \text{ or } \ln(Y_{it}) = \gamma_1 z_{it} + \gamma_2 z_{it}^2 + \beta_1' X_{it} + s_{it}\beta_2' X_{it} + \epsilon_{it} \qquad 11.$$

$$\ln(Y_{it}) = \gamma_1 z_{it} + \gamma_2 z_{it}^2 + \beta' X_{it} + \delta s_{it} + s_{it}\left(\gamma_3 z_{it} + \gamma_4 z_{it}^2\right) + \epsilon_{it} \qquad 12.$$

$$= (\gamma_1 + \gamma_3 s_{it}) z_{it} + (\gamma_2 + \gamma_4 s_{it}) z_{it}^2 + \beta' X_{it} + \delta s_{it} + \epsilon_{it} \qquad 13.$$

$$Y_{it} = (\gamma_1 + \gamma_3 s_{it}) z_{it} + (\gamma_2 + \gamma_4 s_{it}) z_{it}^2 + \beta' X_{it} + \delta s_{it} + \epsilon_{it} \qquad 14.$$

$$Y_{it} - Y_{i,t-1} = Y_{i,t-1}\gamma + \beta' X_{i,t-1} + \delta s_{it} + \epsilon_{it} \qquad 15.$$

$$Y_{it} - Y_{i,t-1} = Y_{i,t-1}(\gamma + \delta s_{it}) + \beta' X_{i,t-1} + \epsilon_{it}. \qquad 16.$$

In Equation 11, which we term a status-resource interaction model, the status affects the level of Y but not its growth rate. However, the effect on Y of other resource variables in X (such as education) depends on s. In the simple CA process of Equation 12, the rate of CA is governed by the coefficients of z and z^2, but the coefficients of these variables depend on s. This is one form of status-dependent

CA. In Equations 14 and 15, s affects the growth rate of Y, and its effect persists over time. However, the effect does not occur through a strict CA process because the effect does not depend on the current magnitude of Y. We might label the process by which one's status directly affects outcomes in Equations 14 and 15 as a cumulative exposure process because continuing exposure to the status or position has a continuing direct effect on the rate of arithmetic increase in Y (in addition to any indirect effect it may have through the CA process involving lagged Y as in Equation 15). It is important to note that a cumulative exposure process can cause mean group differences to rise even as the ratio of group means or other measures of inequality decline over time. In Equation 16, which we refer to as path- and status-dependent CA, the magnitude of the intensity parameter of the path-dependent CA process depends on the status.

As noted above, the Mertonian form of CA is what we have labeled path-dependent CA. Elaborations of this idea by Cole & Singer (1991) to account for gender differences in science are what we have labeled path- and status-dependent CA and amount to an explicit specification of the heterogeneity of the CA parameter discussed in Allison et al. (1982). In contrast, the cumulative exposure and the status-resource interaction models were the essential ideas in Blau and Duncan's characterization of the cumulative disadvantage of being black, which was revealed in the 1962 OCG.

The illustrative models above by no means exhaust the collection of models proposed in the literature in which outcomes are affected by the duration of exposure to some variable. The growing literature on "new causal analysis" conceptualizes causality in the experimental framework, where exposure to some treatment has a potential effect on an outcome. The accelerated-failure-time model describes the expected impact of continuous exposure to a treatment as a shift in the failure time under the counterfactual assumption of no exposure, and Robins et al. (1993) have recently developed the G-estimator for addressing the endogeneity problems that arise in getting consistent estimates of the effects of continuous exposure to the treatment on the outcome. In the accelerated-failure-time model, the arithmetic gap between the failure time of someone continuously exposed and an otherwise identical person who was not exposed grows with exposure time (while the ratio remains constant). Thus, the exposed individual could be said to experience cumulative advantage/disadvantage from the exposure in the same way that Blau and Duncan described race as a source of cumulative advantage/disadvantage.

The distinction between the mechanisms described above can be subtle, although quite important. Thus, race (represented by status s in the equations above) may have a continuing effect on wage growth throughout one's career because of discrimination (as in Equations 12, 14, and 15). The effect of race may also interact with education (as in Equation 11) because the effect of discrimination varies by level of education or because race affects the quality of schooling, which affects the returns of education. Race could also affect wealth accumulation, as in Equation 16, if housing equity grew at a different rate in white and black neighborhoods [however, empirical evidence suggests that the rate of accumulation is in fact similar by race (Scholz & Levine 2004)]. These mechanisms alone or in combination could produce growing inequality between whites and blacks over time. Together, they pick up one or more of the four main connotations of CA that exist in the sociological literature:

1. In a CA process, the growth rate in an outcome variable is a function of current values of that outcome.
2. In a CA process, small advantages or disadvantages at an early stage of a process grow larger over time.
3. When growth rates of an outcome variable vary by some status, such as race or gender, and these status-unequal growth rates persist over time or across multiple stages of the life course, the process is often described as CA (although it is not necessarily strict CA, as noted above).
4. Inequality grows over time as a consequence of a CA process. In strict CA, this always refers to growing population-level inequality. Many sociologists focus instead on growing inequality between groups over time, often without paying attention to the broader question of whether inequality in the population as a whole is growing.[3]

In the two forms of strict CA (simple CA and path-dependent CA), these four conditions are satisfied. It is important to note that not all uses of CA discussed above necessarily satisfy these conditions. Furthermore, some uses of CA lack sufficient formality to determine which of these conditions hold. The frequent lack of clarity in models, mechanisms, and tests is a continuing issue in the sociological literature on CA processes as potential generators of inequality. This lack of clarity can produce incorrect specifications, incorrect estimates, and incorrect interpretations. Lack of clarity can also arise from the failure to model CA when it in fact exists. For example, if group differences grow because of a small initial mean difference between groups that is widened through a group-neutral path-dependent CA process, an analyst who fails to look for this CA process may incorrectly conclude that group membership has a persisting effect on outcomes and that group membership is somehow magnifying group-level differences over time. This incorrect specification could then lead to the search for a continuing group-level mechanism that in fact does not exist.

The sociological literature often makes distinctions between CA and cumulative disadvantage. The literature on scientific careers is typically framed in terms of CA because it focuses on those who become stars in their field, whereas the literature on race inequality or on health is often framed in terms of cumulative

[3]Blau & Khan (1992) show that this omission can be problematic by demonstrating that rising inequality in a population generally produces rising between-group inequality as a by-product. Exclusive attention to group differences can cause a researcher to misidentify a population-level process as a group-specific process.

disadvantage because the focus is on those who experience disadvantaged outcomes. Although rhetorically useful, these alternative framings are mainly devices for indicating the direction of the effect (see also Zuckerman 1988). In the context of a CA process, one is at a relative disadvantage either if one's current stock of the outcome variable is comparatively small, if one's rate of growth in the outcome is comparatively low, or if one has experienced shocks that will adversely affect future outcomes. Similarly in the case of group-based CA, one might refer to the process as cumulative disadvantage if the status in question is generating negative outcomes relative to some other group. Cumulative (relative) disadvantage for some corresponds to cumulative (relative) advantage for others.

