Journal of Quantitative Criminology, Vol. 13, No. 3, 1997

Generality, Continuity, and Change in Offending

Raymond Paternoster,[1,5] Charles W. Dean,[2] Alex Piquero,[3] Paul Mazerolle,[4] and Robert Brame[1]

A number of criminological theories make either implicit or explicit predictions about the empirical relationship between prior and future offending behavior. Some argue that time-stable characteristics such as criminal propensity should account for any positive correlation between past and future criminal behavior for all individuals. Others contend that the positive association between offending behavior at different points in time are partly causal and partly spurious. Still others anticipate that different patterns will emerge for different groups (distinguished by their ciminal propensity) of individuals. Using a longitudinal data set comprised of 848 training school releasees, we test various hypotheses emanating from these different theoretical perspectives. The results indicate that (1) both stability and change have causal implications for one's offending behavior and (2) with but one exception, these effects do not vary between high and low criminal propensity groups.

KEY WORDS: generality; continuity; predictions; offending behavior; change.

1. INTRODUCTION

It has been noted that theoretical controversies in criminology rarely involve disputes about facts; instead they often involve disputes about the *interpretation* of facts (Hirschi and Gottfredson, 1995). One of the most consistent and resilient facts in criminology is the positive correlation between past and future criminal behavior. While there is no dispute about the existence of this fact, there is a great deal of disagreement about what it means. The purpose of this paper is to review briefly some of the prominent

[1]Department of Criminology and Criminal Justice, University of Maryland, College Park, Maryland 20742.
[2]Department of Criminal Justice, University of North Carolina—Charlotte, Charlotte, North Carolina 28223.
[3]Department of Criminal Justice, Temple University, Philadelphia, Pennsylvania 19122.
[4]Division of Criminal Justice, University of Cincinnati, Cincinnati, Ohio 45221.
[5]To whom correspondence should be addressed.

0748-4518/97/0900-0231$12.50/0

explanations for the relationship between past and future crime. These different explanations, though sharing some common ground, differ in ways that have critical implications for the development of theory and public policy. We then attempt to organize these explanations by creating a framework of theory types. While we do not wish to oversimplify the content of these theories, we do think that our proposed framework highlights some important issues and suggests that theory development and testing is at an important crossroad. In devising this vocabulary, we contend that some theories are far less parsimonious than others. The primary research question addressed in this paper is whether the additional complexity of some theories is worth their added weight. We try to answer this question by testing some hypotheses deduced from these more complex theories.

2. EXPLAINING THE RELATIONSHIP BETWEEN PRIOR AND FUTURE OFFENDING

Nagin and his colleagues (Nagin and Paternoster, 1991; Nagin and Farrington, 1992a, b) have suggested two processes that could account for the strong positive correlation between past and future offending. One process implicates differences between individuals in their latent tendency to commit crime (e.g., criminal propensity, impulsivity, and present orientation). These differences are established early in life and are time stable. Whether these individual differences in criminal propensity or criminal potential are due to constitutional factors and personality traits (Wilson and Herrnstein, 1985) or differences in upbringing (Gottfredson and Hirschi, 1990), the common ground among these theories is their view that "bad apples" are created by forces that operate early in life. These "bad apples" begin offending early in life, exhibit great versatility in offending, and are more likely to offend throughout life, resulting in a positive correlation between sundry types of past and future problem behavior. Since the process generating the relationship is due to persistent differences between individuals along an underlying dimension of criminal propensity, this view can be conveniently labeled the "population heterogeneity" position (Hsiao, 1986, pp. 172–180; Nagin and Paternoster, 1991). According to this view, repeated offending among crime-prone individuals is simply a series of continuing realizations of a relatively stable underlying crime producing process. In causal terms, the positive correlation between past and future offending is spurious insofar as variation in both variables is the outcome of a common cause. For those who accept a population heterogeneity explanation, adequate controls for stable criminal propensity should cause the association between prior and future behavior to vanish.

Theories of population heterogeneity do not rule out the possibility that "change" may occur (e.g., previous offenders may find a good job and/or partner, and subsequently desist from crime).[6] What they do rule out, however, is a *causal relationship* between life changes and transitions away from offending behavior. Any observed relationship between life events and desistance from crime is attributed solely to self-selection. That is, once time-stable differences in criminal potential are held constant, the effects of life events of all sorts (including prior episodes of offending behavior) should have no effect on subsequent offending. According to this view, "bad apples" simply do not secure meaningful, stable jobs or emotionally warm and rewarding social relationships. Not quite so bad "bad apples," however, may a priori decide to secure jobs, find meaningful relationships, and cease offending. In other words, the moderate "bad apples" sort themselves into somewhat better situations than their extreme "bad apple" counterparts. Consequently, in a pure population heterogeneity explanation, the very intention of the moderate "bad apples" to make positive changes constitutes evidence that they are not really the worst of the "bad apples," that the "reform" is merely apparent, and that the underlying process is one of self-selection, rather than one of cause and effect (Hirschi and Gottfredson, 1995).

A second explanation for the observed correlation between past and future offending argues that there is a genuine causal link between past and future criminal behavior. This link, commonly referred to as the *state-dependent* effect, suggests that the commission of criminal acts reduces inhibitions and/or strengthens motivations to commit crime. There are any number of specific mechanisms that can account for this state dependent effect. For example, the commission of crimes with impunity may weaken persons' perceptions of the certainty of punishment, weaken their bond to conventional others or their commitments to conventional roles, strengthen their affiliation with deviant others leading to increased social reinforcement for crime and more criminal opportunities, or result in labeling and one's exclusion from the normal routines of life. Clearly, state dependent effects are compatible with numerous extant criminological theories. The common ground they share is the expectation that the commission of criminal acts has a causal effect on the commission of subsequent criminal behaviors. Contrary to heterogeneity explanations, life events (e.g., marriage, job,

[6]Some population heterogeneity theories are compatible with the notion that there may be some *absolute* change in criminal propensity and associated characteristics over time but preclude the possibility of *relative* change (i.e., changes in relative standing within the population in the likelihood of committing crime) (Gottfredson and Hirschi, 1990). Other population heterogeneity theories seem to be more hostile to the idea of absolute change as well as relative change (Wilson and Herrnstein, 1985).

changes in friendship groups, and involvement in criminal behavior) can have a genuine causal effect on future behavior. In other words, committing criminal acts can make things worse for offenders if they desire a more conventional life; moreover, changes in the life situations of those who have offended in the past may subsequently lead to long-term (Sampson and Laub, 1993) or short-term (Horney *et al.*, 1995) redemption.

It can be appreciated that heterogeneity and state-dependent theories offer very different understandings of the processes that result in a positive correlation between past and future criminality. In their pure form, heterogeneity theories (like those of Gottfredson and Hirschi and of Wilson and Herrnstein) can be construed as static processes (Nagin and Farrington, 1992a) because they disallow any genuine causal effect of prior behavior on future behavior. They also discount the possibility that experiences or major life events can alter one's destined pattern of offending after the die is cast.[7] State-dependent theories, in contrast, are dynamic in that they both allow for a causal effect of prior on future behavior and are compatible with the idea that changes in life circumstances can affect one's proclivity to offend.

Thus far, we have drawn a clear distinction between static population heterogeneity theories, on the one hand, and dynamic state dependent theories, on the other. There is now in the literature an ongoing debate as to whether or not the added complexity of a dynamic theory is superior to the simpler static model (Gottfredson and Hirschi, 1990; Sampson and Laub, 1993, 1995; Laub and Sampson, 1993; Hirschi and Gottfredson, 1995). As we argue below, however, there is more than one dimension of complexity to consider in current debates over criminological theory.

Though differing in the causal importance attached to prior criminal behavior and the theoretical importance of change, the kinds of heterogeneity and state-dependent theories we have alluded to above share a common theme—they essentially argue that there is a common explanation of

[7]It should be noted that in Gottfredson and Hirschi's theory, low self-control is the product of failed early childhood socialization practices. Until the end of early childhood, then, interventions that improve those socialization practices should be effective at modifying one's level of self-control. This "window of opportunity" closes around age 8, however, and relative standing on self-control within any given age cohort after that time is predicted to be resistant to intervention. Thus, their theory is dynamic up to about age 8 and is static thereafter.

Gottfredson and Hirschi also claim that self-control does not explain all variation in offending; even those with low self-control need opportunities to commit crimes. In the limiting case of a complete absence of opportunity, they would argue that offending is physically impossible. We make two observations about this: (1) the opportunity to commit some kind of crime is certainly ubiquitous, particularly when one considers that offenders are often versatile creatures (Blumstein *et al.*, 1986, p. 5); and (2) some criminal behaviors may open up new opportunities for crime while closing off those for noncrime, and some noncriminal behaviors may open up opportunities for a conventional life while closing off those for a criminal life. The latter observation, that past crime has important causal consequences for involvement in crime in the future, is of central interest to us in this paper.

crime that applies to all members of the population. As such, they share the characteristic of being general theories of crime. For example, Gottfredson and Hirschi (1990) clearly argue that all crime and analogous acts can be attributed to variations in self-control and available opportunities. Similarly, Wilson and Herrnstein (1985) suggest that what they call serious predatory crime can be linked to a time-stable cluster of characteristics including impulsivity, poor conditionability, and a weak conscience. In addition to these obviously general theories, many theories of crime that are dynamic also posit one or a few causes of crime. For Agnew (1992) it is strain, for Akers (1985) it is differential reinforcement, for Lemert (1951, 1972) it is negative labeling, for Thornberry (1987) it is the dynamic interplay of peers and informal social controls, and for Tittle (1995) it is a balance between control exercised and control experienced.

An interesting example of how theoretical generality can unite those who disagree about the salience of dynamic influences on offending can be found in the dialogue of Gottfredson and Hirschi (1990) and Sampson and Laub (1993). Indeed, their points of agreement are as revealing as their points of contention. To Sampson and Laub (1993), crime can be understood as the product of informal social controls. To be sure, the specific sources of informal control (e.g., family, school, marriage, and employment) change over time. The source of informal control, however, is less important than the overall quantity and quality of control to which one is subject. In short, crime is inhibited when persons are bonded to conventional institutions. Gottfredson and Hirschi's theory, on the other hand, looks to internal controls (what they call self-control) to explain the process that restrains individuals from becoming involved in crime.

A key point of contention between the theory of Gottfredson and Hirschi and that of Sampson and Laub is whether controls on behavior, whatever their source, are subject to variation within individuals over time. Gottfredson and Hirschi argue that within-individual variation (beyond age 8) is not problematic, whereas Sampson and Laub contend that it is central to understanding why some people persist in a life of problem behaviors (including crime) while others turn their lives around. Despite their different perspectives on the importance of static and dynamic processes, however, Sampson and Laub and Gottfredson and Hirschi apparently agree that a *single* theory is sufficient to explain variation in offending behavior throughout the population. Indeed, what the static theories of Gottfredson and Hirschi and Wilson and Herrnstein have in common with the dynamic theories of Agnew, Thornberry, Akers, Tittle, and Sampson and Laub is that all offenders follow a single pathway to crime. Thus, it is possible to articulate either static or dynamic theories of crime that are general in scope—that is, the causes of offending are presumed to be the same for all persons.

