SEDUCTIONS OF METHOD: REJOINDER TO NAGIN AND TREMBLAY'S "DEVELOPMENTAL TRAJECTORY GROUPS: FACT OR FICTION?"*

ROBERT J. SAMPSON
Harvard University
JOHN H. LAUB
University of Maryland

Remember LISREL? Many of us can recall a time when papers were accepted (or rejected) in peer-reviewed journals based on the choice of this then-preferred method. Alas, criminology has witnessed a recurring sequence of hegemonic statistical methods. Before LISREL came path analysis and before that factor analysis was all the rage; in the 1980s event history analysis ruled for a brief stint and since the 1990s HLM has come on strong. It is even common to read in abstracts nowadays that a paper applies (fill in the acronym for method) as if that were a contribution in itself.[1] We might include as well the criminal career paradigm that reigned for years largely on grounds of method (for example, how best to measure lambda).

Enter group-based trajectory modeling, aka TRAJ, self-described by Nagin and Tremblay (2005: 874) as the current "statistical method of choice" for developmental research. In a remarkably short time by academic standards, more than fifty papers have been published using group-based modeling, and many more are surely in the pipeline. Is TRAJ just another statistical fad that will pass or has it transformed our knowledge and the way we conduct research?

---

* Original title of Nagin and Tremblay's accepted comment. We thank Travis Hirschi and Stephen Raudenbush for their comments. We also thank the Editor of *Criminology*, Ray Paternoster, for giving us an opportunity to respond.

1. Stephen Raudenbush, one of the founders of HLM, has frequently commented to one of the present authors that nothing irks him more than the notion that applying HLM on a data set should in itself be considered a motivation, much less a substantive contribution.

With seemingly much at stake we eagerly read Nagin and Tremblay's comment that takes aim at our 2003 *Criminology* paper ("Life-course desisters? Trajectories of crime among delinquent boys followed to age 70") in an apparent attempt to resolve whether groups are (in their phrase) fact or fiction. Expecting to learn, we were instead disappointed by the lack of an original contribution. Indeed pretty much everything in their comment has previously been published in numerous articles familiar to criminologists, in recent exchanges considering our work, especially the *Journal of Quantitative Criminology* (Nagin, 2004), and in Nagin's book (2005). As one example, we cannot help but wonder if Figure 2 ("Using groups to approximate an unknown distribution") has worn out its welcome.[2] The experience for us, and we suspect many readers of *Criminology*, will be one of déjà vu, only this time there is no illusion.

Relatedly, we read the paper thinking of what we could say that did not already appear in our original *Criminology* article, subsequent publications such as *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70* (Laub and Sampson, 2003), or our exchange with Nagin published in the *Journal of Quantitative Criminology* (Eggleston, Laub and Sampson, 2004; Sampson, Laub and Eggleston, 2004). What's more, an entire conference and special volume was recently devoted to many (and more) of the issues that Nagin and Tremblay raised. We specifically refer readers to *Developmental Criminology and Its Discontents: Trajectories of Crime from Childhood to Old Age*, a special issue of the *Annals of the American Academy of Political and Social Science* (November 2005). This publication contains our conference paper, a paper by Nagin and Tremblay, and articles by Michael Gottfredson, Stephen Raudenbush, and Wayne Osgood, among others. All bear directly on the issues under discussion here. We would note especially the commentary by Raudenbush on Nagin and Tremblay—"How do we study 'what happens next'?"—and Nagin's and Tremblay's response.

Given this backdrop, we confess to ambivalence about the necessity of further debate and the proliferation of an apparently unprecedented rejoinder to our rejoinder.[3] We shall nonetheless try to be constructive by focusing on clear and important issues that Nagin and Tremblay may wish to address. More important than a sense of closure is our hope that this round of debate will help clarify the role of statistical methods in research and theorizing about crime.

2. By our count Figure 2 has appeared at least six times in print.
3. Nagin and Tremblay (2005: 874) clearly state: "Our purpose here is to comment on several issues raised by Sampson and Laub's thoughtful analysis and critique."