Estimation problems can impede distinguishing path-dependent CA from simple CA and distinguishing strict CA processes from other processes. Equations 7, 15, and 16 are referred to as dynamic equations in the econometrics literature, and consistent estimation of the parameters of such models is not straightforward because of the presence of endogenous variables (including specifically the lagged value of the outcome variable) as regressors. The standard advice in the methodology literature is to obtain consistent estimates of the model parameters via generalized method-of-moments estimators or some related instrumental variable regression strategy (Baltagi 2001), but the data requirements can be severe. Additional tests are available when the mechanisms producing the hypothesized CA are specified. For example, an advantage of Equation 8 over Equation 7 is that it provides additional tests for the hypothesized causal link between outcomes in one period and resources in the next period. The difficulty of adjudicating between the different models described above is further complicated by the fact that they are not nested, and so deciding among them on the basis of statistical tests is not straightforward. However, methods such as Bayesian model selection (Raftery 1995) may help to adjudicate between competing models. Before examining this point, however, the underlying model needs to be clearly specified.

MECHANISMS

The equations of the previous section may define what CA means, but they do not by themselves explain why it may occur. The most famous example within sociology of the strict CA mechanism comes from Merton's reputation-and-resources model of scientific careers. Merton started with three premises: (*a*) resources in the scientific world are limited, (*b*) scientific talent is difficult to observe directly, and (*c*) allocation of resources in science is governed by the norms of universalism (recognition should be granted on the basis of the quality of scientific work) and communism (resources should be allocated to maximize the overall productivity of the scientific community). Scientific resources are therefore not simply a reward for past productivity, but are given in order to stimulate future productivity. With limited ability to evaluate the great mass of ongoing scientific work, and with limited ability to measure future productivity beforehand, the scientific community favors those who have been most successful in the past, given their additional resources and attention. One consequence of this mechanism is that the gap in rewards between a more able and a less able scientist would grow over time. A second consequence is that chance events [e.g., unequal luck in the draw of reviewers upon submitting a grant proposal for funding (Cole et al. 1981)] would produce a relative advantage for one of two individuals of identical talent, and this relative advantage could persist and increase over time. A third consequence (Merton's so-called Matthew effect) is that scientists with greater reputations would gain greater rewards from work of the same quantity and quality of scientists with lesser reputations.

Rosen's (1981) still widely cited model of superstars is based on an entirely different principle, namely that small differences in intrinsic quality can imply large differences in market prices. Rosen argued that compensation for the highest levels of talent would increase disproportionately (to use his example, a top surgeon who is 10% better than the next most talented surgeon can command considerably more than a 10% premium for his exceptional talent). Rosen argued further that the oversize returns available to the most talented people give them a stronger incentive to produce high quantities of goods and services, and these unequal incentives further magnify the reward gaps at the top of the distribution. Rosen's mechanism is similar in one respect to Merton's in that relative (as opposed to absolute) performance is a central determinant of rewards. In another respect, however, Rosen's model is quite different from Merton's model. Rosen's model placed a large value on innate talent and no value on external resources and focused its attention on an economic rationale for why small differences in talent could lead through market mechanisms to large differences in rewards. Merton's model recognizes potential differences in talent, but additionally explains why zero or even negative differences in talent can lead to large positive differences in rewards because of the importance of external resources in the production of scientific output.

Some models express putative mechanisms behind growth in inequality relatively clearly but remain ambiguous about the role of CA in the process. Frank & Cook's (1995) theory of "winner take all markets" asserts that there are certain markets in which rewards are given not on the basis (or not simply on the basis) of one's absolute productivity, but rather on the basis of relative performance against a group of competitors. They further assert that an important component of the recent rise in inequality has been produced by the spread of the tournament system of compensation to a larger fraction of the labor market and by an increase in the size of the prizes in these career tournaments. However, either an innate-differences model such as that of Rosen or a path-dependent CA model such as that of Merton could produce the outcomes in question.

Cole & Cole (1973) and later Cole & Singer (1991) argued that CA operated partly through social psychological, cultural, and structural mechanisms. These mechanisms could produce a motivation to work harder if one is successful and a motivation to work less hard if one experiences rejection. In their "theory of limited differences," Cole & Singer show how status group inequality could result either through group differences in the distribution of the numerous positive or negative career "kicks" experienced in the early and middle part of their career or through group differences in the distribution of responses to these kicks. Rosen's version of this reinforcement process is based on rational responses to the market value of one's time, which overlaps with Cole & Cole's proposed mechanisms in some respects, but not in others.

Both Merton's path-dependent CA model and Rosen's non-CA model should be distinguished from the simple form of CA (Equation 6), which was most clearly expressed in Mincer's (now standard) human capital earnings function (Mincer 1974) and which is still considered seminal within labor economics (Heckman et al. 2003). Mincer argued that workers gain valuable experience during their careers and that this experience adds to the stock of their human capital. The mechanism producing growth in wages is the combination of growing resources (in Mincer's case, growing human capital) and a positive rate of return on this investment. Divergence in Merton's model can occur through differences in endowments, through chance differences early in one's career, or through any other factors that produce differences in the quality of early career performance (e.g., the quality of mentors or patrons). Divergence in Rosen's model would occur through inequality in endowments. Divergence in Mincer's model would occur through divergent rates of investment in human capital during the work career or unequal rates of return to resources. Mincer's model becomes more similar to Rosen's model if the unequal rates of returns to resources are driven by relative differences in talent, and becomes similar to Merton's model if the trajectory of resource accumulation is driven by prior performance in addition to prior investment.

Other more complex investment processes than described in Mincer's human capital earnings function also produce CA. Pension plans are prime examples of social institutions that—through a complex set of rules that vary both within and across countries—produce heterogeneous levels of investment, rates of accumulation, and levels of wealth. Unintentionally or intentionally, these rule systems can also interact with the distribution of behaviors to produce inequality by gender, race, or other statuses (Allmendinger et al. 1993, Johnson et al. 2003).

