

Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments*

Thomas A. DiPrete
Columbia University

Gregory M. Eirich
Columbia University

November 22, 2005

Abstract

While originally developed by 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. We show that the term "cumulative advantage" has come to have multiple meanings in the sociological literature. We distinguish between these alternative forms, discuss mechanisms that have been proposed in the literature that might produce cumulative advantage, and review the empirical literature in the areas of education, work careers, and related life course processes.

KEYWORDS: Inequality, Life Course, Careers, Cumulative Advantage

1 Introduction

The concept of cumulative advantage (CA) frequently appears in the social scientific literature, including literature on social mobility, poverty, race, crime, education, and human development. Traceable to the work of Merton (Merton 1973b; Merton 1973a; Merton 1988), the "Matthew effect" was defined by Merton 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

*We 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.

(see Zuckerman 1988 for a recent review of Merton's contribution to the development of CA theory). Cole and Cole (1973) generalized these insights to produce a dynamic theory of stratification in scientific careers (see also Cole 1970, Zuckerman 1977). Allison and colleagues (e.g., Allison, Long, and Krauze 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.

The literature frequently invokes CA mechanisms as an explanation for inequality. Life course scholars see CA working in patterns of diverging cohort trajectories (e.g. Ross and Wu 1996; O'Rand 1996; or Mayer, Maas, and Wagner 1999; see O'Rand 2002 and Dannefer 2003 for recent literature reviews). The organizations, markets, and networks literature has analogous terms for cumulative advantage, such as first mover advantage (Kerin, Varadarajan, and Peterson 1992), path-dependent increasing returns (Arthur and Arrow 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 and Banaji 1995), the scarring effects of unemployment (Ellwood 1982), and cumulative discrimination (Blank, Dabady, and Citro 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 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. 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 etc to catch up. Ironically, despite their obvious theoretical and policy importance, and despite widespread references to their existence in the literature, 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, while 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 most simply illustrated by the process of wealth accumulation through the mechanism of compound interest. When an initial sum of money is placed into an account where 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 \$1000 account produces a \$50 accumulation, while a \$100,000 account produces \$5,000. If the interest rate is constant, the numerical difference between two unequal starting fortunes grows exponentially over time, from \$99K at the start to \$126K after five years to \$161K after 10 years and so forth. The yearly increments to these fortunes also grow exponentially. However, the ratio of the fortunes remains constant (at 100:1) no matter how long the process continues. If instead the size of the interest rate depends positively on the size of the principal (e.g., if one could get 7% interest on \$100,000, but only 5% interest on \$1000), then both the difference and the ratio of the two fortunes 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 period of time depends only on the speed and the 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 that is often found in the sociological literature can be traced back to Blau and Duncan's *The American Occupational Structure* (Blau and Duncan 1967). Their analyses of the 1962 *Occupational Changes in a Generation* data showed that African-American males suffered from a "cumulative disadvantage" relative to whites. Their concept of cumulative disadvantage, which emphasized group differences rather than inequality within a group or within an entire population, is – as we shall 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. For them, 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 that they referred to as "perverse" equality). The Blau-Duncan approach can be generalized to variables conceptualized as exposures over some (possibly long) duration to a treatment, such as growing up in a poor vs. a rich family, growing up in a single-parent vs. a two-parent family, growing up in a poor vs. an affluent neighborhood, or being assigned to a low vs. a high academic track in school.

An additional difference between what we have called strict CA and the Blau-Duncan approach to CA concerns over-time growth in inequality. Strict CA – as illustrated by the example of wealth accumulation – will produce growing differences between individuals and growing mean differences between groups over time. However, as Allison, Long, and Krauze (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 in 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

Both Merton's and Blau and Duncan's research on CA are 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 not sufficient 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.

2 Formal Considerations

Evidence for CA processes can take several forms. One class of evidence is distributional in character. It has often been noted that 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

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

Price (1965, 1976) argued that the number of citations to scientific papers follows a Yule process and produces a power-law, while Fox and Kochanowski (2004) found that commercial success by recording artists follows a generalized Lotka distribution, though the success distributions vary by gender and race. As Newman (2005) recently showed, the distribution of citations, the wealth of rich people, and the distribution of sales of books all follow power

¹Suppose 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-type 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} .

laws. Newman further points out that strict CA (which we formalize below) is one mechanism by which a power-law distribution is produced, 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 do a better job of describing 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 would tell 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, Long, and Krauze 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 need not produce growing inequality over time (Allison, Long, and Krauze 1982).