It is perhaps fruitful to view theoretical generality as a type of null hypothesis (Osgood and Rowe, 1994), which asserts that only one theoretical model is needed to account for variation in offending behavior. In deference to the principle of parsimony, we should want to relax the constraint of a single theory only if the additional complexity of a "multiple pathways" theory significantly increases our understanding of the etiology of crime.

Recent work in developmental theory relaxes the assumption of generality and, in so doing, adds further complexity to the theoretical picture. Contemporary developmental theory emerged partly in response to empirical findings from the criminal career literature that suggested the existence of distinct types of offenders such as early starters, late starters, persisters, desisters, occasionals, and chronics (Blumstein *et al.*, 1986, 1988; Barnett *et al.*, 1985, 1989, 1992). In essence, the criminal career literature led to a burgeoning taxonomy of offender types. This work did not extend beyond the classification stage, however, as criminal career researchers were better at identifying offender types than explaining them. Contemporary developmental theories entered this hiatus by rejecting the assumption of general causality and assuming that different offense patterns may need to be understood as resulting from unique causal processes (Osgood and Rowe, 1994). Rather than positing a general causal process that applies to all offenders to varying degrees, developmental theorists hypothesized that offenders differ in kind rather than only in degree (Loeber and LeBlanc, 1990). In this view, a unique causal process is required to explain the behavior of different kinds of offenders. In other words, different offenders have very distinctive sequelae.

In addition to the identification of different offender types, empirical work in the criminal career tradition also noted the simultaneous existence of continuity and change in offending. While it appeared that many offenders consistently committed crimes over long periods of time, others' involvement was more intermittent, beginning later and ending earlier. In addition to explaining behavioral continuity, then, the task taken on by developmental theorists was to explain variations in offending over the life course. The intermittent and seemingly nonrandom patterns of much criminal offending led contemporary developmental theorists to question the static position taken by some general theories of crime. Almost by definition, the developmental view grants a prominent role to changes in life circumstances as well as the experience, rewards, and consequences of prior actions in the explanation of persistent offending. Thus, in focusing attention on how changes in life circumstances affect changes in offending, developmental perspectives are friendly to the idea that changing life circumstances can have causal implications for future involvement in crime. They are also friendly to the idea that involvement in crime in the past can have a causal

impact on whether one continues to offend in the future. In short, a prominent theme of contemporary developmental theories can be found in their focus on dynamic rather than static causes of crime.

In reconciling both continuity and change in criminal offending, some prominent examples of developmental theory have adopted both a dynamic and static perspective to theory by applying each process to different offender types. For example, Moffitt and her colleagues (1993, 1994, 1996) and Patterson and his colleagues (1989, 1993) reject the assumption that there is a general theory of crime and argue for the existence of two distinct offender groups. One of these offender groups is characterized primarily by continuity in offending, the other by change. Moffitt refers to these two groups as life-course persistent and adolescent-limited offenders while Patterson refers to them as early- and late-starters.

Though differing in subtle ways, Moffitt and Patterson offer very comparable explanations for the behavior of persons who begin offending early in life and persist in their problem behaviors over the life course. Both presume that the one who begins offending early and continues offending is characterized by early behavior problems, parent–child conflict, and poor socialization. These poorly socialized children eventually fail in their family life, their school work, and their relationships with others. They "miss out on opportunities to acquire and practice prosocial alternatives at each stage of development" (Moffitt, 1993, p. 683). Essentially, some children are never effectively socialized, never learn to control or channel their antisocial behavior, and act impulsively and aggressively as children, adolescents, and adults.

It is clear from their discussion of etiology that Moffitt's life-course persistent and Patterson's early starting offenders follow a pathway to offending that resembles in many ways the pathway to low self-control developed by Gottfredson and Hirschi (1990). The early-starters and life-course persistent offenders are "bad apples" who exhibit significant deficits in early childhood socialization and rarely get back on track. This explanation for life-course continuity, like the one offered by Gottfredson and Hirschi (1990), allows for some dynamic variation during early childhood which gives way to a largely static process by the beginning of adolescence. Assuming such an explanation is correct, any correlation between dynamic variables and continued offending within this group is likely to be spurious since both will be the result of ineffective socialization.[8]

[8]Our position, that Patterson's explanation for the early-starting offender is a static one, may be controversial. This is because when they have described their theory, Patterson and his colleagues have hinted that the early starting offender may at some point be profoundly and causally affected by other deviant peers. That is, even though they possess a high propensity

Both Moffitt and Patterson adopt a more dynamic position when explaining adolescent-limited or late-starting offending. The stances of Moffitt and Patterson on the explanation of offending that is limited to the adolescent phase of the life course, however, continue to be quite similar to

for and considerable experience with antisocial behavior, early-starters are said to *require* other deviant peers to become delinquent. This was the reading of the theory adopted by Simons *et al.* (1994) in their empirical test of the Patterson model. Essentially, this position asserts that the early-starting route to delinquency is a dynamic one because delinquent peers play an important causal role in transforming a "bad child" into a delinquent. Although a detailed discussion of this issue is beyond the scope of the present paper, we would like to provide some rationale for our position. In a nutshell, for several reasons we are very skeptical of any strong causal effect for delinquent peers in the Patterson early-starter group.

First, we take Patterson *et al.* at their word when they propose that their theory is a typological one. In numerous publications they emphasize the fact that they are positing two different routes to delinquency, i.e., that there are "two very different developmental paths leading to juvenile delinquency . . . they differ both in terms of the determinants that bring them about and in terms of the long-term outcomes of following one path or the other" (Patterson and Yoerger, 1993, p. 166). The "two very different" routes to delinquency is seriously compromised if peers play an important causal role for both the early- and the late-starting offender. Patterson and colleagues are adamant that the primary causal factor in accounting for the late-starting offender is the influence of delinquent peers. Their typological distinction is clear if it is maintained that the primary causal factor for the early-starter is inadequate socialization. Ultimately, we think that their distinction becomes blurred and seriously compromised if they attribute significant causal import to delinquent peers within both the early- and the late-starting group. In this case, the distinction between the early- and the late-starters seems merely to be a matter of degree and not kind.

Second, the precise causal role played by delinquent peers within the early-starting group is not clearly specified, and when sketched out, it remains unconvincing. It should be remembered that those in the early-starting group are the product of ineffective parental socialization. More important, they already exhibit a diverse repertoire of antisocial behavior—they are destructive, aggressive, and confrontational with both other children and adults. They appear to show little regard for either the feelings or the possessions of others. In sum, they appear to us to be both fully capable and experienced in committing antisocial acts. Why do they need other antisocial peers to do things they are both skilled and experienced in doing? Patterson and colleagues seem to answer this question by stating that peers are needed to *amplify* the early starter's antisocial tendency. It is not clear what amplify means in their theory, but we assume it means that peers are needed to broaden the targets for the early-starter's antisocial acts. If all peers do is to broaden the supply of targets, is the only important role they play that of providing opportunities to commit delinquent acts? This position is neither very dynamic nor at odds with the position adopted by Gottfredson and Hirschi (1990), who note that delinquent peers may be important in providing youths with low self-control a steady state of criminal opportunities.

Finally, in describing the role played by delinquent peers among the early-starting group, Patterson and colleagues have suggested that the apparent correlation with delinquent peers may be noncausal. That is, they at times suggest that their empirical finding of a positive correlation between delinquent peers and the delinquency of the early starter may be due to a selection phenomenon—i.e., that "birds of feather flock together." Patterson and Yoerger (1993, pp. 147, 149) have noted, for example, that ". . . problem children tend to be selected as friends by isolated or rejected children" and that the early-starter is "free to shop for social groups that match their own proclivities." If antisocial children "shop around" for other antisocial children with whom to commit delinquent offenses, how is it that they play an important causal role in the continuation or amplification of offending?

	GENERAL	DEVELOPMENTAL
STATIC	Wilson & Herrnstein (1985) Gottfredson & Hirschi (1990)	Moffitt (life course persistent) Patterson (early starter)
DYNAMIC	Sampson and Laub (1993) Agnew (1992) Tittle (1995)	Moffitt (adolescent limited) Patterson (late starter)

Fig. 1. Classification scheme of criminological theory.

each other. To both, individuals who begin offending in adolescence will generally not continue to offend beyond. Within this group, however, whatever offending does occur will depend largely on life circumstances such as conflict with or divorce between parents, parental unemployment, assertions of adolescent independence, and most importantly, falling into the wrong crowd. Thus, the primary causes of offending among those who do not fit into the life course-persistent or early-start group are salient life circumstances. Logically, changes in those circumstances can causally affect the risk of involvement in offending.[9] For individuals in the adolescence-limited and late-starting groups, prior criminal acts may causally affect current and future offending by further alienating parents and conventional peers and diminishing conventional alternatives to crime. Similarly, changes in life circumstances may lead to crime desistance as partners are secured and careers or vocational training initiated. Contrary to the early-starter/life course-persisent offender, then, the behavior of the late-starter/adolescence-limited offender is predicted to be driven by a much more dynamic, state-dependent process.

It would appear from this brief discussion that current criminological theory is in disarray, even when it is directed at the single problem of accounting for the relationship between prior and current offending. We now offer an organizing theme that classifies theory along two dimensions, whether it is static or dynamic and whether it is general or developmental. Figure 1 illustrates our fourfold classification scheme.

Static/general theories of crime presume that there is a general cause of crime for all offenders and that, once the causal process has played out, change is unlikely. This is what characterizes a pure population heterogeneity theory. Examples include Wilson and Herrnstein (1985) and Gottfredson and Hirschi (1990). A dynamic/general theory is slightly more complex. It

[9]According to Moffitt (1993), some dynamic factors such as destructive reactions by others to negative behavior can begin a chain reaction that culminates in a transition from adolescence-limited offending to life course-persistent offending.

maintains the assumption of general causality, but rejects the assumption that changes in life circumstances cannot materially affect one's offending. An example of dynamic/general theory is Sampson and Laub's (1993) theory of age-graded informal social control.

A developmental theory relaxes both assumptions. It assumes that causality is not general and that different causal processes explain different offender types. The causal process involved, however, may be either static, as when it adopts the position that crime throughout the life course is due to persistent individual differences in criminal propensity produced by ineffective socialization, or dynamic, as when it adopts the position that prior crime and changes in life circumstances may causally affect current and future crime. As can be discerned, dynamic/developmental theory is far less parsimonious and more complex than dynamic/general theory, and the latter is less parsimonious than static/general theory. The question motivating this research is whether the additional complexity of some theories is necessary.