Merton's model is one example of how rewards become resources that produce subsequent awards. But rewards can become resources for further rewards through processes other than social recognition. For example, Stanovich (1986) and others (e.g., Bast & Reitsma 1998) have argued that reading competency involves multiple skills such as phonological processing abilities, word recognition skills, vocabulary, and comprehension ability. They argue that reading level has a reciprocal causal relationship with a set of cognitive skills, such that reading increases these skills, and these skills in turn increase reading ability. Reading therefore becomes a resource for further improvements in reading.

Another mechanism is what Dickens & Flynn (2001) term social-multiplier effects. Dickens and Flynn proposed a sophisticated explanation for the considerable rise in measured IQ between the early 1950s and the early 1980s in the Netherlands and nearly 20 other countries. Investigators suggested that this rise must be due to environmental change, yet the magnitude of growth was inconsistent with the, until recently, standard heritability estimates first formulated by Jensen (1973). Dickens and Flynn proposed a resolution of this paradox by hypothesizing a CA process that is in many respects similar in form to Merton's reputation-and-resources model, though the mechanism is different. They start with the frequent observation in contemporary literature that a variety of social factors cause an individual's local environment to be selected on the basis of their IQ. This selection induces a correlation between genetic and environmental factors. This link between IQ and the environment produces further cognitive gains in the subsequent period, and then further environmental selection, and then further IQ gains, and so forth into the future. If the environment (including the IQ of other people in the environment, which plays an explicit role in their theory) is identified with resources and if IQ is the outcome, their model is a version of Equation 10. Their model is more sophisticated because of their explicit specification of how genetic factors (analogous to skill in the Mertonian model) have direct effects on IQ, as well as indirect effects through environmental selection. They show that the social-multiplier effects in their model allow relatively small exogenous changes in the environment to cause surprisingly large changes over time in a population's mean IQ. In our terms, they show that CA can occur at the level of populations as well as at the level of individuals because the total environmental resources—including the IQ of significant others—are partly driven by the CA process. Their paper is an excellent example of how the failure to recognize the possibility that CA is at work can lead to incorrect models and incorrect predictions about the evolution of population distributions over time.

Other scholars have also theorized social-multiplier effects in models of behavior. Any network model of contagion that posits that behavior is influenced by the behavior of peers has this form. For example, in their review of the "neighborhood effects" literature, Jencks & Mayer (1990) noted that neighborhoods can get caught in a cycle of poverty because poverty is caused in part by individual behavior, which is shaped by the behavior of peers. If the number of positive role models in a neighborhood, for instance, dips below a threshold, negative behaviors can dominate the neighborhood and thereby intensify the poverty rate and the consequences of poverty. This type of model is essentially a form of the feedback model of Equation 8. Another example is Adler's (1985) model of stars, which shows how information costs and network effects would—in contrast with Rosen's model—produce stars in the entertainment industry and other markets even if talent levels were constant in the population.

A complete specification of a strict CA process would not just characterize the mechanism producing the effect, but would also address the mechanism that turns the process off, which, in the context of Equations 7 and 10, means that the γ parameters shrink and become negative over time. In some cases, the CA process may occur only over a specific period of time or a specific set of stages of a dynamic process. These cases would apply to processes that function within developmental limits on the human life cycle, as in Stanovich's model of reading. In other cases, position may interact with other attributes of the individual in question, and the nature of this interaction may reduce the impact of the CA process over time. For example, accumulation of human capital in the Mincer model is gradually shut down because of rational calculations that compare the reduced time horizon for returns on human capital as a worker ages with the current cost of further investment. In the case of scientific careers, productivity depends not simply upon resources but also upon skill. Finite skill creates a limit to what a scientist can accomplish even with substantial resources, and the interaction between resources and skill may produce diminishing returns to resources and cause the scientific community to redirect resources and thereby shut down the CA process. A related example is what is known as the Peter Principle, which argues that people tend to be promoted to their own level of incompetence, or in other words, positional resources operate interactively with skill, and limits on skill will ultimately diminish the advantage of positional resources for further promotions. In general, whenever the production function involves factors besides the position in question or the resources made available to that position, returns will diminish. External structural factors may also cause a shutdown of the CA mechanism. For example, a position on a career ladder provides an independent advantage only in the early or middle stages of a process; career ladders have a finite length and (as in the Venture tube model of Stewman & Konda 1983) the value of a position usually declines at a certain point on the hierarchy because the ratio of available positions at the next level declines relative to the number of competitors for these positions. Another structural mechanism for shutting down CA is the competition between generations; institutional leaders may avoid overinvesting in a few stars in order to direct resources toward nurturing future stars in the next generation (Zuckerman 1988). Another possible mechanism is the career mistakes by stars and superstars, which cast doubt on future productivity, violate social mores, lead to the withdrawal of further rewards or resources, and thereby shut down the CA process (Zuckerman 1988). In short, any structural or chance occurrence that produces a set of "negative kicks" (Cole & Singer 1991) can shut down a CA process. Until sociologists better understand the processes that produce CA, our knowledge of how these processes shut down remains more illustrative than systematic.

EVIDENCE FOR CUMULATIVE ADVANTAGE IN SPECIFIC LIFE COURSE DOMAINS

Schooling

In the educational process, progression from each step depends on attainment of and satisfactory performance in the previous step. Educational transitions therefore have a CA character, although clearly the distribution of completed schooling does not show the strong distributional skew typical of CA processes. The lack of right skew in distributions of school grades and many cognitive tests may also argue against the presence of CA mechanisms, although, as Walberg and colleagues (1984) pointed out, so-called absolute measures of cognitive achievement are likely to have highly skewed distributions. Daneman (1991), for example, has shown that the variance in reading scores grows from lower to upper grades, indicating the possibility of a CA process. Psychologists have in fact constructed developmental models of reading ability that are specified as a CA process at the individual level and that have as the main empirical evidence an increased variance in reading ability with age, coupled with structural equation models of the interrelationship between different aspects of reading ability across the early life course (Bast & Reitsma 1998).

In contrast, sociologists have been more concerned about the effects of structural factors on learning and the role of structural factors in producing group-based inequality. A huge body of literature in the sociology and economics of education addresses whether there are "school effects" at all levels of the educational process from preschool to university. Because academic performance is an important determinant for entry into high-quality postgraduate schools, colleges, high schools, and even elementary schools, the existence of school effects implies that access to a high-quality elementary school confers a positional advantage for entry into a high-quality high school and eventually into a high-quality graduate or professional school. School effects could in principle imply a CA process, but, as Attewell (2001) recently showed, high-quality schools may instead have negative effects on the quality of school at the next educational level (e.g., high school to college) because of the negative effect of high school quality on class rank, even if high-quality schools have positive effects on the quality of learning. The difficulty in establishing the strength of school effects has probably inhibited the formulation and testing of CA models driven by this mechanism.