A third form of evidence comes from micro-level dynamic data, which allow the estimation and testing of parametric models. Allison and colleagues (Allison, Long, and Krauze 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 may depend upon prior successes. Allison et al. specified λ as

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

where Y_{it} is the cumulated number of accomplishments of individual i at time t , where α_i describes baseline differences between individuals, and where $\gamma_i > 0$. 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. is similar to the well known (Yule process) differential equation for exponential growth, 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 of 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

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

where for simplicity we assume equally spaced observations (thus obviating the need for a Δt factor on the right hand side). 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, one could replace

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

by

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

which causes the CA process to slow as the time reaches some critical value T^* as might occur in a life course process. One can further note that for small period to period changes,

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

We can therefore propose a model for $\ln(Y_{it})$ in which time dependence is expressed as a quadratic function of a variable that we call z (which could stand for age, time, labor market experience etc.), 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

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

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 produce cumulative change in Y unless the variables in X themselves are accumulating over time. While 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 , or 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, e.g.:

$$Y_{it} = (1 + \gamma)Y_{i,t-1} + \beta X_{it} + \epsilon_{it} \quad (7)$$

This equation expresses a form of *path dependent CA*, by which we mean that 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 as one moves further into the future.² To put it another way, exogenous chance events have long-term consequences, which is not true of the simple form of CA. The idea that exogenous chance events have long-term consequences is an essential characteristic of the cumulative advantage process proposed by Merton, and is to be distinguished 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, produces future productivity. If we let R be the resources that are made available on the basis of prior success, and if these resources are the causes of additions to productivity, then the true model takes the following form:

$$\begin{aligned} Y_{it} - Y_{i,t-1} &= \alpha R_{it} + \beta X_{it} + v_{it} \\ R_{it} &= \delta Y_{i,t-1} + w_{it} \end{aligned} \quad (8)$$

so that

$$\begin{aligned} Y_{it} - Y_{i,t-1} &= \alpha(\delta Y_{i,t-1} + w_{it}) + \beta X_{it} + v_{it} \\ &= \alpha\delta Y_{i,t-1} + \beta X_{it} + \alpha w_{it} + v_{it} \\ &= \gamma Y_{i,t-1} + \beta X_{it} + \epsilon_{it} \end{aligned}$$

where $\alpha\delta = \gamma$ and $\alpha w_{it} + v_{it} = \epsilon_{it}$, and where v and w are stochastic disturbances. It is important to note that equation (8) is the same model as equation (7), but is elaborated to show the mechanism that drives the CA process. Such

²To clarify, suppose that there are no explicit X variables, and 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 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 grow larger as time progresses.

feedback loops could 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, e.g.

$$Y_{it} = \sum_{j=0}^{t-1} \gamma_{tj} Y_{ij} + \beta' X_{it} + \epsilon_{it} \quad (9)$$

where the magnitude of each γ coefficient depends in the general case 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 deeper-past outcomes on current outcomes. It allows, for example, for spectacular performance in an early year of the career (which causes the individual to be labelled a rising "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 while 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) imply increasing inequality over time. Allison et al. argued that γ might vary across scientists because, as they put it, "'good' papers are more likely to bring further advantage than 'bad' papers" (Allison, Long, and Krauze 1982, p. 622). Other sources of heterogeneity in γ could derive from unchanging variables such as race or gender as well as exposure to "treatments" such as living in a poor neighborhood, living in a family with a single parent, being in a high or low academic track while in school, or educational attainment. 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 generally 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 labelled as s , the most relevant forms for our purposes include:

$$Y_{it} \text{ or } \ln(Y_{it}) = \gamma_1 z_{it} + \gamma_2 z_{it}^2 + \beta_1' X_{it} + s_{it} \beta_2' X_{it} + \epsilon_{it} \quad (10)$$

$$\ln(Y_{it}) = \gamma_1 z_{it} + \gamma_2 z_{it}^2 + \beta X_{it} + \delta s_{it} + s_{it} (\gamma_3 z_{it} + \gamma_4 z_{it}^2) + \epsilon_{it} \quad (11)$$

$$= (\gamma_1 + \gamma_3 s_{it}) z_{it} + (\gamma_2 + \gamma_4 s_{it}) z_{it}^2 + \beta X_{it} + \delta s_{it} + \epsilon_{it} \quad (12)$$

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

$$Y_{it} - Y_{i,t-1} = Y_{i,t-1} \gamma + \beta X_{i,t-1} + \delta s_{it} + \epsilon_{it} \quad (14)$$

$$Y_{it} - Y_{i,t-1} = Y_{i,t-1} (\gamma + \delta s_{it}) + \beta X_{i,t-1} + \epsilon_{it} \quad (15)$$

In equation (10), which we call 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 (11), the rate of CA is governed by the coefficients of z and z^2 , but the coefficients of these variables depend upon s . This is one form of *status-dependent CA*. In equations (13) and (14), s affects the rate of growth 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 upon the current magnitude of Y . We might label the process by which one's status directly affects outcomes in equations (13) and (14) as a *cumulative exposure* process because continuing exposure to the 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 (14)). It is important to note, however, 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 (15), which we refer to as *path and status dependent CA*, the magnitude of the intensity parameter of the path-dependent CA process depends upon the status.

As noted above, the Mertonian form of CA is what we have labelled *path-dependent CA*. Elaborations of this idea to account for gender differences in science in Cole and Singer (Cole and Singer 1991) are what we have labelled *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 that was revealed in the 1962 OCG data.