3. OVERVIEW OF PRIOR RESEARCH

Prior research in this area suggests that some complexity above and beyond that anticipated by general static models may be required. For example, Nagin and Paternoster (1991) found that both persistent individual differences in criminal propensity and prior offending behavior exerted significant effects on self-reported offending (theft) patterns in a convenience sample of South Carolina high school students. This finding was replicated within sampling error (using a broader set of delinquent behaviors as outcomes) by a recent analysis of the National Youth Survey self-report data by Paternoster and Brame (1997). In their analysis of the Glueck's long-term follow-up of Boston males, Sampson and Laub (1993) concluded that both stable individual differences and past offending patterns were important in the explanation of future offending behavior.

Nagin and Farrington (1992a) also discovered important heterogeneity effects in the Cambridge Study in Delinquent Development data comprised of a 22 year follow-up of inner-city London males. Their analysis did not, however, find the strong state-dependent effects that were evident in the Nagin and Paternoster, Paternoster and Brame, and Sampson and Laub studies. Consequently, they concluded that persistent individual differences were the predominant cause of variation in future offending behavior. Nevertheless, a close review of their empirical results suggests that state-dependent

influences, though greatly reduced after the introduction of controls for persistent unobserved heterogeneity, were still evident.[10]

In sum, a variety of studies using different outcomes, different samples, different control variables, and different model specifications seems to yield similar conclusions in at least one respect—changes in prior offending behavior appears to have important implications for one's future offending behavior that cannot be explained solely by a process of self-selection.

The extant literature also reports results that bear on the question of whether general theories are sufficient or whether typological models are more realistic. Patterson (1993), for example, predicted that parental discipline and monitoring practices would be associated with antisocial behavior by the fourth grade of school and that changes in antisocial behavior after that time would be associated with changing levels of unsupervised street time and involvement with deviant peers. Using data from his ongoing Oregon Youth Survey, Patterson (1993, pp. 914–915) estimated the parameters of this model and found that it reproduced the covariance matrix of the study variables within sampling error.

Simons *et al.* (1994) examined panel data on 177 midwestern boys from small towns in the midwestern United States. Based on the work of Patterson and his colleagues, they predicted that the process driving arrest outcomes would differ across samples stratified by age at first arrest. Their analysis uncovered evidence consistent with this prediction. Specifically, among those arrested by approximately age 14, they found that parental monitoring and disciplinary practices predicted "a coercive, noncompliant orientation" in interactions with other people. Within this "early-onset group," those possessing this oppositional/defiant orientation were found to be more likely to succumb to involvement with deviant peers and, in turn, experience more frequent future involvement with the criminal justice system (Simons *et al.*, 1994, p. 260).

Among those arrested later on (after age 14), however, Simons and his colleagues (1994, p. 262) found that effective parental monitoring and

[10]Another study using the Cambridge data conducted by Nagin and Land (1993) also estimated a state-dependent effect. However, the state-dependent effect was apparently estimated within the framework of a probit model where observed, but not unobserved, heterogeneity was controlled. Review of the tables of Nagin and Land (1993) indicates that the probit coefficient for prior convictions is very similar to the probit coefficient for prior convictions without controls for persistent unobserved heterogeneity reported by Nagin and Farrington (1992a). A recent reanalysis of the Cambridge data by Land and Nagin (1996), however, reports a strong effect for "any prior convictions" on contemporary conviction frequency with a model that incorporates a nonparametric specification of unobserved heterogeneity. Taken together, the various analyses of the Cambridge data appear to suggest that both stable individual differences and prior offending behavior have important implications for future offending behavior.

disciplinary practices were associated with lower levels of involvement with deviant peers and, in turn, lower frequencies of future contact with the criminal justice system. Unlike the early-onset group, oppositional/defiant orientation exerted no effect on future criminal justice system involvement in the later-onset group. On the basis of this evidence, Simons and his colleagues concluded that distinct processes were responsible for the future offending behavior of those who began offending early and those who began offending after age 14.

Dean *et al.* (1996) assessed whether the correlates of post-age 16 offending persistence (defined as the time to first arrest after release from training school) varied significantly between groups of North Carolina training school releasees stratified by age at first adjudication. They found that child abuse victimization and number of juvenile adjudications exerted different effects between the early (onset occurred before 12 years of age) and the late (onset occurred at or after 12 years of age) groups. Nevertheless, they also discovered that these results were sensitive to the age cutoff points used to define the groups and that when other cutting points were used the differences vanished.

More recently, Paternoster and Brame (1997) used the National Youth Survey data to examine the effects of prior offending behavior and delinquent peer exposure on participation and frequency of contemporary offending. This study was limited because onset ages were not available in the data. Nevertheless, they did use the subsample of respondents who were age 11 or 12 at the first wave of the survey to create "antisocial propensity" groups based on both attitudinal and behavioral measures. They then compared the effects of delinquent peer exposure and prior offending behavior on the outcomes between high and low antisocial propensity groups. Their analysis revealed that the effects of both variables were comparable across the groups.

Nagin and his colleagues (Nagin and Land 1993; Nagin *et al.*, 1995) have taken a somewhat different approach to the study of whether different etiological processes explain offending behavior. Using data from the Cambridge Study in Delinquent Development and a model that probabilistically assigned individuals to distinct offending "trajectories," they inductively identified four different longitudinal offending patterns. One group, the "high-rate chronics," exhibited relatively high conviction frequencies over a sustained period of time. A second group, the "adolescence-limiteds," were convicted during adolescence but tended to cease offending as they moved into adulthood. Nagin and his colleagues concluded that these two groups bore strong resemblance to the life course-persistent and adolescence-limited offenders described by Moffitt (1993). A third group that offended at low rates beginning in late adolescence and persisting into adulthood, not "anticipated" by Moffitt and Patterson, was also discovered. The fourth group was

comprised of individuals in the Cambridge study who were not convicted at all.

A key aspect of the analyses reported by Nagin and Land (1993, 1996) as well as Nagin *et al.* (1995) was an assessment of whether individuals in the various groups differed from each other on behavioral measures that did not include offending. A number of results from this assessment seem noteworthy. First, at ages 14 and 18, the adolescence-limited and high-rate chronic groups reported greater drug use, smoking, and sexual promiscuity than their low-level chronic and never convicted counterparts.

At age 32, the three groups of individuals who had been convicted at least once exhibited greater levels of fighting as well as greater levels of drug and alcohol use than the never convicted group. Interestingly, by age 32, the adolescent-limited and never-convicted groups were much more successful in the labor market than the chronic low- and high-rate groups. The adolescent-limited and never-convicted groups also appeared to have higher-quality attachment to their families by age 32 than the chronic offending groups.[11] One anomalous result was that adolescent-limited offenders did not always desist completely. Within this group, nontrivial levels of theft and embezzlement were still evident well into adulthood.

Nagin *et al.* (1995) observed, in summary, that drug use, alcohol use, theft, and rowdy behavior (e.g., fighting) appear to comprise a pattern of "circumscribed deviance" for the adolescent-limited group, while more malignant forms of deviance were in evidence for the chronic offending groups. The adolescent-limited group, thus, exhibits some deviance but generally it is the kind of deviance that is "less likely to result in official sanction or disrupt intimate attachments" (Nagin *et al.*, 1995, p. 132). Alternatively, say Nagin and his colleagues (1995, p. 136), such activities may not be regarded as serious infractions within the "working class subculture in which the AL's live."

Moffitt and her colleagues (1996) recently assessed the antisocial conduct problems of a retrospective 1972–1973 birth cohort of 457 Dunedin, New Zealand, males from age 3 to age 18. Five groups of individuals were cobbled together based on a number of measures of conduct problems within the longitudinal data: (1) individuals who engaged in extreme antisocial behavior during both childhood and adolescence (about 7% of the sample), (2) individuals who engaged in extreme antisocial behavior as adolescents but not as children (about 24% of the sample), (3) individuals who engaged

[11]In considering these findings of "attachment quality," Nagin *et al.* (1995) noted that the adolescent-limited group might still have been quite susceptible to divorce and family breakup in the years most approximate to their peak period of offending. They suggest this in the wake of their finding that divorce patterns by age 32 did not vary among the three groups of offenders.

in extreme childhood antisocial behavior but abstained from such behavior during adolescence (about 6% of the sample), (4) individuals who refrained from extreme antisocial behavior throughout the entire follow-up period (about 6% of the sample), and (5) a "normal" group that included 58% of the birth cohort and did not meet the criteria for inclusion in any of the other groups. Comparison of the early- and late-onset groups revealed that they differed in their tendencies to drop out of school, strength of attachments to their families, and personality profiles.

In sum, the various studies we have examined provide mixed signals on the propriety of studying different developmental pathways to offending behavior. While some analyses suggest that distinct causal processes may be responsible for the continued offending behavior of early- and late-onset groups, others indicate that there may be a great deal of similarity as well. Still other studies speak to the question of whether groups defined on the basis of the offending behavior they exhibit are comparable on a number of characteristics that do not include offending. These studies generally suggest that offending groups resembling those described by Moffitt and Patterson may indeed differ from each other in interesting ways. Nevertheless, they do not resolve the question of whether these differences are of causal significance and whether a theory that anticipates differences in degree rather than differences in kind can account for them as well.

4. HYPOTHESES

In this paper, we derive some hypotheses that constitute a test of some of the implications of developmental and general theory. Because of the centrality of the issue, each of the hypotheses concerns the possible causal impact of prior on current criminal offending. Starting from the premise that, other things equal, the most parsimonious theory is the preferred one, we ask whether (1) prior behavior has a causal effect on future behavior once time-stable differences in criminal propensity have been controlled and (2) whether there are unique causal pathways to offending in that prior behavior has a causal impact on the behavior of one type of offender but not another.

From the general perspective of Gottfredson and Hirschi (1990, pp. 154–168) and Wilson and Herrnstein (1985), we derive the following prediction:

> *H_1: The relationship between prior and future offending behavior will be positive, but after controlling for persistent individual differences, this positive relationship will vanish.*

This hypothesis is derived from a pure or "strong" version of heterogeneity theory, as it predicts that all continuity in offending is due to time-stable individual differences in criminal propensity. As argued by Nagin and Paternoster (1991) and Nagin and Farrington (1992a), many other theories, such as differential association, social control, and strain theory, would make a different prediction:

> *H_2: The relationship between prior and future offending behavior will be positive and will continue to be positive even after persistent individual differences have been controlled.*

Confirmation of this second hypothesis would constitute evidence that "change matters" (Laub and Sampson, 1993). In other words, changes in life circumstances may have causal significance for offending that cannot be ignored. We view the second hypothesis as a mixed model, or a "weak" form of the previous heterogeneity hypothesis, since it allows for both population heterogeneity and significant causal change. Finally, the typological theories recently advanced by Moffitt (1993) and Patterson *et al.* (1989) would make yet another prediction:

> *H_3: For those with high criminal propensity, the causal effect of prior behavior on future behavior should be trivial while, for those with low criminal propensity, the effect should be substantial.*

Confirmation of this third hypothesis would constitute evidence that continuity and change in offending may be subject to different causal processes for different types of individuals. Persistence in offending may be due to time-stable differences in criminal propensity among a small group of "hard-core" offenders (the "bad apples"), while the intermittent offending of a larger pool of offenders may be due to state-dependent processes of change and adaptation to change. We characterize this as a "category-dependent" hypothesis since population heterogeneity is predicted to dominate in one group (those who onset early), while state-dependent processes dominate in the other group (those who onset later).[12]

5. METHODS

5.1. Data

The data for this study were comprised of 838 releasees from North Carolina Division of Youth Services training schools during the 1988–1989

[12]We thank one of the reviewers of this article for suggesting these descriptive characterizations of our hypotheses.