The study of CA in the educational literature has focused mainly on tracking and its effects, and specifically on whether tracking produces growing inequality in educational outcomes over time. Some scholars have found that tracking increases the variance in academic outcomes (Gamoran & Mare 1989), and this fact, combined with the possibility that track placement directly affects future track placement (Lucas 1999), produces growing inequality over time. Kerckhoff's work (Kerckhoff 1993, Kerckhoff & Glennie 1999) has been perhaps the most explicit. Kerckhoff and Glennie's interpretation of CA as a cumulative exposure process closely follows the formulation of CA by Blau & Duncan (1967), and they see CA at work when the effects of being in a particular educational track at a specific point in the educational career has persisting direct effects after controlling for later outcomes. However, the characterization of tracking as a cumulative exposure–type CA process is by no means a settled issue, as the scholarly literature has continued to debate whether tracking systems do in fact produce growing inequality over time relative to nontracked systems. The recognition that tracking is not a unitary institution (Sorensen 1970, Gamoran 1992), and that, as Lucas (2001) argues, schools generally no longer assign students to single overarching tracks, makes theorizing about CA mechanisms in the context of school effects and to identifying them empirically more difficult.

Family and Neighborhood

Much of the recent literature about the effects of poverty asserts that CA is at work in the form of a status-resources interaction model (e.g., Hannon's 2003 identification of CA with the existence of an interaction effect of delinquency and poverty on educational attainment). Other literature on the impact of poverty while growing up (Alexander & Entwisle 1988, Duncan et al. 1998) has the character of a cumulative exposure process. Krein & Beller (1988) and McLanahan & Sandefur (1994) argue that the negative effects of growing up with a single parent increase with duration and that the duration effects vary by the age of the child (Garasky 1995). Duncan et al. (1994) have shown that children who live in persistent poverty have decreased IQ and increased behavior problems up to at least age 5 relative to children who experienced only short poverty spells or to children who never experienced poverty. Meanwhile, a large body of literature has grown around the idea that living in a poor neighborhood may produce negative effects via a cumulative exposure process. Studies based on panel data and on data from social policy experiments suggest that neighborhood characteristics have a causal effect on developmental outcomes, although producing a powerful test of this conjecture even in an experimental context has been difficult (Brookes-Gunn et al. 1994, Sampson et al. 2002, Kling & Liebman 2004), and we are unaware of research that establishes the extent to which neighborhood effects have the character of a cumulative exposure process.

Work and Careers

Merton and associates' early research on CA in sociology involved careers, and so it would seem natural that the sociological community would have researched the role of CA in the broader labor market and investigated the specific mechanisms that may drive whatever CA processes are found to exist. However, Rosenfeld's (1992) review of this literature omits any explicit mention of CA because most of the attention of career researchers in sociology was focused in other directions. The topic of CA has not been totally ignored, however, and an important segment of sociological research on careers involves CA mechanism, at least implicitly.

One important early example concerns what is often referred to as state dependence. The general model of state dependence asserted that the probability of being in a given state was a function of previous residence or duration in that state (Heckman & Borjas 1980). Previous unemployment arguably scars a worker and creates future unemployment because it reduces a worker's human capital, reduces a worker's psychological readiness for work, and makes a person less attractive to prospective employers. The early microdynamic literature on unemployment demonstrated an empirical connection between previous unemployment and future unemployment risk (Heckman & Borjas 1980, DiPrete 1981, Ellwood 1982). Such an effect is in fact a manifestation of path-dependent CA; growing inequality in life chances for employment security is affected by the resources or liabilities that workers obtain in their early careers, and this resource/liability acquisition occurs partly for reasons independent of their personal characteristics. However, as with CA processes in general, it is difficult to separate the path-dependent CA effects (pure state dependence) from unmeasured heterogeneity, which would generate an apparent effect of past unemployment on future unemployment even if path-dependent CA was not occurring. The methodological difficulties of distinguishing these possibilities may be a reason why the sociological literature has not devoted more attention to this question in recent years (see Gangl 2004 for an important exception).

Early sociological theories on internal labor markets contained CA mechanisms. As already noted, Rosenbaum (1979) proposed that careers within corporations follow a path-dependent CA process in which the outcomes of early promotion tournaments impact subsequent events. This hypothesis still has great theoretical appeal, although his empirical research did not clearly distinguish the extent to which unequal promotion rates are due to the outcomes of early promotion tournaments or whether they are instead driven by quality differences among workers. Althauser (1989) similarly argued that internal labor markets linked jobs within firms that had related firm-specific skill requirements and that location in entry-level positions of internal labor markets was a career advantage that increased over time through the mechanism of positive duration dependence. Positive duration dependence is a CA process, but most empirical research on duration dependence has found its effect to be negative rather than positive (Rosenfeld 1992).

Elman & O'Rand's (2004) recent work is an effort to construct a CA theory for explaining the experiences of late bloomers. They show that people who attain a high level of (educational) resources earlier in the life course have better outcomes than those who attain these resources later in adulthood, and they attribute this fact to the cumulative career advantages that early bloomers enjoy.

Bielby & Bielby's (1996) study of careers in the film industry is clearer than most empirical studies in sociology in its attempt to specify a status (gender)-based CA process. They define gender-based CA to exist whenever the gender interactions in Equations 12 or 14 are nonzero. They find that for Hollywood screenwriters (but not for television writers), the effects of experience on earnings is greater for men than for women (i.e., a simple status-dependent CA process), which produces widening gender-based inequality over time. They do not find evidence of path- and status-dependent CA. However, they do not investigate whether a CA process is in fact occurring either in the overall population or separately within the population of male screenwriters and female screenwriters, and this omission creates ambiguity about the correct interpretation of their results. Although a gender-dependent CA process may cause the observed growing inequality, the growing gender gap over the life course may also be a product of a gender-neutral CA process operating on small initial mean differences between men and women that arise from other causes.

Research on scientific careers has continued to develop since the early work of Merton (1973a,b, 1988), the Coles (1973), Zuckerman (1977), and Allison and colleagues (1974, 1982). Long & Fox's (1995) literature review, along with Xie & Shaumann's (2003) recent book, demonstrate considerable differences in the rewards for male and female scientists. Both Long & Fox and Reskin & Hargens (1979) found evidence that women's disadvantage grows during the early career as a result of CA processes that magnify their early disadvantages, although Long & Fox treat this evidence as more tentative than conclusive. Meanwhile, Long (1992) found sex differences in rates of rank promotion even after controlling for productivity, which suggests that—at least for the cohorts he studied—women experienced a continued disadvantage in career advancement apart from any cumulative disadvantage related to their lower starting position. Long's results, however, showed that gender differences grew only during the first 10 years of the career of the research chemists that he studied. After that, gender differences declined, which suggests that if CA was the mechanism for growing gender inequality, it shuts down by the end of the first third of a scientist's working life. Xie & Shauman argue that gender differences in science are diminishing and that remaining differences are largely explained by the rank of the academic position, the quality of the current institution, and research resources. Although their finding that productivity differences are associated with resource differences coincides with the predictions of a CA model, they do not conduct any explicit tests for whether inequality in scientific careers is in fact a consequence of a CA mechanism.