The illustrative models above by no means exhaust the collection of models that have been 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 of no exposure,

and Robins, Blevins, Ritter, and Wulfsohn (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-Duncan described race as a source of cumulative advantage/disadvantage.

The distinction between the mechanisms described above can be subtle, though still quite important. Thus race (acting as status s in the equations above) may have a continuing effect on wage growth throughout the career because of discrimination (as in equations (11), (13) or (14)). The effect of race may also work in interaction with education (as in equation (10)) because the effect of discrimination varies by education, or because race affects the quality of schooling which affects the returns to education. Race would also affect wealth accumulation as in equation (15) 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 and 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, namely:

1. In a CA process, the rate of growth 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 (though it is not necessarily strict CA as we 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. At the risk of being repetitious, 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 about 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

³That this omission can be problematic is shown by Blau and Kahn's (1992) demonstration that rising inequality in a population generally produces rising between-group inequality as a byproduct. Exclusive attention to group differences can cause a researcher to misidentify a population-level process as a group-specific process.

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 itself has a persisting effect on outcomes, and that group-membership itself 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 literature often makes distinctions between cumulative advantage and cumulative disadvantage. The literature on scientific careers is typically framed in terms of cumulative advantage because it focuses on those who become stars in their field, while the literature on race inequality or on health is often framed in terms of cumulative disadvantage because the focus is on those who experience disadvantaged outcomes. While rhetorically useful, these alternative framings are for the most part a device for indicating the direction of the effect (see also Zuckerman 1988). In the context of a CA process, one is at a disadvantage relative to someone else 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 negative 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 the ability of the analyst to distinguish path-dependent CA from simple CA and to distinguish strict CA processes from other processes. Equations (7), (14), or (15) 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 (GMM) estimators or some related instrumental variable regression strategy (Baltagi 1995), 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 of 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 it is not straightforward to decide among them on the basis of statistical tests. However, methods such as Bayesian model selection (Raftery 1994) may help to adjudicate between competing models. Before one gets to this point, however, the underlying model needs to be clearly specified.

3 Mechanisms

The equations of the previous section may do a good job of defining what CA means, but they do not by themselves explain why it might 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. The first was that resources in the scientific world were limited. The second was that scientific talent was difficult to observe directly. The third was that allocation of resources in science was governed by the norms of universalism (recognition should be granted based on the quality of scientific work) and communism (resources should be allocated in order to maximize the overall productivity of the scientific community). Scientific resources were therefore not simply a "reward" for past productivity, but were 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 *ex ante*, the scientific community favored those who had been most successful in the past with additional resources and attention. One consequence of this mechanism was that the gap in rewards between a more able and a less able scientist would grow over time. A second consequence was that chance events (e.g., unequal luck in the draw of reviewers upon submitting a grant proposal for funding (Cole, Cole, and Simon 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) was that scientists with greater reputations would gain greater rewards from work of a given quantity and quality than would 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 the rents for the highest levels of talent would increase more than proportionally (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 placed a large value on innate talent and no value on external resources in his model, and focused his attention on an economic rationale why small differences in talent could lead through market mechanisms to large differences in rewards. While recognizing potential differences in talent, Merton's mechanism 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

and 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 and Cole (1973) and later Cole and Singer (1991) argued that CA operated partly through social psychological, cultural and structural mechanisms. These mechanisms could produce a motivation to work harder if successful, and a motivation to work less hard if one experiences rejection. In their "theory of limited differences," Cole and 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 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 and 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, Lochner, and Todd 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 to this investment. Divergence in Merton's model can occur through differences in endowments, through chance differences early in the career, or through any other factors (e.g., the quality of mentors or patrons) that produce differences in the quality of early career performance. 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 cumulative advantage. 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 be-

haviors to produce inequality by gender, race, or other status (Allmendinger, Brueckner, and Brueckner 1993; Johnson, Sambamoorthi, and Crystal 2002).

Merton's model is one example of how rewards become resources that produce subsequent awards. But rewards can become resources for further rewards through other processes than social recognition. For example, Stanovich (1986) and others (e.g. Bast and 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 and Flynn (2001) call "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. This fact must be due to environmental change, and yet the magnitude of growth is inconsistent with the until-recently standard heritability estimates first formulated by Jensen (1974). 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 the contemporary literature that a variety of social factors cause an individual's local environment to be selected on the basis of IQ. This selection induces a correlation between genetic and environmental factors. This matching between IQ and the environment produces further cognitive gains in the subsequent period, and then further environmental selection, 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 (9). Their model is made more sophisticated still by 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 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 over-time evolution of population distributions.