Table I. Descriptive Statistics

Variable	Mean	SD	Range
Demographic characteristics ($N=838$)			
Sex (0=female/1=male)	0.882	0.323	0 to 1
Race (0=white/1=nonwhite)	0.549	0.498	0 to 1
Prior offending behavior ($N=838$)			
Age at first adjudication	14.307	1.537	7 to 17
Number of prior adjudications	2.189	1.650	0 to 8
Time incarcerated (months)	10.845	7.772	0 to 52
Indicator variables (0=no/1=yes) ($N=838$)			
Reported child abuse	0.195	0.396	0 to 1
Learning disability designation	0.394	0.489	0 to 1
Family receives government assistance	0.383	0.486	0 to 1
Not living in a two-parent household	0.829	0.376	0 to 1
Criminal history among other family	0.365	0.482	0 to 1
Mental deficit and assaultive behavior	0.080	0.271	0 to 1
Time-varying variables ($N=4529$)			
Time index	3.235	1.588	1 to 6
Arrested in previous year (0=no/1=yes)	0.236	0.425	0 to 1
Count of arrests in current year	0.465	0.963	0 to 11

calendar years. All subjects were between 16 and 18 years of age at the time of their release (average age=16.210 years), and the analysis includes over 98% of the individuals in this age range who were released in 1988–1989 (a small number of cases were deleted due to missing data on one or more covariates). Individuals who were younger than 16 at the time of their release were not included in the analysis because subsequent arrest information was not uniformly available for these youngsters until they reached age 16. A search of arrest records for these subjects was conducted in November 1994 and the current analysis includes information on yearly arrest counts between the date of release and the search date.

A number of background variables were collected from the official Division case files on each of the 838 individuals. Table I presents descriptive statistics for all of the variables included in the analysis. The majority of the subjects were male and nonwhite. The average age of first adjudication was slightly greater than 14 years and the mean number of prior adjudications was about 2.2, although there was considerable variation about the mean for both of these variables. The average length of confinement in the juvenile training school was just under 11 months (but the sample median was eight months).

Table I also shows that about 20% of the sample had experienced at least one incident of child abuse that had been reported to authorities. More commonly, nearly 40% of the sample had been diagnosed with some type

of learning disability and about 40% of the sample lived in a household that received some form of government assistance (e.g., welfare payments, food stamps, etc.). The vast majority of youngsters in the release cohort did not live in a two-parent household and slightly over one-third had parents and/or siblings who possessed a criminal record. A small minority of releasees had been diagnosed with a severe disorder that consists of both mental or intellectual deficiencies and serious episodes of assaultive behavior. Children classified into this category are referred to as "Willie M" cases and receive intensive treatment or services in areas where particular needs are identified (details are discussed by Weisz *et al.*, 1990). These data were used to create a multiwave panel that varied from 4 to 6 years in length. Thus, some individuals were followed for as few as 4 full years ($n=47$), others were followed for 5 years ($n=405$), while still others were followed for 6 full years ($n=386$).

A key component of the tests performed in this analysis is the definition of high and low criminal propensity groups. Unfortunately, there is no single widely accepted definition of criminal propensity (Simons *et al.*, 1994, p. 267; Moffitt, 1994, pp. 28–29; Patterson *et al.*, 1989, p. 331, 1992, pp. 335–341; Patterson, 1993, pp. 913–914; Nagin *et al.*, 1995, p. 112; Nagin and Farrington, 1992b, pp. 504–506). Nevertheless, Nagin and Farrington (1992b) argue that "[i]t is well documented that early onset of delinquent behavior is predictive of more persistent future offending (p. 503; see also Farrington *et al.*, 1990, pp. 290–293; Patterson *et al.*, 1992, pp. 339–340; Moffitt, 1993, p. 694; Sampson and Laub, 1995, pp. 149–150; Horney *et al.*, 1995, p. 671). Thus, while criminal propensity itself may be difficult to define there seems to be considerable agreement that an early onset of offending behavior is an indicator of relatively high propensity to engage in crime and delinquency. For example, according to Gottfredson and Hirschi (1990), those who initiate offending early in life exhibit comparatively low levels of self-control (i.e., high criminal propensity) throughout the life course (see also Gottfredson and Hirschi, 1986, p. 223, 1988, pp. 39–40). Similar themes are expressed in the work of other theorists as well (see, e.g., Moffitt, 1993, p. 694, 1994, p. 46; Loeber and LeBlanc, 1990, pp. 394–395).

In short, onset age appears to be a useful index for criminal propensity. While it is not yet clear whether onset age is itself causally important, all seem agreed that early onset of criminal behavior is not a harbinger of good things to come.[13] As such, those with early onset ages are generally viewed as being at highest risk for long-term chronic offending patterns, while those with later onset ages are generally viewed as being at somewhat lower risk

[13]In sum, while it may seem that the age at which one onsets offending is a mere calendrical event, it proxies for latent criminal propensity.

for these malignant outcomes. For theorists such as Gottfredson and Hirschi and Sampson and Laub, factors that predict future problem behavior in early- and late-onset groups are hypothesized to be the same—those who onset late just possess them to a lesser degree. For taxonomic theorists such as Moffitt and Patterson, however, the factors that lead to offending are actually thought to be very different between the two groups. The question for our research is whether the difference between these groups is one of degree or kind.

Finally, the data facilitate calculation of the number of arrest events that occurred during each full year of the follow-up period (see Table I). They also provide a reasonable base of information upon which to address the questions discussed above. First, the analysis attempts to identify the effect of prior arrests on future arrests under the restrictive assumption that a single process generated the data for all 838 individuals. Then, using varying onset ages as cutting points, this equal-effects constraint is relaxed to test the hypothesis that prior arrests exert similar effects on subsequent arrests regardless of one's membership in an early- or late-onset group.

5.2. Analysis Methods

The dependent variable is a count of arrest events within each year of the panel. Thus, a statistical model that takes the discrete features of the dependent variable and the lack of independence of multiple observations on the same individual into account is required. As Land (1992) argues, it is conventional to assume that data such as these are generated by a Poisson process with discrete density

$$\text{pr}(n_{it}) = e^{-\lambda it}\lambda_{it} y_{it} / y_{it}!, \qquad y_{it} = 0, 1, \ldots \tag{1}$$

where λ is the Poisson parameter and is allowed to depend on vectors of time-stable and time-varying covariates, by the equality

$$\lambda_{it} = \exp(x_i'\theta + t\gamma + z_{it}\eta) \tag{2}$$

where

$$z_{it} = \begin{cases} 1, & \text{if individual } i \text{ was arrested in previous year} \\ 0, & \text{otherwise} \end{cases}$$

t is a time index which ranges from 1 to 6 and indexes one's temporal position within the follow-up period, and x_i is a vector of covariates that do not change over time (i.e., sex, race, etc.).

In the special case where λ adequately captures both the mean and the variance of the outcome variable y_{it}, then the parameters of Eq. (2) can be estimated by maximum-likelihood methods for Poisson regression. In

regression analyses of crime count outcomes (including the current analysis), however, the mean-variance equality is often violated by substantial over dispersion in the data (see Table I). A common remedy for this problem is to augment Eq. (2) to include a random disturbance term,

$$\lambda_{it} = \exp(x_i'\theta + t\gamma + z_{it}\eta + \varepsilon_{it}) \tag{3}$$

under the assumption that $\exp(\varepsilon)$ is distributed as $\Gamma(1, \alpha)$ where the parameters θ, γ, η, and α are estimated from the data by maximum-likelihood methods (see, e.g., Land *et al.*, 1996, p. 396; Hausman *et al.*, 1984, p. 922). The probability of the number of events for individual i at time t to enter the likelihood is

$$\begin{aligned} p(n_{it}) &= \int_0^\infty p(n_{it}|\lambda_{it})\, f(\lambda_{it})\, d\lambda_{it} \\ &= \left(\frac{\Gamma(n_{it}+\nu)}{n_{it}!\,\Gamma(\nu)}\right)\left(\frac{\nu}{\nu+\lambda_{it}}\right)^{\nu}\left(\frac{\lambda_{it}}{\nu+\lambda_{it}}\right)^{n_{it}} \end{aligned} \tag{4}$$

where $\Gamma(\cdot)$ is the gamma function, which is easily evaluated using most statistical computing packages and $\nu = 1/\alpha$ (see, e.g., SAS Institute, 1990, p. 551).[14] This specification provides a baseline to which more complex models that take the lack of independence among within-subject observations into account may be compared.

To confront the lack of independence in the data due to multiple observations per individual, a random-effects specification of Eq. (3) is employed. The specification for this problem is similar to that of Eq. (3) except that we now enter a different probability into the likelihood. Following Hausman *et al.* (1984, p. 927), we assume that the ratio $\nu/(1+\nu)$ is distributed as a beta random variable, $B(a, b)$, with parameters a and b estimated from the data. The probability to enter the likelihood is then given by

$$\begin{aligned} p(n_{il}, \ldots, n_{iT}) = &\left(\frac{\Gamma(a+b)\Gamma(a+\Sigma\lambda_{it})\Gamma(b+\Sigma n_{it})}{\Gamma(a)\Gamma(b)\Gamma(a+b+\Sigma\lambda_{it}+\Sigma n_{it})}\right) \\ &\times\left(\prod_{t=1}^{T}\frac{\Gamma(\lambda_{it}+n_{it})}{\Gamma(\lambda_{it})\Gamma(n_{it}+1)}\right) \end{aligned} \tag{5}$$

Maximizing the likelihood based on the probability in Eq. (5) facilitates the control of persistent unmeasured covariates of λ. A number of covariates, as discussed above, are also included in the specification to control for

[14]We note that this specification is equivalent to the NEGBIN II model discussed by Cameron and Trivedi (1986).

observable persistent individual differences that affect λ. The likelihood function and computational details for estimating the parameters of the negative binomial model and the random-effects negative binomial model are discussed by Greene (1995, pp. 570–571).[15]

6. RESULTS

The core problem to be addressed by this analysis is twofold. First, we wish to consider whether the effect of prior arrests on future arrests is spurious. This is equivalent to testing whether the parameter estimate of η in Eq. (3) is different from zero after controlling for both persistent observed and unobserved differences between individuals. The analysis was also designed to assess whether the maximum-likelihood estimate of η differs between those who were first adjudicated at an early age and those who were first adjudicated later on.