Huber has recently published a series of articles (Huber 1998, 2002) that take issue with previous evidence for CA in scientific publication. Huber argues that a CA process implies constant proportional growth, but his research instead finds that scientists have a constant production rate over their active careers, and he concludes that the familiar skewed distribution of scientific publications arises from "the skewed distribution of talent and tenacity" (Huber 2002, p. 218), which is very similar to Rosen's model. Huber's evidence is persuasive, although it does not preclude the possibility that CA is at work. For example, CA could drive scientific careers in their initial stage, but then the effect could be dampened so that over longer periods growth rates in total publications appear roughly constant. Such a possibility may be consistent with Long's (1992) finding that gender differences in science grow only during the first third of the research career.

The specific form of the female career disadvantage has received extensive attention in the scholarly literature. A significant fraction of the gender career disadvantage arises from gender differences in curriculum choices (Jacobs 1996, Xie & Shauman 2004), whereas Petersen & Morgan (1995) have shown that sex segregation at the combined occupation-establishment level explains nearly 90% of gender wage inequality. Petersen's (2004) recent study of a large corporation in the service sector found considerable gender differences in starting pay even when education was controlled (although he could not rule out that these differences were related to gender differences in prior experience). The CA thesis—namely that initial differences grow over time—was strongly refuted, however, for this specific corporation. Instead, gender differences gradually declined after initial hire. The other major source of female disadvantage is the career penalty from cumulative exposure to motherhood. The human capital model explains how withdrawal from the labor market has negative effects on wages and argues that these negative effects grow with the length of the withdrawal (Becker 1985). Waldfogel (1998) has provided evidence that the quality of maternity leave influences the size of the effect of exposure to nonwork on career wage trajectories.

Gould's (2002) model for the growth of inequality in society arises from patterns of social interaction and clearly incorporates a CA mechanism. Gould argues that social hierarchies emerge through a set of four factors: intrinsic quality differences among people, a presumption that an individual's welfare is enhanced by associating with individuals of higher quality, a presumption that quality is not fully observable, and a presumption that people infer quality on the basis of the attention an individual receives from others. This last hypothesis stimulates a CA process in which social attention paid to an individual generates more attention from others and then still more attention from others as the compounding attention increases perceived quality and therefore the welfare gains from associating with that person. Gould argues further that a person's interest in associating with another person of high quality is tempered by a disinclination to associate with people who do not reciprocate the attention, and shows that sufficiently large costs from asymmetric association can produce an equilibrium hierarchy that is based partly on intrinsic quality differences and partly on CA. Although both Merton's and Gould's models are based on reputational processes, they differ in an important respect; Merton saw early reputation as attracting resources that facilitate higher productivity, even as that reputation allowed the scientist to get a bigger career payoff from any given level of productivity. Gould's model is not based on a reputation-resource-productivity-reputation feedback loop, but rather sees reputation as directly producing higher levels of reputation through a process of differential association. Gould's evidence for his model pertains only to small group research carried out in 1961 or earlier (Blatz 1937, Newcomb 1961, Bales 1970); more recent evidence for larger groups is needed for a stronger test.

Wealth accumulation is another process that should be governed by CA processes, and here the mechanism is straightforward, as wealth accumulates through the mechanism of compound returns to investment. A growing body of literature has developed around the role of wealth as both a measure and a determinant of black-white inequality in the United States. Because historically whites have more wealth than blacks, they start each generation with an advantage in initial wealth. Therefore, CA works to maintain white advantage (Oliver & Shapiro 1995, Conley 1999, Scholz & Levine 2004).

Other Social Domains

CA processes play a potentially important role in other life course processes. Criminal careers are frequently seen as outcomes of a cumulative disadvantage process. Cohort members display different types of offending trajectories over their lifetimes (which is analogous to the fanning pattern produced by CA processes), typically with a gradual cessation of crime during the adult years (Sampson & Laub 1997, 2003). Moreover, criminal history negatively affects labor market processes via a combination of scarring and possibly (though statistical evidence is not strong) race-specific levels of scarring from a criminal history (Pager 2003). Sampson & Laub (1997, 2003) argue that most criminals received a series of negative pushes that increased their probability of becoming more criminal via deviant labeling, harsh penal processing, and deflections associated with getting in trouble with the law. Meanwhile, other at-risk youth were pushed away from crime through strong social ties, marriage, and more stable employment. Their path-dependent CA theory differs from Moffitt's (1993) dual taxonomy theory in the postulated timing of the CA process in the life course. Moffitt sees distinctive criminal paths formed by childhood and early adolescence, whereas Sampson & Laub (2003) see external shocks continuing to shift individuals between different career trajectories later in the life course. In terms of Equation 10, these theories disagree over the age pattern of the γ parameters, with Moffitt seeing young age shocks as having primary importance and Sampson & Laub seeing later adolescent shocks to be important as well. Superimposed on these individual-specific factors are global (non-CA) processes that suppress criminal behavior later in life, along with questions about whether these suppressing factors have a homogeneous impact or whether they interact with a person's characteristics and prior criminal career. Although these and other questions remain to be answered, researchers have made considerable strides in understanding the path-dependent character of criminal careers.

The health literature also contains many discussions about CA, including such topics as the cumulative impact of exposure to risk factors such as education (Ross & Wu 1996) or obesity (Ferraro & Kelley-Moore 2003) on health or the cumulative effects of discrimination on race differences in health (Krieger 1994). Singer and colleagues (1998) advocate modeling the heterogeneity in mental-health path-dependent CA processes by means of a person-centered biographical approach rather than a variable-centered approach. Although Sampson & Laub's (2003) recent work is a good example of a biographical approach to the study of careers, the specifics of the Singer et al. (1998) strategy for studying CA processes has not yet been widely implemented.