Other scholars have also theorized social multiplier effects in models of behavior. Any network model of contagion which posits that behavior is influenced by the behavior of peers has this form. For example, in their review of the "neighborhood effects" literature, Jencks and Mayer (1990) noted that neighborhoods can get caught into 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 (1997) 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 characterize not just the mechanism that is producing the effect, but would also address the mechanism that turns the process off, which, in the context of equations (7) or (9) 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, the impact of position may operate interactively in connection with other attributes of the individual in question, and the nature of this interaction may reduce the size 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 into 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. As a general observation, it is likely that diminishing returns will occur whenever the "production function" involves other factors besides the position in question or the resources that are made available on the basis of position. External structural factors may also be the source of a shutdown of the CA mechanism. For example, position on a career ladder only provides an independent advantage through the early or middle stages of a process; career ladders have a finite length and (as in the Venture tube model of Stewman and Konda 1983) the value of position usually declines at a certain point on the hierarchy because the ratio of available positions at the next higher level falls 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 the nurturing of the future stars in the next generation (Zuckerman 1988). Still another possible mechanism is the commission of career mistakes by stars and superstars, which cast doubt on future productivity or 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 and

Singer 1991) can shut down a CA process. Until sociologists better understand the processes that produce CA, our knowledge of how these processes get shut down will remain more illustrative than systematic.

4 Evidence for CA in Specific Life Course Domains

4.1 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, though clearly the distribution of completed schooling does not show the strong distributional skew that is typical of CA processes. Even though status and positional advantages clearly play an important role in school performance, the lack of right skew in distributions of school grades or many cognitive tests might also appear to argue against CA mechanisms. However, as Walberg et al. (1984) point out, cognitive achievement can also be measured via "absolute measures" of cognitive knowledge or skill, and these measures are likely to have 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 where the main empirical evidence consists of evidence of 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 and 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 literature in the sociology and economics of education addresses whether there are "school effects" in the process of education at all levels of the educational process from pre-school to university. Because academic performance is an important determinant for entry into high quality post-graduate schools, colleges, high schools, and even elementary schools, the existence of school effects would seem to imply that access to a high quality elementary school confers a positional advantage for entry into a high quality high school, which confers a positional advantage for entry into a high quality college, which confers a positional advantage for entry into a high quality graduate or professional school. School effects could in principle imply a CA process, but, as Attewell recently showed (Attewell 2001), high quality schools might 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 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 that would be driven by this mechanism.

Much of the focus on CA in the educational literature has been on tracking and its effects, and specifically whether tracking produces growing inequality in educational outcomes over time. Some scholars have found that tracking increases the variance in academic outcomes (Gamoran and 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 and 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 in Blau and Duncan (Blau and 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 has grown more murky as the scholarly literature has continued to debate whether tracking systems do in fact produce growing inequality over time relative to non-tracked 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 over-arching tracks, makes it more difficult both to theorize about CA mechanisms in the context of school effects, and to identify them empirically.

4.2 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 (Entwisle and Alexander 1988; Duncan, Yeung, Brooks-Gunn, and Smith 1998) has the character of a cumulative exposure process. Krein and Beller (1988) and McLanahan and 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, Brooks-Gunn, and Klebanov (1994) have shown that children who live in persistent poverty have decreased IQ and increased behavior problems at least up to age 5 relative to children who experienced only short poverty spells or to children who never experienced poverty. Meanwhile, a large 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, though it has been difficult even in an experimental context to produce a powerful test of this conjecture (Brookes-Gunn and Duncan 1994; Sampson, Morenoff, and Gannon-Rowley 2002; Kling and Liebman 2004), and we are unaware of research that establishes the extent to which neighborhood effects have the character of a cumulative exposure process.

4.3 Work and Careers

The early research on CA in sociology by Merton and his associates 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 might drive whatever CA processes are found to exist. However, Rosenfeld's early 1990s review (Rosenfeld 1992) of this literature omits any explicit mention of cumulative advantage, 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 and Borjas 1980). Previous unemployment arguably "scars" a workers and creates future unemployment because it reduces a worker's human capital, because it reduces a worker's psychological readiness for work, and because it makes a person less attractive to prospective employers. The early micro-dynamic literature on unemployment demonstrated an empirical connection between previous unemployment and future unemployment risk (Heckman and 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 are affected by the resources or liabilities that workers obtain in their early careers, and this resource/liability acquisition occurs partly for reasons that are 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 (though see Gangl (2004) for an important exception).

Early sociological work on internal labor markets contained theories that contained CA mechanisms. As was already noted, Rosenbaum (1979) proposed that careers within corporations follow a path-dependent CA process in which the outcomes of early promotion tournaments have an impact on subsequent events. This hypothesis still has great theoretical appeal, though his empirical research did not clearly distinguish the extent to which unequal promotion rates were due to the outcomes of early promotion tournaments or whether they are instead driven by quality differences among workers. Althausser (1989) similarly argued that internal labor markets linked jobs together 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 and O’Rand’s recent work (Elman and O’Rand 2004) 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 attribute this fact to the cumulative career advantages that early bloomers enjoy.