We begin by considering whether the effect of prior arrests on subsequent arrests varies between models that make very different assumptions: (1) there are no persistent unobserved differences between individuals that relate to arrest frequency, and (2) such differences are present and must be taken into account in order to assess properly the process that generates the data. Table II presents Models 1 and 2, which are estimated with controls for the full set of covariates (with the exception of age at first adjudication) described in Table I.[16] Figure 2 graphically displays the corresponding confidence intervals for the structural effects reported in Table II for Model 1 and Model 2. From these figures one can easily see the full range of the expected effect for each estimated parameter.

Model 1 is a negative binomial regression analysis based on the assumption that the specification includes no important omitted time stable covariates [i.e., estimated according to Eq. (3) above]. Model 2 relaxes this

[15]Although useful for our purposes, this specification makes assumptions about the process generating persistent heterogeneity. Incorrect specification of this process affects the efficiency of the estimator resulting in deflated standard errors and, therefore, inflated t ratios. Land and Nagin have developed a semiparametric estimator that makes no assumptions about the distribution of persistent unobserved heterogeneity in the population (Nagin and Land, 1993; Land *et al.*, 1996; Land and Nagin, 1996). Unfortunately, the software to estimate their semiparametric mixed Poisson model is not widely available. To assess whether our model provides a reasonable approximation to the observed frequencies, we generated the theoretical relative frequency distribution from Eq. (5) based on a specification that included only indicator variables for the time index. In support of the specification we have chosen, we found considerable congruence between the observed frequencies and those expected from evaluating Eq. (5). We report the results of this investigation in the Appendix.

[16]Age at first adjudication was not included as a covariate in this analysis because its values were used as the basis for dividing the sample into low and high criminal propensity groups.

Table II. Full Sample Negative Binomial Regression Models Estimated Under Assumptions of Zero Persistent Unobserved Heterogeneity (Model 1) and Nontrivial Persistent Unobserved Heterogeneity (Model 2)

	Model 1. No control for persistent unobserved heterogeneity ($N=838$)			Model 2. Persistent unobserved heterogeneity controlled ($N=838$)		
Variable	Parameter estimate	SE	$\|t\|$ ratio	Parameter estimate	SE	$\|t\|$ ratio
Constant	−1.738	0.138	12.58	−1.199	0.211	5.67
Number of juvenile adjudications	0.036	0.018	1.93	0.034	0.023	1.52
Race = nonwhite	0.304	0.064	4.71	0.421	0.075	5.64
Sex = male	0.787	0.100	7.86	0.962	0.129	7.44
Reported child abuse	0.159	0.075	2.12	0.141	0.090	1.57
Learning disability	0.024	0.064	0.37	0.073	0.075	0.97
Mental deficit/assaultive behavior	−0.019	0.121	0.16	0.090	0.131	0.68
No two-parent household	0.230	0.083	2.75	0.248	0.099	2.49
Family criminal history	0.244	0.061	4.03	0.231	0.073	3.16
Government assistance	−0.030	0.064	0.47	−0.044	0.076	0.57
Time incarcerated	0.003	0.004	0.76	−0.002	0.005	0.39
Time index (γ)	−0.183	0.019	9.75	−0.165	0.019	8.82
Arrested in previous year (η)	0.631	0.071	8.82	0.228	0.074	3.09
α	1.713	0.116	14.75			
a				6.714	1.040	6.47
b				2.884	0.544	5.30
Log-likelihood	−3942.64			−3886.14		
Number of parameter estimates	14			15		

assumption at the cost of one parameter estimate (i.e., one degree of freedom). A test of whether the log-likelihoods of Models 1 and 2 are different is, therefore, a test of the plausibility of the exhaustive specification assumption (i.e., that there is no persistent unobserved heterogeneity). Twice the difference between these log-likelihood values is distributed as a chi-square random variable with one degree of freedom, and in this case, the test statistic is statistically significant [$\chi^2_{(1)} = 113.0$, $P < 0.05$]. On the basis of this result, the hypothesis of no persistent unobserved heterogeneity seems to be inconsistent with the data.

The next item of interest in Table II is the variable effect of η under the assumption of no time-stable individual differences (Model 1) and after that assumption has been relaxed (Model 2). The η term, it will be remembered, measures the effect of prior arrests on future arrests. Under Hypothesis 1, we would expect this to be zero once measured and unmeasured sources of persistent heterogeneity have been controlled. In particular, the quantity $\delta = \eta_{M1} - \eta_{M2}$ is of interest and an examination of Table II indicates

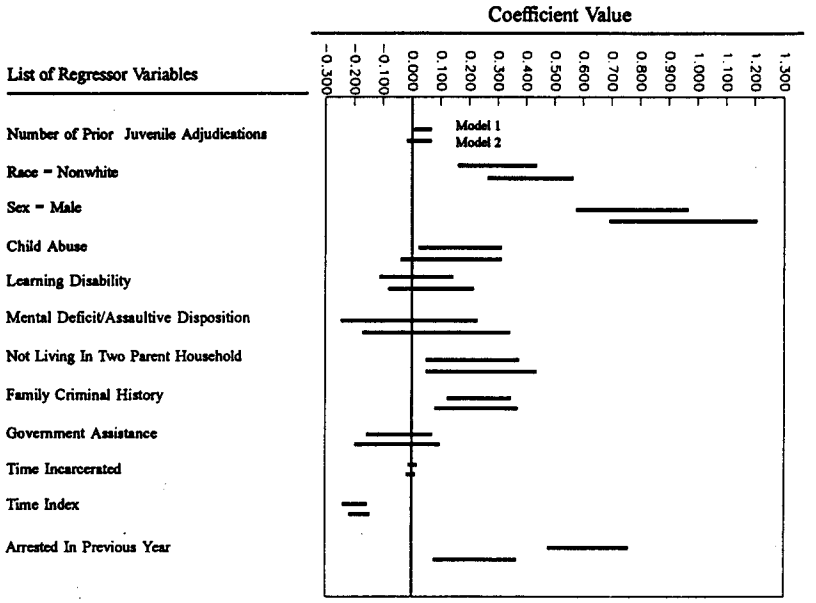

Fig. 2. Ninety-five percent confidence intervals for parameter estimates associated with Models 1 and 2. The 95% intervals are computed under the assumption that repeated random sampling from the target population would yield a sampling distribution for each regression coefficient. This sampling distribution is assumed to be approximately normal. For each coefficient the 95% interval is bounded by the 2.5th percentile and the 97.5th percentile of that sampling distribution.

that δ is sizable and of substantive importance ($\delta = 0.631 - 0.228 = 0.402$). The imposition of statistical controls for persistent unobserved heterogeneity leads to a dramatic deterioration in the estimate of the effect of prior arrests on future arrests [i.e., $(\delta/\eta_{M1}) \times 100 = 63.9\%$ reduction in the value of η].[17] Nevertheless, consistent with Hypothesis 2, the estimate of η was positive and statistically significant under both specifications. The most reasonable conclusion on the basis of this evidence is that, *ceteris paribus*, prior arrests increase the frequency of future arrests. Moreover, these results run counter to the predictions of theories such as Gottfredson and Hirschi's (1990) and

[17]Ideally, we would conduct a formal test of the hypothesis that $\delta = 0$. The test statistic would be distributed as a standard normal variable z. Unfortunately, the existing theory for testing this hypothesis requires either (1) that one know the covariance between the coefficients whose difference is being tested or (2) that the covariance can be assumed to be zero. When the samples on which the coefficients are estimated are independent, the second assumption is reasonable. In this case, however, the models are estimated on the same sample, the coefficients used in the computation of δ are not independent, and we do not know their covariance.

Table III. Expected Relative Frequency Distributions for Arrests During a 1-Year Period Showing Estimated Effect of Prior Behavior With and Without Controls for Persistent Unobserved Heterogeneity

	Model 1. No control for persistent unobserved heterogeneity		Model 2. Persistent unobserved heterogeneity controlled	
Number of arrests	Arrest in prior period = 0	Arrest in prior period = 1	Arrest in prior period = 0	Arrest in prior period = 1
0	0.760	0.643	0.755	0.706
1	0.166	0.199	0.162	0.187
2	0.049	0.084	0.049	0.062
3	0.016	0.038	0.018	0.024
4	0.005	0.018	0.007	0.010
5+	0.004	0.018	0.009	0.011

Note. Expected relative frequency distributions were calculated by constraining all variables to their sample means (except the indicator variable for one or more arrests in the previous period) and evaluating the probability density for a negative binomial model and a random-effects negative binomial model.

Wilson and Herrnstein's (1985), models that anticipate static causal processes. Our findings in this regard corroborate those of Land and Nagin (1996), who also reported evidence of persistent heterogeneity (continuity) and state dependence (change).

Table III provides some insight into the difference in the effect of prior behavior between Model 1 and Model 2. To construct this table, all variables except the indicator variable signifying whether one was arrested in the previous year were constrained to their mean values; we then multiplied the mean variable vector by the estimated regression coefficients with the prior behavior variable set to 0.0 and 1.0, respectively. To secure the probabilities associated with Model 1, we evaluated Eq. (4), and to secure the probabilities associated with Model 2, we evaluated Eq. (5). As might be expected from the results in Table II, the effect of being arrested in the previous period is attenuated as one moves from Model 1 to Model 2.

Although the models presented in Table II provide evidence in favor of dynamic rather than static theoretical perspectives, they do not confront the possibility accommodated by developmental theorists that there are qualitatively different groups whose behaviors are generated by distinct causal processes. In particular, these theories predict that the continued offending behavior of those with extremely high levels of criminal propensity will be caused by a different set of factors than those associated with low-criminal propensity individuals. Among the most prominent of these factors is the effect of prior behavior on future behavior.

To test the state dependence/persistent heterogeneity implications of the Moffitt and Patterson developmental theories, we divided the sample

Table IV. Negative Binomial Regression Models Comparing Those Adjudicated at or Before Age 11 (Model 3) and Those Adjudicated at or After Age 12 (Model 4)

	Model 3 ($N=52$)			Model 4 ($N=786$)		
Variable	Parameter estimate	SE	$\|t\|$ ratio	Parameter estimate	SE	$\|t\|$ ratio
Constant	0.939	0.603	1.56	0.464	0.122	3.79
Time index (γ)	−0.239	0.076	3.12	−0.155	0.019	8.13
Arrested in previous year (η)	0.284	0.264	1.07	0.199	0.076	2.62
a	6.340	4.398	1.44	5.269	0.759	6.93
b	2.393	1.446	1.65	1.865	0.302	6.16
Log-likelihood	−290.76			−3646.12		
Number of parameter estimates	5			5		

into two groups: those who were first adjudicated at an early age and those who were first adjudicated later on. While the literature does not offer specific guidance about the exact cutoff point that should be used to define the low and high propensity groups, the age cutoffs herein are consistent with what has been used by others (Simons *et al.*, 1994, p. 267; LeBlanc and Frechette, 1989, p. 105; Loeber and LeBlanc, 1990, p. 395; Patterson *et al.*, 1992, p. 336; Farrington *et al.*, 1990, p. 325; Stattin and Magnusson, 1995, pp. 418–419, 423). Consequently, models were estimated using four cutoff points: (1) age 11 and younger, (2) age 12 and younger compared to age 13 and older, (3) age 13 and younger compared to age 14 and older, and (4) age 14 and younger compared to age 15 and older.