Some of the literature on CA cuts across multiple social domains. Most notably, the literature on cumulative discrimination as a mechanism for racial inequality cuts across life course domains. This literature, which continues to grow from its roots in Blau & Duncan's 1967 book, was recently reviewed by a National Research Council Panel as part of a broader effort to measure racial discrimination (Blank et al. 2004). Some authors focused on the accumulation of negative direct effects of being black, as found in a cumulative exposure model (Williams & Neighbors 2001). Some authors postulated sources of discrimination that stem from status-resource interaction processes (Yang 1999). Finally, some postulated the type of feedback mechanisms found in path-dependent CA (Johnson & Neal 1998). Although the literature on CA as applied to discrimination is growing, the editors of the National Academy Research Council report describe it as very thin at the empirical level, partly because of the lack of efforts to construct formal models and partly because of the difficulty of estimating such models (Blank et al. 2004, p. 224).

CONCLUSION

Cumulative advantage as either an explicit or implicit inequality-generating process has pervaded the sociological literature for several decades. Aggregate data suggest that CA is at work in many social domains. Thus, Allison & Stewart (1974) found the Gini coefficient for scientific productivity grew with career age in their sample of scientists; it started at .42, grew by .0084 per year after one attained a PhD, and reached .69 in the eldest age stratum. Using synthetic cohort data on Protestant ministers, Broughton & Mills (1980) estimated that the Gini coefficient for ministers' command of resources (as measured by religious congregation size) started at .15 and moved to .34 over the 30-year careers of Protestant ministers. Kerckhoff & Glennie (1999) found that the relative educational deflection caused by one's tenth-grade placement created a gap of 14.7 percentile points after 2 years, 33.5 points after 6 years, and some 46.5 points after 12 years. Bielby & Bielby (1996) found that women television writers suffered a relatively constant 11%–25% penalty at every stage of their career, but for Hollywood screenwriters, the gender gap was 4%–6% at one year of experience, more than 20% at 5 years of experience, and more than 40% at 15 years of experience. Little is currently known about the reasons for these different growth rates in inequality.

To make progress, there is a need for more explicit attention to mechanisms, more formal theorizing, and greater attention to methodological issues in the specification and testing of theories. To rectify these shortcomings, we have called attention to the distinction between strict CA and the CA theories that are primarily focused on changes in between-group inequality over time. Strict CA theories typically take two forms, which we have referred to as simple and path-dependent CA, whereas CA theories of between-group inequality generally take the form of cumulative exposure and status-resource interaction models. It is important to be explicit about these different models because they imply different inequality-generating mechanisms and because they predict different trajectories for the distribution of population and group-specific outcomes. More precise theorizing and more systematic empirical study of these mechanisms are needed in order to move beyond the use of CA as a descriptive characterization of diverging trajectories toward a deeper understanding for the reasons why trajectories diverge at both the group and the individual level of observation.

Given the limits of space, this review cannot examine the development and testing of CA in all subfields of sociology, not to mention the broader social scientific literature. In our view, however, attention to this broader literature is important. Just as population biology became a source of inspiration for ecological models of organizational populations, CA and related processes in art markets (Fox & Kochanowski 2004, Salganik et al. 2005), models of phase transitions in physics and self-organized criticality (Newman 2005), models of network growth (Newman 2003), models of IQ growth in populations over time (Dickens & Flynn 2001), and other applications may provide inspiration for new approaches and greater understanding of inequality-generating processes in society.

ACKNOWLEDGMENTS

The authors thank Jonathan Cole, Douglas Massey, Karl Ulrich Mayer, Angela O'Rand, Matt Salganik, Michael Sobel, and an anonymous reviewer for helpful comments on an earlier draft of this manuscript.

The *Annual Review of Sociology* is online at http://soc.annualreviews.org

LITERATURE CITED

Adler M. 1985. Stardom and talent. *Am. Econ. Rev.* 75:208–12

Alexander KL, Entwisle DR. 1988. Achievement in the first 2 years of school: patterns and processes. *Monogr. Sociol. Res. Child Dev* 53:1–157

Allison PD, Long JS, Krauze TK. 1982. Cumulative advantage and inequality in science. *Am. Sociol. Rev.* 47:615–25

Allison PD, Stewart JA. 1974. Productivity differences among scientists: evidence for accumulative advantage. *Am. Sociol. Rev.* 39: 596–606

Allmendinger JE, Brueckner H, Brueckner E. 1993. The production of gender disparities over the life course and their effects in old age: results from the West German life history study. In *Age, Work, and Social Security*, eds. AB Atkinson, MH Rein, pp. 188–223. New York: St. Martin's

Althauser RP. 1989. Internal labor markets. *Annu. Rev. Sociol.* 15:143–61

Arthur WB. 1994. *Increasing Returns and Path Dependence in the Economy*. Ann Arbor: Univ. Michigan Press

Attewell P. 2001. The winner-take-all high school: organizational adaptations to educational stratification. *Sociol. Educ.* 74:267–95

Bales RF. 1970. *Personality and Interpersonal Behavior.* New York: Holt, Rinehart & Winston

Baltagi BH. 2001. *Econometric Analyses of Panel Data.* New York: Wiley. 3rd ed.

Bast J, Reitsma P. 1998. Analyzing the development of individual differences in terms of Matthew effects in reading: results from a Dutch longitudinal study. *Dev. Psychol.* 34:1373–99

Becker G. 1985. Human capital, effort, and the sexual division of labour. *J. Labor Econ.* 3:33–38

Bielby DD, Bielby WT. 1996. Women and men in film: gender inequality among writers in a culture industry. *Gend. Soc.* 10:248–70

Blank RM, Dabady M, Citro CF, eds. 2004. *Measuring Racial Discrimination*. Washington, DC: Natl. Acad.

Blatz WE, Chant N, Charles MW, Fletcher MI, Ford NHC, et al. 1937. *Collected Studies on the Dionne Quintuplets.* Toronto: Univ. Toronto Press

Blau FD, Kahn LM. 1992. The gender earnings gap: learning from international comparisons. *Am. Econ. Rev.* 82:533–38

Blau PM, Duncan OD. 1967. *The American Occupational Structure.* New York: Wiley

Brookes-Gunn J, Duncan GJ, Aber JL, eds. 1994. *Neighborhood Poverty: Context and Consequences for Children*, Vol. 1. New York: Russell Sage Found.

Broughton W, Mills JEW. 1980. Resource inequality and accumulative advantage: stratification in the ministry. *Soc. Forces* 58:1289–301

Cole JR, Cole S. 1973. *Social Stratification in Science.* Chicago: Univ. Chicago Press

Cole JR, Singer B. 1991. A theory of limited differences: explaining the productivity puzzle in science. In *The Outer Circle: Women in the Scientific Community*, Chapter 13, eds. H Zuckerman, JR Cole, JT Bruer, pp. 277–310. New York: W.W. Norton and Co.