Bielby and 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 our equations (11) or (13) 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 screen writers and female screen-writers, and this omission creates ambiguity about the correct interpretation of their results. While a gender-dependent CA process may be the cause of the observed growing inequality, it is also possible that the growing gender gap over the life course is 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, the Coles, Zuckerman, and Allison and colleagues. Long and Fox’s (1995) literature review along with Xie and Shaumann’s (2004) recent book demonstrate considerable differences in the rewards of male and female scientists. Both Long and Fox and Reskin and Hargens (1979) found evidence that women’s disadvantage grows during the early career as a result of CA processes that magnify their early disadvantages, though Long and Fox evaluate this evidence as more tentative than conclusive. Meanwhile, Long (1992) found sex differences in rates of rank promotion even after controlling for productivity, which suggest 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 scientists’s working life. Xie and 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. While 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; Huber 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 the model of Rosen. Huber's evidence is persuasive though 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 might be consistent with Long's finding that gender differences in science only grow 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 (Xie and Shauman 2004; Jacobs 1996), while Petersen and Morgan (1996) have shown that sex segregation at the combined occupation-establishment level explains nearly 90% of gender wage inequality. Petersen's (1992) recent study of a large corporation in the service sector found considerable gender differences in starting pay even when education was controlled (though he could not rule out that these differences were related to gender differences in prior experience). The CA thesis – namely that initial differences would 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 has a major impact on the effect of this "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 argued that social hierarchies emerge through a set of four factors: intrinsic quality differences amongst people; a presumption that an individual's welfare is enhanced more by associating with individuals of higher quality than by associating with individuals of lower quality; a presumption that quality is not fully observable; and a presumption that people infer quality based upon the attention that they see others paying to specific alters. 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. While both Merton and Gould's model 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 from 1961 and earlier; more recent evidence for larger groups is needed for a stronger test.

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

4.4 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 and Laub 1997; Sampson and Laub 2003a). 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 and Laub (1997; 2003a) 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, while Sampson and Laub see external shocks continuing to shift individuals between different career trajectories later in the life course (Sampson and Laub 2003b). In terms of equation (9), these theories disagree over the age pattern of the γ parameters, with Moffitt seeing young age shocks of primary importance and Sampson and 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. While 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 cumulative advantage, including such topics as the cumulative impact on health of exposure to risk factors such as education (Ross and Wu 1996) or obesity (Ferraro and Kelley-Moore 2003), or the cumulative effects of discrimination on race differences in health (Krieger 1993). Singer et al. (Singer, Ryff, Carr, and Magee 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. While Sampson and Laub's (2003a) recent work is a good example of a biographical approach to the study of careers, the specifics of the Singer et al. strategy for studying cumulative advantage 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 and Duncan's 1967 book, was recently reviewed by a National Research Council Panel as part of a broader effort to measure racial discrimination (Blank, Dabady, and Citro 2004). Some authors focused on the accumulation of negative direct effects of being black as found in a cumulative exposure model (Williams and 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 and Neal 1998). While the literature on CA as applied to discrimination is growing, the editors of the National Academy 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, Dabady, and Citro 2004, p. 224).

5 Conclusion

Cumulative advantage as either an explicit or implicit inequality-generating process has pervaded the sociological literature for several decades. Aggregate data suggests that CA is at work in many social domains. Thus, Allison and 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 since Ph.D., and reached .69 in their oldest age stratum. Using synthetic cohort data on Protestant ministers, Broughton and Mills (1980) estimated that the Gini coefficient for minister's 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 and Glennie (1999) found that the relative educational deflection caused by 10th-grade position 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 and 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 experience and over 40% at fifteen years of experience. Little is currently known about the reasons for these different growth rates in inequality.

In order to make progress, there is a need for more explicit attention to mechanisms, a need for more formal theorizing, and a need for greater attention to methodological issues in the specification and testing of theories. In an effort to rectify this shortcoming, we have called attention to the distinction between strict CA and the CA theories that are primarily focused on over-time changes in between-group inequality. Strict CA theories typically take two forms, which we have referred to as simple and path-dependent CA, while 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 is needed in order to move beyond the use of CA as a descriptive characterization of diverging trajectories towards 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 has not been able to examine the development and testing of CA in all subfields of sociology, to say nothing about 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, so cumulative advantage and related processes in art markets (Fox and Kochanowski 2004; Salganik, Dodds, and J.Watts 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 and Flynn 2001), and other applications may provide inspiration for new approaches and greater understanding of inequality-generating processes in society.

References

- Adler, M. (1997). Stardom and talent. *American Economic Review* 75, 208–212.
- Allison, P. D., J. Long, and T. Krauze (1982). Cumulative advantage and inequality in science. *American Sociological Review* 47, 615–625.
- Allison, P. D. and J. A. Stewart (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review* 39, 596–606.
- Allmendinger, J., H. Brueckner, and E. Brueckner (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 A. B. Atkinson and M. H. Rein (Eds.), *Age, Work, and Social Security*, pp. 188–223. St. Martin's Press:

New York.