Using the first age cutoff point, we divided the sample into two groups: those who were first adjudicated at or before age 11 and those who were adjudicated at or after age 12. The results of this estimation are presented in Table IV as Models 3 and 4. Only 52 individuals could be classified as early starters using this scheme and the estimation routine failed to reach convergence using the full set of covariates. We were able to achieve convergence with a model that included the intercept, a time trend indicator, and a prior arrest indicator (as well as the random effects parameters). The log-likelihood for this specification on the entire sample was −3940.12 (with five parameter estimates) while the sum of the log-likelihood values in the separate group equations was −3936.88 (with 10 parameter estimates). Twice the difference between the summed log-likelihood of Models 3 and 4 and the constrained log-likelihood is distributed as a chi-square random variable with $10-5=5$ degrees of freedom. The test statistic provides no support for rejecting the null hypothesis that the population coefficients are the same for each of the groups [$\chi^2_{(5)}=6.48$, $P>0.05$].

In words, the dynamic effect of prior on future behavior is comparable between the low and the high criminal propensity groups. This result is counter to the predictions of the Moffitt and Patterson theories and Hypothesis 3, which anticipate that the effect of prior behavior will be contingent on established levels of criminal propensity. Consistent with the predictions of general theorists, the results point to invariance in the process leading to future offending behavior. As demonstrated below, this conclusion is robust to the choice of cutoff points.

Despite this result, it might be argued that the likelihood difference test is inappropriate when interest centers on a difference between specific coefficients because the chi-square statistic must be very large to reject the null hypothesis when there are many degrees of freedom involved. To address this possibility, a direct test of the hypothesis that $\delta = 0$ (ignoring the possibility of other similarities and differences) was calculated using a large sample test statistic, z, which is distributed as a standard normal random variable,

$$z = |\delta| / [\sigma^2(\eta_{Mi}) + \sigma^2(\eta_{Mj})]^{1/2} \tag{6}$$

where $i \neq j$ and, in the case at hand, $\delta = \eta_{M3} - \eta_{M4} = 0.085$. This result follows the convention of calculating the difference and the variance of the difference between two independent random normal variables (Clogg *et al.*, 1995, pp. 1276–1277). In the case of the hypothesis that $\delta = \eta_{M3} - \eta_{M4} = 0$, the obtained test statistic is not statistically significant at conventional levels ($z = 0.31$, $P > 0.05$). As such, there is no basis in the current data for concluding that the effects of prior arrests vary between those who were first adjudicated at or before age 11 and those first adjudicated later on.

The results of this analysis using the second cutoff point (those age 12 or younger compared to those 13 or older) are presented in Table V as Models 5 and 6. These models were estimable with the full set of covariates. Two tests were conducted to assess the plausibility of the hypothesis that the effect of prior arrests on future arrests is the same between the early and late first adjudication groups (i.e., $\delta = \eta_{M5} - \eta_{M6} = 0$). The first assessment was a global test of the hypothesis that none of the differences between parameter estimates in Models 3 and 4 was greater than zero. To conduct this test, we sum the log-likelihood values of Models 5 and 6. This summed log-likelihood [$\log(L) = -3877.08$] is based on a total of 30 parameter estimates. Twice the difference between the summed log-likelihood of Models 5 and 6 and the constrained log-likelihood of Model 2 is distributed as a chi-square random variable with $30 - 15 = 15$ degrees of freedom. The test statistic provides no support for rejecting the null hypothesis that the population coefficients are the same for each of the groups [$\chi^2_{(15)} = 18.12$, $P > 0.05$].

Table V. Negative Binomial Regression Models Comparing Those Adjudicated at or Before Age 12 (Model 5) and Those Adjudicated at or After Age 13 (Model 6)

	Model 5 ($N=101$)			Model 6 ($N=737$)		
Variable	Parameter estimate	SE	$\|t\|$ ratio	Parameter estimate	SE	$\|t\|$ ratio
Constant	0.509	0.836	0.61	−1.415	0.227	6.22
Number of juvenile adjudications	−0.022	0.066	0.33	0.039	0.026	1.51
Race = nonwhite	0.377	0.229	1.65	0.428	0.078	5.45
Sex = male	−0.033	0.544	0.06	1.057	0.137	7.68
Reported child abuse	0.568	0.256	2.22	0.057	0.102	0.56
Learning disability	0.168	0.203	0.83	0.055	0.079	0.69
Mental deficit/assaultive behavior	0.234	0.348	0.67	0.052	0.143	0.37
No two-parent household	0.317	0.347	0.91	0.232	0.103	2.24
Family criminal history	0.086	0.218	0.40	0.256	0.077	3.28
Government assistance	−0.109	0.250	0.43	−0.010	0.081	0.13
Time incarcerated	−0.006	0.014	0.44	0.000	0.005	0.09
Time index (γ)	−0.229	0.060	3.80	−0.154	0.020	7.66
Arrested in previous year (η)	0.183	0.183	1.00	0.246	0.081	3.03
a	9.388	4.96	1.89	6.819	1.182	5.77
b	2.775	1.53	1.82	3.136	0.675	4.64
Log likelihood	−514.66			−3362.42		
Number of parameter estimates	15			15		

A direct test of the hypothesis that $\delta = \eta_{M5} - \eta_{M6} = 0$ indicates that the difference is not statistically significant at conventional levels ($z = 0.32$, $P > 0.05$). As such, there is no basis in the current data for concluding that the effects of η vary between those who were first adjudicated early and those first adjudicated later on.

We now turn to a third version of the test for differences between groups distinguished by age at first adjudication. In this third version, the cutoff point for the early first adjudication group was raised to include all individuals who were age 13 or younger at the time of their first adjudication. All individuals who were older than 13 at the time of their first adjudication were included in the late first adjudication group. The models for the early and late first adjudication groups, so defined, are presented in Table VI as Models 7 and 8, respectively. The sum of the log-likelihood values for these two group-specific regressions was −3877.69 (based on 30 parameter estimates). A comparison of this summed log-likelihood value to the baseline value of Model 2 yields a nonsignificant test statistic [$\chi^2_{(15)} = 16.90$, $P > 0.05$]. Furthermore, a direct comparison of the estimated values of η between the early and the late adjudication group [using the z test in Eq. (5)] provides no support for the conclusion that $\delta = 0$ in the population ($z = 0.07$, $P > 0.05$).

Table VI. Negative Binomial Regression Models Comparing Those Adjudicated at or Before Age 13 (Model 7) and Those Adjudicated at or After Age 14 (Model 8)

	Model 7 ($N=197$)			Model 8 ($N=641$)		
Variable	Parameter estimate	SE	$\|t\|$ ratio	Parameter estimate	SE	$\|t\|$ ratio
Constant	−0.774	0.455	1.70	−1.217	0.262	4.65
Number of juvenile adjudications	0.029	0.042	0.68	0.016	0.031	0.53
Race = nonwhite	0.295	0.145	2.04	0.465	0.092	5.04
Sex = male	0.388	0.262	1.48	1.088	0.153	7.08
Reported child abuse	0.322	0.152	2.12	0.074	0.115	0.64
Learning disability	0.067	0.135	0.49	0.065	0.089	0.72
Mental deficit/assaultive behavior	0.219	0.242	0.91	−0.000	0.162	0.00
No two-parent household	0.437	0.233	1.88	0.180	0.115	1.56
Family criminal history	0.123	0.132	0.93	0.265	0.089	2.96
Government assistance	−0.162	0.143	1.14	0.015	0.094	0.16
Time incarcerated	−0.002	0.008	0.25	−0.003	0.008	0.38
Time index (γ)	−0.187	0.042	4.48	−0.155	0.022	7.15
Arrested in previous year (η)	0.221	0.136	1.62	0.232	0.088	2.62
a	6.909	2.246	3.08	6.914	1.267	5.46
b	3.918	1.599	2.45	2.684	0.582	4.61
Log-likelihood	−1064.36			−2813.33		
Number of parameter estimates	15			15		

Finally, we consider a fourth test for differences between groups distinguished by age at first adjudication. The early first adjudication group included all individuals who were age 14 or younger at the time of their first adjudication. All individuals who were 15 or older at the time of their first adjudication were included in the late first adjudication group. The models for the two groups are presented in Table VII as Models 9 and 10. The sum of the log-likelihood values for these two equations was −3879.35 (based on 30 parameter estimates). A comparison of this summed log-likelihood value to the baseline value of Model 2 yields a nonsignificant test statistic [$\chi^2_{(15)} = 13.58$, $P > 0.05$]. Direct comparison of the estimated values of η between the early and the late adjudication group [using the z test in Eq. (5)] provides no support for the conclusion that $\delta = 0$ in the population ($z = 0.90$, $P > 0.05$).

Thus far, the results provide support for the idea that both stable individual differences and the effect of prior arrests (a state dependent effect) have implications for the yearly frequency of offending in these data. The results, thus far, do not support the idea that prior behavior exerts different effects between the early and late onset groups. In sum, the effect of prior arrests on future arrests appears to be positive, statistically significant, and

Table VII. Negative Binomial Regression Models Comparing Those Adjudicated at or Before Age 14 (Model 9) and Those Adjudicated at or After Age 15 (Model 10)

	Model 9 ($N=375$)			Model 10 ($N=463$)		
Variable	Parameter estimate	SE	$\|t\|$ ratio	Parameter estimate	SE	$\|t\|$ ratio
Constant	−1.119	0.312	3.58	−1.263	0.311	4.05
Number of juvenile adjudications	0.028	0.031	0.89	0.013	0.044	0.29
Race = nonwhite	0.416	0.108	3.83	0.429	0.108	3.98
Sex = male	0.886	0.178	4.97	1.031	0.194	5.30
Reported child abuse	0.236	0.123	1.92	0.041	0.133	0.31
Learning disability	−0.034	0.106	0.33	0.147	0.107	1.38
Mental deficit/assaultive behavior	0.136	0.183	0.74	0.014	0.194	0.07
No two-parent household	0.416	0.152	2.74	0.101	0.139	0.72
Family criminal history	0.212	0.102	2.07	0.235	0.106	2.22
Government assistance	−0.117	0.109	1.07	0.057	0.114	0.50
Time incarcerated	−0.007	0.006	1.27	0.008	0.010	0.81
Time index (γ)	−0.176	0.027	6.38	−0.154	0.026	5.83
Arrested in previous year (η)	0.159	0.103	1.54	0.294	0.109	2.69
a	6.980	1.618	4.31	6.749	1.455	4.64
b	3.307	0.943	3.51	2.689	0.711	3.78
Log-likelihood	−1876.83			−2002.52		
Number of parameter estimates	15			15		

of similar magnitude regardless of whether one was first adjudicated early or late.