(MIS)CONCEPTIONS

Nagin and Tremblay initially identify three common misconceptions (apparently by us and many others) about group-based developmental trajectories. They are as follows. *Misconception #1*: Individuals actually belong to a trajectory group; *Misconception #2*: The number of trajectory groups in a sample is immutable; and *Misconception #3*: Trajectories of group members follow the group-level trajectory in lock-step. In Nagin and Tremblay's correction they argue, sometimes stated explicitly, that there is no such thing as a group (for example, that it is a mere "word"; an "internal substance of the head"), that the number does not really matter, and that people in the trajectory group do not always follow the pattern. Now imagine a thought experiment where you apply these corrections to a group-based theory such as Moffitt (1993). To our mind, some rather strange implications follow: first, her dual taxonomy theory by definition explains something that does not exist; second, there may be two or ten offender groupings, it does not matter; and, three, Life-Course Persisters desist and Adolescent Limiteds persist—anything is possible! We leave it for the reader to consider the wisdom of these implications for clarifying our knowledge.[4]

We would add one misconception of our own—in the larger sense of the term and outside of criminology: we believe it is reasonable to say some groups do exist. Moreover, nowhere have we said that chronic criminal offenders do not exist. The latter have been around since time immemorial—these individuals are more than words or apparitions. The questions of relevance are whether chronic offenders are categorically distinct in terms of causal mechanisms (as Moffitt and many others explicitly claim), whether there are two distinct groups (as Moffitt also claims), and whether future career criminals can be robustly predicted—prospectively—among juvenile delinquents or those at risk. The entire selective incapacitation paradigm is based on the belief that the answer to the last question is "yes," a belief that remains alive and well in criminology (Sampson and Laub, 2003). We believe the evidence is in, that it is roundly negative, and that nothing in Nagin and Tremblay (2005) proves otherwise. Why criminologists continue to search for a small

4. Consider also Nagin and Tremblay's argument (p. 877) that people respond in categorical terms to questions such as "Did you have a good day?" or "How is it going?" because the query requires an efficient and time-saving heuristic that all parties understand. Days are rarely always good or always bad, we agree. But rather than a "succinct summary" that captures a near-truth, why do we more often say "yes" or "fine" when in fact we feel like hell or are in despair? Erving Goffman (1959) provided a classic answer, of course, one we find far more persuasive. In impression management, the categorical answer (group) is false rather than a heuristic.

number of groups and remain seduced by the idea of ever more-perfect prediction from the distant past is something we leave for future historians of the field.

There are two additional points worthy of comment. One, Nagin and Tremblay stress in Misconception #3 that development is not immutable and that offenders do not follow trajectories in "lock step." We obviously agree but find these claims hard to reconcile with the dominant language and logic of developmental groups. Underlying the notion of distinct developmental types is a teleological assumption—behavior unfolds according to a prior design and toward something. How else are we to make sense of Nagin and Tremblay's rejection of photography because it cannot tell us where the person "*is destined to go*" (2005: 875, emphasis added; unfortunately, the only place we are all destined to go is in fact immutable). The language of destiny reveals the categorical imperative and retrospective strategy of typological models—only when a sequence of actions is completed does it make sense to say that at a prior moment (picture) in the action a later one was predestined, but even here the logic is misleading. As Raudenbush (2005) responded to Nagin and Tremblay in a recent forum on developmental criminology, as we live life forward in time, "Shift Happens."[5]

Second, Nagin and Tremblay contend that group-based statistical modeling "is not bound at the hip" with taxonomic theories. This becomes their *Misconception #4*: Group-based modeling is linked to testing taxonomic theories. This misconception seems to us to backtrack in light of previous statements that the group-based approach was designed for testing theories with a taxonomic dimension, not to mention somewhat odd considering that TRAJ as a method is based on the assumption of groups. For example, Nagin (2004: 33) writes that a central mission of group-based modeling is "testing whether there are distinctive predictors of the clusters (aka groups)." Nagin and Tremblay (1999) take it one step further and assert that one of the principal advantages of the group-based modeling approach is that it is "well suited to analyzing questions about developmental trajectories that are *inherently categorical*—do *certain types of people* tend to have *distinctive* developmental trajectories?" (1182-3, emphasis added; see also Nagin, 1999: 140; Nagin and Tremblay, 2001: 29; Nagin, 2005: 5, 8). We may well be thick headed, but it is hard for us to separate these messages and claims from taxonomic or typological logic.

Another interpretation is that Nagin and Tremblay are now claiming a win-win outcome (double standard?) in their fourth "misconception."

5. We would further note that from a strict logical standpoint, *if* groups are following a pre-ordained or unfolding path, photography at the right moment is cheaper than cinema and will allow the correct conclusion.

Groups are both fact and fiction, with TRAJ best able to finesse the contradiction. They explicitly argued that "because hierarchical and latent growth curve modeling assume a continuous distribution of trajectories within the population... it would be awkward to use these methods to address research questions that contrast categories of developmental trajectories" (1999: 1183). Mix in Nagin and Tremblay (2005) and logically one can only infer the claim that statistical models assuming a continuous distribution are appropriate for testing nontaxonomic theory, whereas Nagin's group-based statistical method is appropriate to test all theories, taxonomic and nontaxonomic alike. Yet as noted (see also Sampson et al., 2004: 41), the Nagin model begins with the assumption that groups exist, leading users to then discover groups even though a model cannot be said to discover what it assumes. And, as Nagin and Tremblay correctly note, TRAJ will estimate ("find") pseudo groups from an underlying continuous distribution.