- Althausser, R. P. (1989). Internal labor markets. *Annual Review of Sociology*. 15, 143–161.
- Arthur, W. and K. Arrow (1994). *Increasing Returns and Path Dependence in the Economy*. Ann Arbor: University of Michigan Press.
- Attewell, P. (2001). The winner-take-all high school: Organizational adaptations to educational stratification. *Sociology of Education* 74, 267–295.
- Baltagi, B. H. (1995). *Econometric Analyses of Panel Data*. New York: Wiley.
- Bast, J. and P. Reitsma (1998). Analyzing the development of individual differences in terms of matthew effects in reading: Results from a dutch longitudinal study. *Developmental Psychology* 34, 1373–1399.
- Becker, G. (1985). Human capital, effort, and the sexual division of labour. *Journal of Labour Economics* 3, 33–38.
- Bielby, D. D. and W. Bielby (1996). Women and men in film: Gender inequality among writers in a culture industry. *Gender and Society* 10, 248–270.
- Blank, R. M., M. Dabady, and C. F. Citro (2004). *Measuring Racial Discrimination*. Washington: The National Academies Press.
- Blau, F. D. and L. M. Kahn (1992). The gender earnings gap: Learning from international comparisons. *American Economic Review, Papers and Proceedings* 82, 533–538.
- Blau, P. and O. Duncan (1967). *The American Occupational Structure*. New York: Wiley.
- Brookes-Gunn, J. and G. Duncan (1994). *Consequences of Growing Up Poor*. New York: Russell Sage.
- Broughton, W. and E. W. Mills (1980). Resource inequality and accumulative advantage: Stratification in the ministry. *Social Forces* 58, 1289–1301.
- Cole, J. R. and S. Cole (1973). *Social Stratification in Science*. Chicago: The University of Chicago Press.
- Cole, J. R. and B. Singer (1991). A theory of limited differences: Explaining the productivity puzzle in science. In J. C. H[arriet] Zuckerman and J. Bauer (Eds.), *The Outer Circle: Women in the Scientific Community*, Chapter 13, pp. 277–340. New York: W.W. Norton and Co.
- Cole, S. (1970). Professional standing and the reception of scientific discoveries. *American Journal of Sociology* 76, 286–306.
- Cole, S., J. Cole, and G. Simon (1981). Chance and consensus in peer review. *Science* 214, 881–886.
- Conley, D. (1999). *Being Black, Living in the Red: Race, Wealth and Social Policy in America*. Los Angeles: University of California Press.
- Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M. L. Kamil, P. B. Mosenthal, and P. D. Pearson (Eds.), *Handbook of reading research, Vol. 2*, pp. 512–538. New York: Longman.