One possible criticism of our analysis as presented thus far is that our theoretical interest centers on the invariance of the state-dependent effect, while our models have been estimated to allow for differences between the groups on all effects. In order to examine the robustness of our results to an alternative specification, we estimated models including an indicator variable for one's age at first adjudication (0 = late first adjudication; 1 = early first adjudication), an indicator variable for whether one was arrested in the previous year (0 = not arrested in the previous year; 1 = arrested in the previous year), a product of the age at first adjudication indicator and the prior arrest indicator, and all of the covariates included in the previous specifications. A test of the hypothesis of invariant state-dependent effects in this specification is given by a z test on the null hypothesis that the product term coefficient is equal to zero in the population.

Table VIII presents the results of this analysis using the age division schema employed in our previous tests. Models 11, 12, and 13 all suggest that the effect of the interaction between whether one was adjudicated early or late and arrest in the previous period is only trivially different from zero

Table VIII. Investigation of Interactions Between First Adjudication Age and Prior Offending Behavior

Variables	Model 11		Model 12		Model 13		Model 14		Model 15	
	Parameter estimate	$\|t\|$ ratio	Parameter estimate	$\|t\|$ ratio	Parameter estimate	$\|t\|$ ratio	Parameter estimate	$\|t\|$ ratio	Parameter estimate	$\|t\|$ ratio
Constant	−1.183	5.57	−1.189	5.58	−1.177	5.50	−1.228	5.83	−1.173	6.09
Number of juvenile adjudications	0.030	1.30	0.030	1.24	0.022	0.89	0.025	1.02		
Race = nonwhite	0.420	5.62	0.422	5.63	0.415	5.56	0.419	5.62	0.400	5.59
Sex = male	0.959	7.39	0.956	7.37	0.944	7.29	0.956	7.39	0.974	7.73
Reported child abuse	0.138	1.45	0.141	1.49	0.141	1.54	0.139	1.54		
Learning disability	0.067	0.88	0.071	0.94	0.067	0.90	0.066	0.89		
Mental deficit/assaultive behavior	0.083	0.63	0.086	0.65	0.089	0.68	0.094	0.72		
No two-parent household	0.246	2.47	0.248	2.48	0.244	2.45	0.247	2.48	0.245	2.52
Family criminal history	0.230	3.13	0.231	3.14	0.226	3.09	0.229	3.14	0.239	3.31
Government assistance	−0.046	0.60	−0.045	0.59	−0.040	0.53	−0.044	0.57		
Time incarcerated	−0.002	0.46	−0.002	0.46	−0.003	0.69	−0.003	0.68		
Time index (γ)	−0.165	8.82	−0.165	8.82	−0.165	8.85	−0.165	8.82	−0.165	8.85
Arrested in previous year (η)	0.228	3.01	0.241	3.08	0.276	3.34	0.339	3.57	0.331	3.49
First adjudicated at age										
≤11	0.116	0.71								
≤12			0.098	0.80						
≤13					0.172	1.71				
≤14							0.146	1.66	0.184	2.33
Arrested in previous year × age at first adjudication group	−0.001	0.01	−0.088	0.55	−0.145	1.15	−0.210	1.80	−0.205	1.76
a	6.714	6.44	6.692	6.49	6.811	6.38	6.794	6.41	6.64	6.44
b	2.881	5.28	2.871	5.31	2.945	5.24	2.944	5.24	2.84	5.31
Log-likelihood	−3885.81		−3885.84		−3884.66		−3884.16		−3887.16	
Number of parameter estimates	17		17		17		17		11	

in the population. The estimates obtained from Model 14, however, lead us to a somewhat more qualified conclusion on this matter.

The estimated effect of prior arrests on future arrests in Model 14 does appear to depend on whether one was first adjudicated at or before age 14. Inspection of the coefficients reveals that the estimated value of η for those who were first adjudicated at or after age 15 was 0.339. The comparable effect among those who were first adjudicated at or before age 14, however, was 0.129. The z value associated with the product of the early first adjudication indicator and the prior arrest indicator is -1.80, which is statistically significant at the 95% confidence level (using a one-tailed test).

To investigate this result more closely, we eliminated the variables that were not statistically significant at the 95% confidence level from Model 14 and estimated the reduced version of the specification as Model 15.[18] The results of this specification led us to exactly the same conclusions as those obtained under Model 14. The effect of prior arrests was significantly weaker within the group of individuals who were first adjudicated at or before age 14 compared to those who were first adjudicated at a later age.

6.1. Summary of Results

Because we have covered much ground in this paper, we briefly summarize our results in this section. It should be recalled that there were two primary questions driving the analysis. First, we examined the relative utility of static and dynamic models of offending frequency. The decrease in the magnitude of prior arrest effects after persistent individual differences were controlled was substantial. This finding is testimony to the important role that continuity and stability play in the process driving offending careers (Gottfredson and Hirschi, 1990; Nagin and Paternoster, 1991; Nagin and Farrington, 1992). Consistent with Hypothesis 2 (and counter to Hypothesis 1—the "strong" heterogeneity hypothesis), the effect of prior arrest on future arrest was positive, statistically significant, and substantively important even after controlling for persistent individual differences. This result provides empirical support for the predictions of dynamic theorists such as Sampson and Laub (1993, 1995) and Loeber and LeBlanc (1990), which anticipate that transitions and change will also play important roles in the process driving offending careers. On balance, these results highlight the theoretical importance of studying *both* continuity and change. A mixed model (the "weak" heterogeneity hypothesis) that combines both heterogeneity and state dependence would seem to be consistent with the data.

[18] A test of whether the log-likelihood of Model 15 differs significantly from the summed log-likelihood values of Models 9 and 10 yields $\chi^2_{(19)} = 15.62$, which is not statistically significant at a 95% confidence level.

Although the analysis provided strong support for the importance of dynamic models that adequately partition stability and change, the results of our assessment of Hypothesis 3 were somewhat more equivocal. The majority of our analyses failed to substantiate the more complex arguments offered by developmental theorists (the category-dependent hypothesis). We do note, however, that one of our specifications did suggest that state-dependent effects are less salient for those who were first adjudicated at or before age 14 compared to those who were first adjudicated at or after age 15. This result strikes us as being consistent with the predictions offered by developmental theorists, but the lack of robustness in this result to slight variations in the early/late onset sample division scheme leaves us with some question about whether the result is artifactual. Indeed, we can think of no good reason for believing a priori that this difference should occur yet not be apparent with any other age division schemes.[19]

In sum, our analysis questions the assumption of developmental theorists that prior behavior will exert different effects between high and low criminal propensity groups. The statistical models estimated herein revealed relatively robust and stable effects of prior arrests on future arrests regardless of onset age.

7. DISCUSSION AND CONCLUSIONS

At the inception of this paper, we suggested that criminological theories can usefully be organized within a two-dimensional framework. Some theories, such as those of Wilson and Herrnstein (1985) and Gottfredson and Hirschi (1990), can best be thought of as static/general theories. Pure static/general theories adopt the position that differences in crime for all offenders are due to time-stable differences in an underlying trait or characteristic that can be termed "criminal propensity." Once formed, criminal propensity is an enduring characteristic that affects the life events of individuals but is not, in turn, affected by those events. Other theories, such as Sampson and Laub's (1993) theory of age-graded informal controls, also apply to all individuals but relax the stability constraint that is a defining feature of Gottfredson and Hirschi's model. Still other theories add to this theoretical complexity by relaxing both the assumption of a general causal process and static causality. These developmental theories anticipate multiple pathways to offending. Within different developmental typologies, the process leading to offending may be either static or dynamic.

[19]We also estimated a model that examined the interaction using a cutoff that classified individuals into the early start group if they were first adjudicated at or before age 15. This analysis revealed no support for the hypothesis that prior arrests exerted different effects on future arrests between the early- and the late-onset groups.

As can be discerned, the static/general theory is far more parsimonious than the dynamic and developmental models. Our purpose in this paper has been to determine if the added complexity of these dynamic and developmental theories is necessary. We have tested key hypotheses bearing on the validity of all of these approaches. One unequivocal conclusion from our analyses is that purely static or purely dynamic models of criminal offending do not appear to fit the facts. The evidence at hand clearly indicates that the relationship between past and future criminal offending cannot be attributed solely to persistent individual differences in criminal propensity nor solely to state-dependent processes of change and adaptation to change. We conclude, therefore, that continuity *and* change matter and that observed change cannot be attributed to a process of self-selection. Our findings, therefore, are more compatible with the theories of Sampson and Laub (1993; Laub and Sampson 1993) and Nagin and Paternoster (1993, 1994), which recognize the theoretical importance of stability but do not trivialize the possibility that people can be profoundly affected by the changes in their lives.

Our reading of the evidence with respect to general vs. developmental models of offending lead us to believe that the complexity inherent in developmental models is probably not necessary, though we remain open on this important question. When separate models of criminal offending were estimated for those with low and high criminal propensity (as measured by age of onset), we found more similarities than differences. Most important, we found no evidence that state dependent effects were more pronounced within the group low in criminal propensity. This finding was robust with respect to the cutoff point for the onset age of offending. It was also relatively robust with respect to method. When product terms of age of onset by prior offending were included in pooled models, the interaction terms were, with one exception, insignificant. The exception was when the age of onset was 14 years old and younger and 15 years and older.[20] From our reading of

[20]As to why an onset age of 14 would be so critical relative to other ages, we can offer only some conjecture at this point. It should be remembered that our data reflect the onset age of *adjudication*. It may well be that a first adjudication by age 14 follows a substantial amount of prior involvement in crime not captured by official statistics. Thus, adjudication by age 14 simply reflects the criminal justice system's response to established antisocial behavioral tendencies. Given our indicator of propensity, and the expectation that the causes of early adolescent delinquency may reflect antisocial tendencies, it is not surprising that the effect of prior arrest on future arrests is lower in this group of delinquents. In other words, adjudication catches up with latent criminal propensity by age 14.

In contrast, the stronger state-dependent effect for official delinquents whose first adjudication occurred at or after age 15 may coincide with time-specific criminogenic influences such as a pronounced influence of peers as one becomes free of parental control, critical period events (entering high school), biological transformations in the transition to puberty, and changes in social experiences (having sexual intercourse). In sum, the volatility of life during adolescence may accentuate the negative effects of an arrest on future behavior, generating a strong state-dependent effect for those who onset later.