We thus find it hard to escape the conclusion that in some fundamental sense Nagin and Tremblay want it both ways—the proverbial having their cake and eating it too. This is a natural impulse, to be sure, but our concern is how such a stance can fail to illuminate the important questions at hand.

WHAT ARE THE SUBSTANTIVE LESSONS LEARNED?

Much of the published commentary on group-based statistical modeling and continued in Nagin and Tremblay (2005) appeals to its hypothetical benefits. We are told about groups that might exist or patterns that if they did exist would change our knowledge. Hence another question that logically needs to be asked is: What new substantive knowledge have we learned from analyses using group-based statistical modeling (or HLM, or LISREL, as the case may be)? We believe that the yield on substantive findings about the causes of crime and its developmental course is the larger and more critical issue for any method. Based on our reading of their paper, Nagin and Tremblay seem to present at least five lessons from applications of TRAJ.

The first concerns risk factors. We learn here that the more there are, the higher the rate is of offending, including physical aggression (Nagin and Tremblay, 2005: 883–885, Table 1, and Figure 1, Panel B). This confirms what we learned across a wide range of previous research conducted by the Gluecks, Robins, Rutter, West and Farrington, and Loeber and colleagues, among many others.

Second, we learn that less educated mothers and teen mothers produce delinquents (Nagin and Tremblay, 2005: 886). This too confirms what we learned in research by the Gluecks in *500 Criminal Careers* and *500 Delinquent Women* and others such as Furstenberg, Brooks-Gunn and Morgan's *Adolescent Mothers in Later Life*.

Third, we learn that chronic offenders report disproportionately more violence (Nagin and Tremblay, 2005: 891 and Table 2). This confirms what we learned from Wolfgang, Figlio and Sellin's *Delinquency in a Birth Cohort* and a long line of criminal-career research.

Fourth, we learn that there is variability around the mean (Nagin and Tremblay, 2005: 892 and Figure 5, Panel B). This repeats a basic principle of statistics.[6]

Fifth, we learn that different definitions of desistance (the outcome) lead to different offenders being identified as desisters (Nagin and Tremblay, 2005: 896–897 and Figure 6). This applies a basic principle of research design. Moreover, the proposed "dynamic" definition of desistance—pitted against an easily criticized alternative—that is reproduced in Figure 6 is hardly one of clarity to our eye. We see a mishmash of groups (now up to 7) and are led to question the specific measures and the nature of the sample rather than our knowledge of crime. For example, how many times have these findings (for example, "slow uptake chronics," "late starters") been replicated on other data with a different design? Furthermore, what theory of crime is the definition of desistance testing and what claim is falsified that has import for revising our theories of crime?

We would thus argue that the five substantive lessons offered are venerable and not tied to the use of any particular statistical method, and that most of the real problems Nagin and Tremblay noted are better addressed with a proper design rather than post-hoc application of a statistical method. The argument about counterfactuals in Head Start research offered as a brief for TRAJ only reinforces the latter point. We are told that because only low-income children are eligible for Head Start, any analysis that includes ineligible higher-income children to test its effect is "pure extrapolation, pure speculation" (896). Although true, this is a matter that a proper research design should have solved at the outset. Specifically, the Head Start example suggests the need for an experiment or well-designed quasi-experiment. If that is not possible then the next best solution is propensity score matching or stratification, which

6. Nagin and Tremblay critique our prospective prediction of crime trajectories from child risk factors because there *could* be meaningful variability (i.e., groups) within 'a priori' defined risk groups. Although possible hypothetically, in fact and upon further inspection, the prospective factors provided to us by theory still do not predict distinctive later trajectories—observed *or* expected—as widely claimed. In any case, searching for criminal groups after the fact seems to defeat the whole point of prediction (for further analysis see Sampson and Laub, 2005).

Rosenbaum and Rubin (1983) pioneered to solve exactly the problem Nagin and Tremblay identified.[7]

In short, one does not necessarily need TRAJ or another statistical program to avoid sloppy or inappropriate comparisons. Once again the arguments in favor of the group-based method lean toward preachiness, where the hypothetical benefits are many and risks of sin great.