Cole S. 1970. Professional standing and the reception of scientific discoveries. *Am. J. Sociol.* 76:286–306

Cole S, Cole JR, Simon GA. 1981. Chance and consensus in peer review. *Science* 214:881–86

Conley D. 1999. *Being Black, Living in the Red: Race, Wealth and Social Policy in America.* Los Angeles: Univ. California Press

Daneman M. 1991. Individual differences in reading skills. In *Handbook of Reading Research*, Vol. 2, eds. R Barr, ML Kamil, PB Mosenthal, PD Pearson, pp. 512–38. White Plains: Longman

Dannefer D. 1987. Aging as intracohort differentiation: accentuation, the Matthew effect and the life course. *Sociol. Forum* 2:211–36

Dannefer D. 2003. Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* 58:S327–38

Dickens WT, Flynn JR. 2001. Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychol. Rev.* 108:346–69

DiPrete TA. 1981. Unemployment over the life cycle: racial differences and the effect of changing economic conditions. *Am. J. Sociol.* 87:286–307

Duncan GJ, Brooks-Gunn J, Klebanov PK. 1994. Economic deprivation and early-childhood development. *Child Dev.* 65:296–318

Duncan GJ, Yeung WJ, Brooks-Gunn J, Smith JR. 1998. How much does childhood poverty affect the life chances of children? *Am. Sociol. Rev.* 63:406–23

Ellwood D. 1982. Teenage unemployment: permanent scars or temporary blemishes? In *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*, eds. RB Freeman, DA Wise, pp. 349–90. Chicago: Univ. Chicago Press

Elman C, O'Rand AM. 2004. The race is to the swift: socioeconomic origins, adult education, and wage attainment. *Am. J. Sociol.* 110:123–60

Faia M. 1975. Productivity among scientists: a replication and elaboration. *Am. Sociol. Rev.* 40:825–29

Ferraro KF, Kelley-Moore JA. 2003. Cumulative disadvantage and health: long-term consequences of obesity? *Am. Sociol. Rev.* 68:707–29

Fox MA, Kochanowski P. 2004. Models of superstardom: an application of the Lotka and Yule distributions. *Pop. Music Soc.* 27:507–22

Frank RH, Cook PJ. 1995. *The Winner Take All Society: How More and More Americans Compete for Ever Fewer and Bigger Prizes, Encouraging Economic Waste, Income Inequality, and an Impoverished Cultural Life*. New York: Free Press

Gamoran A. 1992. The variable effects of school tracking. *Am. Sociol. Rev.* 57:812–28

Gamoran A, Mare RD. 1989. Secondary school tracking and educational equality: compensation, reinforcement, or neutrality? *Am. J. Sociol.* 94:1146–83

Gangl M. 2004. Welfare states and the scar effects of unemployment: a comparative analysis of the United States and West Germany. *Am. J. Sociol.* 109:1319–64

Garasky S. 1995. The effects of family structure of educational attainment: Do the effects vary by the age of the child? *Am. J. Econ. Sociol.* 54:89–105

Gould R. 2002. The origins of status hierarchies: a formal theory and empirical test. *Am. J. Sociol.* 107:1143–78

Greenwald AG, Banaji MR. 1995. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol. Rev.* 102:4–27

Hannon L. 2003. Poverty, delinquency, and educational attainment: cumulative disadvantage or disadvantage saturation? *Sociol. Inq.* 73:575–94

Heckman JJ, Borjas G. 1980. Does unemployment cause future unemployment? Definitions, questions, and answers from a continuous time model of heterogeneity and state dependence. *Economica* 47:247–83

Heckman JJ, Lochner LJ, Todd PE. 2003. *Fifty years of Mincer earnings regressions. IZA discussion paper no. 775, May 2003*. NBER Work. Pap. 9732, Natl. Bur. Econ. Res., Inc. http://ideas.repec.org/p/nbr/nberwo/9732.html

Hirsch F. 1976. *Social Limits to Growth*. Cambridge, MA: Harvard Univ. Press

Huber JC. 1998. Cumulative advantage and success-breeds-success: the value of time pattern analysis. *J. Am. Soc. Inform. Sci.* 49: 471–76

Huber JC. 2002. A new model that generates Lotka's law. *J. Am. Soc. Inform. Sci. Technol.* 53:209–19

Jacobs J. 1996. Gender inequality and higher education. *Annu. Rev. Sociol.* 22:153–85

Jencks C, Mayer SE. 1990. The social consequences of growing up in a poor neighborhood. In *Inner-City Poverty in the United States*, eds. JL Lynn, MGH McGreary, pp. 111–86. Washington, DC: Natl. Acad.

Jensen AR. 1973. *Educability and Group Differences.* New York: Harper Row

Johnson RW, Sambamoorthi U, Crystal S. 2003. Gender differences in pension wealth and their impact on late-life inequality. In *Annual Review of Gerontology and Geriatrics, Volume 22: Economic Outcomes in Later Life: Public Policy, Health and Cumulative Advantage*, eds. S Crystal, D Shea, pp. 116–37. New York: Springer

Johnson WR, Neal D. 1998. Black skills and the black-white earnings gap. In *The Black-White Test Score Gap*, eds. C Jencks, M Phillips, pp. 480–98. Washington, DC: Brookings Inst.

Kerckhoff AC. 1993. *Diverging Pathways: Social Structure and Career Deflections.* Cambridge, MA: Cambridge Univ. Press.

Kerckhoff AC, Glennie E. 1999. The Matthew effect in American education. *Res. Sociol. Educ. Social.* 12:35–66

Kerin RA, Varadarajan PR, Peterson RA. 1992. First-mover advantage: a synthesis, conceptual framework, and research propositions. *J. Mark.* 56:33–52

Kling J, Liebman J. 2004. Experimental analysis of neighborhood effects on youth. KSG Work. Pap. No. RWP04-034, NBER, Princeton Univ., Harvard Univ.

Krein SF, Beller AH. 1988. Educational attainment of children from single parent families: differences by exposure, gender, and race. *Demography* 25:221–34

Krieger N. 1994. Epidemiology and the web of causation: Has anyone seen the spider? *Soc. Sci. Med.* 39:887–903

Laub JH, Sampson RJ. 2003. *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge, MA: Harvard Univ. Press

Long JS. 1992. Measures of sex differences in scientific productivity. *Soc. Forces* 71:159–78

Long JS, Fox MF. 1995. Scientific careers: universalism and particularism. *Annu. Rev. Sociol.* 21:45–71

Lotka AJ. 1926. The frequency distribution of scientific productivity. *J. Wash. Acad. Sci.* 16:317–23

Lucas SR. 1999. *Tracking Inequality: Stratification and Mobility in American High Schools*. New York: Teach. Coll.