APPENDIX

Table AI. Comparison of Observed and Expected Arrest Proportions by Year of Follow-Up

Number of arrests	Year 1		Year 2		Year 3		Year 4		Year 5		Year 6	
	Observed proportion	Expected proportion	Observed proportion	Expected proportion	Observed proportion	Expected proportion	Observed proportion	Expected proportion	Observed proportion	Expected proportion	Observed proportion	Expected proportion
0	0.683	0.671	0.667	0.675	0.712	0.704	0.746	0.744	0.812	0.812	0.782	0.786
1	0.168	0.197	0.196	0.195	0.159	0.183	0.158	0.164	0.126	0.128	0.158	0.143
2	0.075	0.070	0.084	0.069	0.070	0.061	0.062	0.051	0.043	0.034	0.028	0.041
3	0.041	0.029	0.033	0.029	0.039	0.025	0.018	0.019	0.009	0.012	0.016	0.015
4	0.019	0.014	0.011	0.013	0.010	0.011	0.006	0.008	0.008	0.005	0.008	0.006
5	0.011	0.007	0.007	0.007	0.002	0.006	0.007	0.004	0.003	0.002	0.003	0.003
6	0.001	0.004	0.001	0.004	0.006	0.003	0.001	0.002			0.003	0.001
7	0.002	0.002	0.000	0.002	0.001	0.002	0.001	0.001			0.000	0.001
8			0.000	0.001			0.001	0.001			0.000	0.001
9			0.001	0.001							0.000	0.000
10											0.000	0.000
11											0.003	0.000
Total	838		838		838		838		791		386	
θ	0.445		0.428		0.298		0.104		−0.277		−0.121	

Note. θ is the estimated value of an intercept term for each year of the follow-up period, and the beta parameters were estimated as $a = 4.8228$ and $b = 1.4791$. Test of whether observed and expected frequencies differ from each other: $\chi^2_{(441)} = 50.57$, $P > 0.05$.

developmental theory we have no reason for attributing particular significance to those who onset by age 14. In view of this latter finding, however, and the evidence in support of developmental theory from other research discussed earlier in this paper, the jury may still be out on the importance of unique pathways to criminal offending.

In sum, what is clear and noncontroversial is that evidence is mounting that a mixture of both continuity and change are required to account for long-term patterns in criminal offending (Nagin and Farrington, 1992a, b; Nagin and Land, 1993; Sampson and Laub, 1993; Horney *et al.*, 1995; Farrington and West, 1995). We therefore dissent from the thoughtful positions of Gottfredson and Hirschi (1986) with respect to both the idea that change is causally unimportant and its methodological implication that longitudinal research is unnecessary. Indeed, the findings from this research, and other analyses leading to similar conclusions, could not have been derived from cross-sectional studies. We believe that findings like ours have important things to say for the refinement of criminological theory.

REFERENCES

Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology* 30: 47–87.

Akers, R. (1985). *Deviant Behavior*, Wadsworth, Belmont, CA.

Barnett, A., and Lofaso, A. J. (1985). Selective incapacitation and the Philadelphia cohort data. *J. Quant. Criminol.* 1: 3–36.

Barnett, A., Blumstein, A., and Farrington, D. P. (1989). A prospective test of a criminal career model. *Criminology* 27: 373–385.

Barnett, A., Blumstein, A., Cohen, J., and Farrington, D. P. (1992). Not all criminal career models are equally valid. *Criminology* 30: 133–140.

Blumstein, A., Cohen, J., Roth, J., and Visher, C. (1986). *Criminal Careers and "Career Criminals,"* National Academy Press, Washington, DC.

Blumstein, A., Cohen, J., and Farrington, D. P. (1988). Criminal career research: Its value for criminology. *Criminology* 26: 1–35.

Cameron, A. C., and Trivedi, P. K. (1986). Econometric models based on count data: Comparisons and applications of some estimators and tests. *J. Appl. Econometr.* 1: 29–53.

Clogg, C. C., Petkova, E., and Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *Am. J. Sociol.* 100(5): 1261–1293.

Dean, C. W., Brame, R., and Piquero, A. R. (1996). Criminal propensities, discrete groups of offenders, and persistence in crime. *Criminology* 34: 547–574.

Farrington, D. P., and West, D. J. (1995). The effects of marriage, separation, and children on offending by adult males. In Blau, Z. S., and Hagan, J. (eds.), *Current Perspectives on Aging and the Life Cycle. Vol. 4: Delinquency and Disrepute in the Life Course: Contextual and Dynamic Analyses.* JAI Press, Greenwich, CT.

Farrington, D., Loeber, R., Elliott, D., Hawkins, J. D., Kandel, D., Klein, M., McCord, J., Rowe, D., and Tremblay, R. (1990). Advancing knowledge about the onset of delinquency and crime. In Lahey, B., and Kazdin, A. (eds.), *Advances in Clinical Child Psychology*, Plenum Press, New York.

Gottfredson, M., and Hirschi, T. (1986). The true value of lambda would appear to be zero: An essay on career criminals, criminal careers, selective incapacitation, cohort studies, and related topics. *Criminology* 24: 213–234.

Gottfredson, M., and Hirschi, T. (1988). Science, public policy, and the career paradigm. *Criminology* 26: 37–55.

Gottfredson, M., and Hirschi, T. (1990). *A General Theory of Crime*, Stanford University Press, Stanford, CA.

Greene, W. (1995). *LIMDEP Version 7.0 User's Manual*, Econometrics Software, Belport, NY.

Hausman, J., Hall, B. H., and Griliches, Z. (1984). Econometric models for count data with an application to the patents–R&D relationship. *Econometrica* 52(4): 909–938.

Hirschi, T. (1969). *Causes of Delinquency*, University of California Press, Berkeley.

Hirschi, T., and Gottfredson, M. (1995). Control theory and the life-course perspective. *Stud. Crime Crime Prevent. Biann. Rev.* 4: 131–142.

Horney, J., Osgood, D. W., and Marshall, I. (1995). Criminal careers in the short term: Intra-individual variability in crime and its relation to local life circumstances. *Am. Sociol. Rev.* 60: 655–673.

Hsiao, C. (1986). *The Analysis of Panel Data*, Cambridge University Press, New York.

Land, K. C. (1992). Models of career criminals: Some suggestions for moving beyond the current debate. *Criminology* 30: 149–155.

Land, K. C., and Nagin, D. S. (1996). Micro-models of criminal careers: A synthesis of the criminal careers and life course approaches via semiparametric mixed Poisson regression models, with empirical applications. *J. Quant. Criminol.* 12(1): 163–191.

Land, K. C., McCall, P. L., and Nagin, D. S. (1996). A comparison of Poisson, negative binomial, and semiparametric mixed Poisson regression models, with empirical applications to criminal careers data. *Sociol. Methods Res.* 24: 387–442.

Laub, J., and Sampson, R. (1993). Turning points in the life course: Why change matters to the study of crime. *Criminology* 31: 301–326.

LeBlanc, M., and Frechette, M. (1989). *Male Criminal Activity from Childhood Through Youth*, Springer-Verlag, New York.

Lemert, E. (1951). *Social Pathology*, McGraw-Hill, New York.

Lemert, E. (1972). *Human Deviance, Social Problems, and Social Control*, Prentice-Hall, Englewood Cliffs, NJ.

Loeber, R., and LeBlanc, M. (1990). Toward a developmental criminology. In Tonry, M., and Morris, N. (eds.), *Crime and Justice: An Annual Review of Research, Vol. 12*, University of Chicago Press, Chicago, IL.

Moffitt, T. (1993). Adolescent-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychol. Rev.* 100: 674–701.

Moffitt, T. (1994). Natural histories of delinquency. In Weitekamp, E., and Hans-Jurgen, K. (eds.), *Cross-National Longitudinal Research on Human Development and Criminal Behavior*, Kluwer Academic, Dordrecht, The Netherlands.

Moffitt, T., Caspi, A., Dickson, D., Silva, P., and Stanton, W. (1996). Childhood-onset vs. adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Dev. Psychopathol.* 8: 399–424.

Nagin, D. S., and Farrington, D. (1992a). The stability of criminal potential from childhood to adulthood. *Criminology* 30: 235–260.

Nagin, D. S., and Farrington, D. (1992b). The onset and persistence of offending. *Criminology* 30: 501–523.

Nagin, D., and Land, K. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology* 31: 327–362.

Nagin, D., and Paternoster, R. (1991). On the relationship of past and future participation in delinquency. *Criminology* 29: 163–190.

Nagin, D., and Paternoster, R. (1993). Enduring individual differences and rational choice theories of crime. *Law Soc. Rev.* 27: 467–496.

Nagin, D., and Paternoster, R. (1994). Personal capital and social control: The deterrence implications of a theory of individual differences in criminal offending. *Criminology* 32: 581–606.

Nagin, D., Farrington, D., and Moffitt, T. (1995). Life-course trajectories of different types of offenders. *Criminology* 33: 111–139.

Osgood, D. W., and Rowe, D. C. (1994). Bridging criminal careers, theory, and policy through latent variable models of individual offending. *Criminology* 32: 517–554.

Paternoster, R., and Brame, R. (1997). Multiple routes to delinquency? A test of developmental and general theories of crime. *Criminology* 35: 49–84.

Patterson, G. (1993). Orderly change in a stable world: The antisocial trait as a chimera. *J. Consult. Clin. Psychol.* 61: 911–919.

Patterson, G., and Yoerger, K. (1993). Developmental models for delinquent behavior. In Hodgins, S. (ed.), *Mental Disorder and Crime*, Sage, Newbury Park, CA.

Patterson, G., DeBaryshe, B., and Ramsey, E. (1989). A developmental perspective on antisocial behavior. *Am. Psychol.* 44: 329–335.

Patterson, G., Crosby, L., and Vuchinich, S. (1992). Predicting risk for early police arrest. *J. Quant. Criminol.* 8: 335–355.

Sampson, R., and Laub, J. (1993). *Crime in the Making: Pathways and Turning Points Through Life*, Harvard University Press, Cambridge, MA.

Sampson, R., and Laub, J. (1995). Understanding variability in lives through time: Contributions of life-course criminology. *Stud. Crime Crime Prev. Biann. Rev.* 4: 143–158.

SAS Institute (1990). *SAS Language*, SAS Institute, Cary, NC.

Simons, R., Wu, C.-I., Conger, R., and Lorenz, F. (1994). Two routes to delinquency: Differences between early and late starters in the impact of parenting and deviant peers. *Criminology* 32: 247–275.

Stattin, H., and Magnusson, D. (1995). Onset of official delinquency: Its co-occurrence in time with educational, behavioral, and interpersonal problems. *Br. J. Criminol.* 35: 417–449.

Thornberry, T. (1987). Towards an interactional theory of delinquency. *Criminology* 25: 863–891.

Tittle, C. R. (1995). *Control Balance: Toward a General Theory of Deviance*, Westview Press, Boulder, CO.

Weisz, J., Walter, B., Weiss, B., Fernandez, G., and Mikow, V. (1990). Arrests among emotionally disturbed violent and assaultive individuals following minimal vs. lengthy intervention through North Carolina's Willie M Program. *J. Consult. Clin. Psychol.* 58: 72–78.

Wilson, J., and Herrnstein, R. (1985). *Crime and Human Nature*, Simon and Schuster, New York.