CONCLUSION

Like LISREL and HLM before it, TRAJ is a tool and we have no disagreement that it has proper, even important, uses. We have used it before (including in collaboration with Nagin) and will again. And, although it may come as a surprise to readers, we actually agree with Nagin and Tremblay on several key points, especially the scientific value of description and pattern recognition. Many of the problems we note in this rejoinder stem from unreflective application and lack of attention to assumptions, a typical scenario when methods diffuse widely. For example, we suspect that by now there are more boilerplate analyses using HLM than TRAJ.

Thus, to our mind, the bigger problem is that criminology seems obsessed with tools rather than keeping its eye on the prize. Indeed we worry less about groups being reified than the methods themselves;[8] when methods rule we lose focus on the fundamental processes that explain crime and its persistence and cessation over the life course. For us, the bottom line of sound research design and basic scientific inquiry is that methods are inextricably linked to, and the servant of, theory. We would hope that Nagin and Tremblay agree but are troubled by their conclusion that "we need to accumulate large quantities of data on large representative samples," an effort said to be greater than coding the human genome or exploring Mars (2005: 898). Whereas they are excited about the many things TRAJ can be turned loose on in such Herculean efforts, from brain images to genetic information to environmental changes, we are not convinced that massive data mining absent a good idea is going to advance the field. So we close on a somewhat different note: if the aim of TRAJ is "to provide the statistical equivalent of cinema

7. The propensity score in this case is the conditional probability of assignment to Head Start given prior observables. Matching or stratification on the propensity score reveals the extent to which a data set can provide meaningful comparisons on a causal variable such as Head Start and drastically reduces extrapolation if there is enough useful prior information. Using TRAJ to identify latent strata is an interesting extension but not required for this task.

8. Lest readers think we overstate our concern, a colleague once asserted to us that, in effect, "to study developmental change, one has to use TRAJ."

to describe and *explain* developmental trajectories" (Nagin and Tremblay, 2005: 876, emphasis added), we are all in favor but await a theory-inspired Oscar performance.

REFERENCES

Eggleston, Elaine P., John H. Laub and Robert J. Sampson

2004 Methodological sensitivities to latent class analysis of long-term criminal trajectories. *Journal of Quantitative Criminology* 20: 1–26.

Goffman, Erving

1959 *The Presentation of Self in Everyday Life.* New York: Anchor.

Laub, John H. and Robert J. Sampson

2003 *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70.* Cambridge, MA: Harvard University Press.

Moffitt, Terrie E.

1993 Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review* 100: 674–701.

Nagin, Daniel S.

1999 Analyzing developmental trajectories: A semi-parametric, group-based approach. *Psychological Methods* 4: 139–157.

2004 Response to 'Methodological sensitivities to latent class analysis of long-term criminal trajectories.' *Journal of Quantitative Criminology* 20: 27–35.

2005 *Group-Based Modeling of Development over the Life Course*. Cambridge, MA: Harvard University Press.

Nagin, Daniel S. and Richard E. Tremblay

1999 Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development* 70: 1181–1196.

2001 Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods* 6: 18–34.

2005 Developmental trajectory groups: Fact or fiction? *Criminology* 43(4): 873–904.

Raudenbush, Stephen W.

2005 How do we study 'What happens next?' *Annals of the American Academy of Political and Social Science,* 602 (November): 131-144.

Rosenbaum, Paul R. and Donald B. Rubin
1983 The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.

Sampson, Robert J. and John H. Laub
2003 Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology* 41: 555–592.

Sampson, Robert J. and John H. Laub, editors
2005 (Special Editors). Developmental criminology and its discontents: Trajectories of crime from childhood to old age. *Annals of the American Academy of Political and Social Science.* Vol. 602, November.

Sampson, Robert J., John H. Laub and Elaine Eggleston
2004 On the robustness and validity of groups. *Journal of Quantitative Criminology* 20: 37–42.

Robert J. Sampson is chairman of the Department of Sociology and the Henry Ford II Professor of the Social Sciences at Harvard University.

John H. Laub is professor of criminology and criminal justice in the Department of Criminology and Criminal Justice at the University of Maryland, College Park. He is also an affiliate faculty member in the Department of Sociology and a faculty associate at the Maryland Population Research Center, both at the University of Maryland. During the 2005–2006 academic year, he is a visiting scholar at the Institute for Quantitative Social Science at Harvard University. His areas of research include crime and deviance over the life course, juvenile delinquency and juvenile justice, and the history of criminology. He has published widely including *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*, co-authored with Rob Sampson, Harvard University Press, 2003. This book has received three major awards: The Albert J. Reiss, Jr, Distinguished Book Award from the American Sociological Association's Crime, Law, and Deviance Section; the Outstanding Book Award from the Academy of Criminal Justice Sciences; and the Michael J. Hindelang Book Award from the American Society of Criminology.