- Dannefer, D. (1987). Aging as intracohort differentiation: Accentuation, the matthew effect and the life course. *Sociological Forum* 2, 211–236.
- Dannefer, D. (2003). Cumulative advantage/disadvantage and the life course: Cross-fertilizing age and social science. *J Gerontol B Psychol Sci Soc Sci* 58, 327–337.
- de S. Price, S. J. (1965). Networks of scientific papers. *Science* 149, 510–515.
- de S. Price, S. J. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society of Information Sciences* 27, 292–306.
- Dickens, W. T. and J. R. Flynn (2001). Heritability estimates versus large environmental effects: The iq paradox resolved. *Psychological Review* 108, 346–369.
- DiPrete, T. A. (1981). Unemployment over the life cycle: Racial differences and the effect of changing economic conditions. *American Journal of Sociology* 87, 286–307.
- Duncan, G. J., J. Brooks-Gunn, and P. Klebanov (1994). Economic deprivation and early-childhood development. *Child Development* 6, 296–318.
- Duncan, G. J., W.-J. J. Yeung, J. Brooks-Gunn, and J. Smith (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review* 63, 406–423.
- Ellwood, D. (1982). Teenage unemployment: Permanent scars or temporary blemishes? In R. B. Freeman and D. A. Wise (Eds.), *The youth Labour Market Problem: Its Nature Causes and Consequences*, pp. 349–390. Chicago: University of Chicago Press.
- Elman, C. and A. M. O’Rand (2004). The race is to the swift: socioeconomic origins, adult education, and wage attainment. *American Journal of Sociology* 110, 123–160.
- Entwisle, D. and K. Alexander (1988). *Achievement in the first 2 years of school : patterns and processes*. Chicago: University of Chicago Press.
- Faia, M. (1975). Productivity among scientists: a replication and elaboration. *American Sociological Review* 40, 825–829.
- Ferraro, K. F. and J. A. Kelley-Moore (2003). Cumulative disadvantage and health: Long-term consequences of obesity? *American Sociological Review* 68, 707–729.
- Fox, M. A. and P. Kochanowski (2004). Models of superstardom: An application of the lotka and yule distributions. *Popular Music and Society* 27, 507–522.
- Frank, R. H. and P. J. Cook (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. *American Sociological Review* 57, 812–828.
- Gamoran, A. and R. D. Mare (1989). Secondary school tracking and educational equality: Compensation, reinforcement, or neutrality? *American Journal of Sociology* 94, 1146–1183.
- Gangl, M. (2004). Welfare states and the scar effects of unemployment: A comparative analysis of the united states and west germany. *American Journal of Sociology* 109, 1319–1364.
- Garasky, S. (1995). The effects of family structure of educational attainment: do the effects vary by the age of the child? *American Journal of Economics and Sociology* 54, 89–105.
- Gould, R. (2002). The origins of status hierarchies: A formal theory and empirical test. *American Journal of Sociology* 107, 1143–1178.
- Greenwald, A. and M. Banaji (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? *Sociological Inquiry* 73, 575–594.
- Heckman, J. J. and G. Borjas (1980). Does unemployment cause future unemployment? definitions, questions, and answers from a continuous time model of heterogeneity and state dependence. *Economica* 47, 247–283.
- Heckman, J. J., L. J. Lochner, and P. E. Todd (2003). Fifty years of mincer earnings regressions. Iza discussion paper no. 775, may 2003.
- Hirsch, F. (1976). *Social Limits to Growth*. Cambridge: Harvard University Press.
- Huber, J. C. (1998). Cumulative advantage and success-breeds-success: The value of time pattern analysis. *Journal of the American Society for Information Science and Technology* 49, 471–476.
- Huber, J. C. (2002). A new model that generates lotka’s law. *Journal of the American Society for Information Science and Technology* 53, 209–219.
- Jacobs, J. (1996). Gender inequality and higher education. *Annual Review of Sociology* 22, 153–185.
- Jencks, C. and S. E. Mayer (1990). The social consequences of growing up in a poor neighborhood,. In J. L[awrence] E. Lynn and M. G. McGreary (Eds.), *Inner-City Poverty in the United States*, pp. 111–186. Washington: National Academy Press.
- Jensen, A. (1974). *Educability and Group Differences*. New York: Harper and Row.
- Johnson, R. W., U. Sambamoorthi, and S. Crystal (2002). Gender differences in pension wealth and their impact on late-life inequalit. In S. Crystal, D. Shea, and K. W. Schair (Eds.), *Annual Review of Gerontology*

- and Geriatrics. *Focus on Economic Outcomes in Later Life: Public Policy, Health, and Cumulative Advantage*, pp. 116–137. Springer Publishing Company: New York.
- Johnson, W. R. and D. Neal (1998). Black skills and the black-white earnings gap. In C. Jencks and M. Phillips (Eds.), *Social Inequality*, pp. 480–498. Washington, DC: Brookings Institution Press.
- Kerckhoff, A. (1993). *Diverging Pathways: Social Structure and Career Deflections*. Cambridge: Cambridge University Press.
- Kerckhoff, A. C. and E. Glennie (1999). The matthew effect in american education. *Research in Sociology of Education and Socialization*. 12, 35–66.
- Kerin, R., P. Varadarajan, and R. Peterson (1992). First-mover advantage: A synthesis, conceptual framework, and research propositions. *Journal of Marketing* 56, 33–52.
- Kling, J. and J. Liebman (2004). Experimental analysis of neighborhood effects on youth. Ksg faculty research working papers series rwp04-034.
- Krein, S. and A. Beller (1988). Educational attainment of children from single parent families: differences by exposure, gender, and race. *Demography* 25, 221–234.
- Krieger, N. (1993). Epidemiology and the web of causation: Has anyone seen the spider. *Psychological Review* 100, 674–701.
- Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces* 71, 159–178.
- Long, J. S. and M. F. Fox (1995). Scientific careers: Universalism and particularism. *Annual Review of Sociology* 21, 45–71.
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* 16, 317–323.
- Lucas, S. (1999). *Tracking Inequality: Stratification and Mobility in American High Schools*. New York: Teacher’s College Press.
- Lucas, S. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. *American Journal of Sociology* 106, 1642–1690.
- Mayer, K., I. Maas, and M. Wagner (1999). Socioeconomic conditions and social inequalities in old age. In P. B. Baltes and K. U. Mayer (Eds.), *The Berlin Aging Study: Aging from 70 to 100*, pp. 227–258. Cambridge University Press: Cambridge and New York.
- McLanahan, S. and G. Sandefur (1994). *Growing Up with a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard University Press.
- Merton, R. K. (1973b). The normative structure of science. In N. Storer (Ed.), *The Sociology of Science*, pp. 267–278. Chicago: The University of Chicago Press.