Lucas SR. 2001. Effectively maintained inequality: education transitions, track mobility, and social background effects. *Am. J. Sociol.* 106:1642–90

Mayer KU, Maas I, Wagner M. 1999. Socioeconomic conditions and social inequalities in old age. In *The Berlin Aging Study: Aging from 70 to 100*, eds. PB Baltes, KU Mayer, pp. 227–58. Cambridge, UK: Cambridge Univ. Press

McLanahan S, Sandefur G. 1994. *Growing Up with a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard Univ. Press

Merton RK. 1973a. The Matthew effect in science. In *The Sociology of Science*, ed. NW Storer, pp. 439–59. Chicago: Univ. Chicago Press

Merton RK. 1973b. The normative structure of science. In *The Sociology of Science*, ed. NW Storer, pp. 267–78. Chicago: Univ. Chicago Press

Merton RK. 1988. The Matthew effect in science, ii: cumulative advantage and the symbolism of intellectual property. *Isis* 79:606–23

Mincer J. 1974. *Schooling, Experience, and Earnings*. New York: Columbia Univ. Press

Mitzenmacher M. 2003. A brief history of generative models for power law and lognormal distributions. *Internet Math* 1:226–51

Moffitt TE. 1993. Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychol. Rev.* 100:674–701

Newcomb TM. 1961. *The Acquaintance Process*. New York: Holt, Rinehart & Winston

Newman MEJ. 2003. The structure and function of complex networks. *SIAM Rev.* 45:167–256

Newman MEJ. 2005. Power laws, Pareto distributions, and Zipf's law. *Contemp. Phys.* 46:323–51

Oliver ML, Shapiro TM. 1995. *Black Wealth/White Wealth: A New Perspective on Racial Inequality*. New York: Routledge

O'Rand AM. 1996. The precious and the precocious: understanding cumulative disadvantage and cumulative advantage over the life course. *Gerontologist* 36:230–38

O'Rand AM. 2002. Cumulative advantage theory in life course research. *Annu. Rev. Gerontol. Geriatr.* 22:14–20

Pager D. 2003. The mark of a criminal record. *Am. J. Sociol.* 108:937–75

Petersen T, Morgan LA. 1995. Separate and unequal: occupation-establishment sex segregation and the gender wage gap. *Am. J. Sociol.* 101:329–61

Petersen T, Saporta I. 2004. The opportunity structure for discrimination. *Am. J. Sociol.* 109:852–901

Price DJ de S. 1965. Networks of scientific papers. *Science* 149:510–15

Price DJ de S. 1976. A general theory of bibliometric and other cumulative advantage processes. *J. Am. Soc. Inform. Sci.* 27:292–306

Raftery AE. 1995. Bayesian model selection in social research. *Sociol. Methodol.* 25:111–63

Reskin BF, Hargens LL. 1979. Scientific advancement of male and female chemists. In *Discrimination in Organizations*, eds. R Alvarez, KG Lutterman, pp. 100–22. San Francisco: Jossey-Bass

Robins JM, Blevins D, Ritter G, Wulfsohn M. 1993. G-estimation of the effect of prophylaxis therapy for pneumocystis carinii pneumonia on the survival of AIDS patients. *Epidemiology* 3:319–36

Rosen S. 1981. The economics of superstars. *Am. Econ. Rev.* 71:845–58

Rosenbaum JE. 1979. Tournament mobility: career patterns in a corporation. *Adm. Sci. Q.* 24:220–41

Rosenfeld RA. 1992. Job mobility and career processes. *Annu. Rev. Sociol.* 18:39–61

Ross CE, Wu CL. 1996. Education, age, and the cumulative advantage in health. *J. Health Soc. Behav.* 37:104–20

Salganik MJ, Dodds PS, Watts DJ. 2006. An experimental study of inequality and unpredictability in an artificial cultural market. *Science*. In press

Sampson RJ, Laub JH. 1997. A life-course theory of cumulative disadvantage and the stability of delinquency. *Adv. Criminol. Theory: Dev. Theories Crime Delinq.* 7:133–61

Sampson RJ, Laub JH. 2003. Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology* 41:555–92

Sampson RJ, Morenoff JD, Gannon-Rowley T. 2002. Assessing 'neighborhood effects': social processes and new directions in research. *Annu. Rev. Sociol.* 28:443–78

Scholz JK, Levine K. 2004. U.S. black-white wealth inequality. In *Social Inequality*, ed. KM Neckerman, pp. 895–929. New York: Russell Sage Found.

Simon HA. 1955. On a class of skew distribution functions. *Biometrika* 42:425–40

Singer B, Ryff CD, Carr D, Magee WJ. 1998. Linking life histories and mental health: a person-centered strategy. *Sociol. Methodol.* 28:1–51

Sorensen AB. 1970. Organizational differentiation of students and educational opportunity. *Sociol. Educ.* 43:355–76

Stanovich KE. 1986. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21:360–407

Stewman S, Konda SL. 1983. Careers and organizational labor markets: demographic models of organizational behavior. *Am. J. Sociol.* 88:637–85

Sutton J. 1997. Gibrat's legacy. *J. Econ. Lit.* 35:40–59

Walberg HJ, Strykowski BF, Roval E, Hung SS. 1984. Exceptional performance. *Rev. Educ. Res.* 54:87–112

Waldfogel J. 1998. The family gap for young women in the United States and Britain: Can maternity leave make a difference? *J. Labor Econ.* 16:505–45

Williams DR, Neighbors HW. 2001. Racism, discrimination, and hypertension: evidence and needed research. *Eth. Dis.* 11:800–16

Wilson EO. 1975. *Sociobiology: The New Synthesis*. Cambridge, MA: Harvard Univ. Press

Xie Y, Shauman K. 2003. *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard Univ. Press

Yang R. 1999. Race, status attainment, and depression: intervening effects of consequential life events. *Race Soc.* 2:195–216

Zuckerman H. 1977. *Scientific Elite.* New York: Free Press

Zuckerman H. 1988. Accumulation of advantage and disadvantage: the theory and its intellectual biography. In *Robert K. Merton and Contemporary Sociology*, eds. C Mongardini, S Tabboni, pp. 139–62. New York: Trans. Publ.

Annual Review of Sociology
Volume 32, 2006

Contents

Errata

An online log of corrections to *Annual Review of Sociology* chapters (if any, 1997 to the present) may be found at http://soc.annualreviews.org/errata.shtml