- Merton, R. K. (1973a). The matthew effect in science. In N. Storer (Ed.), *The Sociology of Science*, pp. 439–59. Chicago: The University of Chicago Press.
- Merton, R. K. (1988). The matthew effect ii: Cumulative advantage and the symbolism of intellectual property. *Isis* 79, 606–623.
- Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- Mitzenmacher, M. (2003). A brief history of generative models for power law and lognormal distributions. *Internet Math* 1, 226–251.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review* 100, 674–701.
- Newman, M. (2003). The structure and function of complex networks. *SIAM Review* 45, 167–256.
- Newman, M. E. J. (2005). Power laws, pareto distributions, and zipf’s law. *Complexity Digest* 2005.02, 1–27.
- Oliver, M. L. and T. M. Shapiro (1995). *Black Wealth/White Wealth: A New Perspective on Racial Inequality*. New York: Routledge.
- O’Rand, A. M. (1996). The precious and the precocious: Understanding cumulative disadvantage and cumulative advantage over the life course. *The Gerontologist* 36, 230–238.
- O’Rand, A. M. (2002). Cumulative advantage theory in life course research. *Annual Review of Gerontology and Geriatrics* 22, 14–20.
- Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology* 108, 937–975.
- Petersen, T. and L. A. Morgan (1996). Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *American Journal of Sociology* 101, 329–361.
- Petersen, T. and I. Saporta (1992). The opportunity structure for discrimination. *American Journal of Sociology* 71, 159–178.
- Raftery, A. E. (1994). Bayesian model averaging. *Sociological Methodology* 24, 111–196.
- Reskin, B. F. and L. Hargens (1979). Scientific advancement of male and female chemists. In K. Lutterman (Ed.), *Discrimination in Organizations*, pp. 100–122. San Francisco: Jossey-Bass.
- Robins, J. M., D. Blevins, G. Ritter, and M. Wulfsohn (1993). G-estimation of the effect of prophylaxis therapy for pneumocystis carinii pneumonia on the survival of aids patients. *Epidemiology* 3, 319–336.
- Rosen, S. (1981). The economics of superstars. *American Economic Review* 71, 845–858.
- Rosenbaum, J. E. (1979). Tournament mobility: Career patterns in a corporation. *Administrative Science Quarterly* 24, 220–241.

- Rosenfeld, R. A. (1992). Job mobility and career processes. *Annual Review of Sociology* 18, 39–61.
- Ross, C. E. and C.-L. Wu (1996). Education, age, and the cumulative advantage in health. *Journal of Health and Social Behavior* 37, 104–120.
- Salganik, M. J., P. S. Dodds, and D. J. Watts (2005, August). Cascades of culture: Exploring extreme variation and extreme unpredictability with modeling and experiments. Presented at the American Sociological Association Annual Meeting, Philadelphia, PA.
- Sampson, R. and J. H. Laub (2003a). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge: Harvard University Press.
- Sampson, R. J. and J. H. Laub (1997). A life-course theory of cumulative disadvantage and the stability of delinquency. *Developmental Theories of Crime and Delinquency: Advances in Criminological Theory* 7, 133–161.
- Sampson, R. J. and J. H. Laub (2003b). Life-course desisters? trajectories of crime among delinquent boys followed to age 70. *Criminology* 41, 555–592.
- Sampson, R. J., J. D. Morenoff, and T. Gannon-Rowley (2002). Assessing neighborhood effects: Social processes and new directions in research. *Annual Review of Sociology* 28, 443–478.
- Scholz, J. K. and K. Levine (2004). U.s. black-white wealth inequality. In K. M. Neckerman (Ed.), *Social Inequality*, pp. 895–929. New York: Russell Sage.
- Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika* 42, 425–440.
- Singer, B., C. D. Ryff, D. Carr, and W. J. Magee (1998). Linking qualitative and quantitative methods linking life histories and mental health: A person-centered strategy. *Sociological Methodology* 28, 1–51.
- Sorensen, A. B. (1970). Organizational differentiation of students and educational opportunity. *Sociology of Education* 43, 355–376.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly* 21, 360–407.
- Stewman, S. and S. L. Konda (1983). Careers and organizational labor markets: demographic models of organizational behavior. *American Journal of Sociology* 88, 637–685.
- Sutton, J. (1997). Gibrat’s legacy. *Journal of Economic Literature* 35, 40–59.
- Walberg, H. J., B. F. Strykowski, E. Roval, and S. S. Hung (1984). Exceptional performance. *Review of Educational Research* 54, 87–112.
- Waldfoegel, J. (1998). The family gap for young women in the united states and britain: Can maternity leave make a difference? *Journal of Labor Economics* 16,3 (July 1998): 505-545 16, 505–545.

- Williams, D. R. and H. Neighbors (2001). Racism, discrimination, and hypertension: Evidence and needed research. *Ethnicity and Disease* 11, 800–816.
- Wilson, E. O. (1975). *Sociobiology: the New Synthesis*. Cambridge: Harvard University Press.
- Xie, Y. and K. Shauman (2004). *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.
- Yang, R. (1999). Race, status attainment, and depression: Intervening effects of consequential life events. *Race and Society* 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 C. Mongardini and S. Tabboni (Eds.), *Robert K Merton and Contemporary Sociology*, pp. 139–162. New York: Transaction Publishers.