

## Inferring temporal patterns of preservation, origination, and extinction from taxonomic survivorship analysis

Michael Foote

*Abstract.*—Apparent variation in rates of origination and extinction reflects the true temporal pattern of taxonomic rates as well as the distorting effects of incomplete and variable preservation, effects that are themselves exacerbated by true variation in taxonomic rates. Here I present an approach that can undo these distortions and thus permit estimates of true taxonomic rates, while providing estimates of preservation in the process. Standard survivorship probabilities are modified to incorporate variable taxonomic rates and rates of fossil recovery. Time series of these rates are explored by numerical optimization until the set of rates that best explains the observed data is found. If internal occurrences within stratigraphic ranges are available, or if temporal patterns of fossil recovery can otherwise be assumed, these constraints can be exploited, but they are by no means necessary. In its most general form, the approach requires no data other than first and last appearances. When tested against simulated data, the method is able to recover temporal patterns in rates of origination, extinction, and preservation. With empirical data, it yields estimates of preservation rate that agree with those obtained independently by tabulating internal occurrences within stratigraphic ranges. Moreover, when empirical occurrence data are artificially degraded, the method detects the resulting gaps in sampling and corrects taxonomic rates. Preliminary application to data on Paleozoic marine animals suggests that some features of the apparent record, such as the forward smearing of true origination events and the backward smearing of true extinction events, can be detected and corrected. Other features, such as the end-Ordovician extinction, may be fairly accurate at face value.

*Michael Foote.* Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637. E-mail: mfoote@midway.uchicago.edu

Accepted: 8 May 2001

### Apparent versus Real Variation in Taxonomic Rates

Secular variation in origination and extinction rates is a major feature of the history of life, and the documentation and interpretation of this variation continue to motivate paleontological research on many temporal and spatial scales and to inform our understanding of the dynamics of diversification (Newell 1967; Raup and Sepkoski 1982, 1984; Van Valen 1984; Raup 1986; Miller 1997a,b, 1998; McKinney and Drake 1998; Sepkoski 1998; Budd and Johnson 1999; Kirchner and Weil 2000; Foote 2000b, 2001a). An important issue of which paleontologists are quite aware, but which is seldom confronted directly in rate studies (but see Koch 1991; Alroy 1996, 1998; Johnson and McCormick 1999; Alroy et al. 2000; Smith 2001), is that this variation is distorted by variability in the quality of the fossil record and our knowledge of it. Stratigraphic intervals with a more complete record than adjacent intervals will tend to exhibit spuriously high

origination and extinction rates, whereas the converse is true for intervals with more depauperate records. Likewise, apparent extinction rates may be artificially high upon approach to a poorly sampled interval, with apparent origination rates spuriously high afterwards. It seems intuitively reasonable that, if we had prior knowledge of how the quality of the fossil record varied over time, we should be able to correct the apparent origination and extinction record, i.e., to obtain an estimate of the true pattern of origination and extinction underlying the observed data (Wagner 2000). One goal of this paper is therefore to develop a general method to convert observed data on first and last stratigraphic occurrences to estimates of true origination and extinction rates, given some estimate or assumption of the temporal sequence of rates of fossil preservation.

We generally do not know precisely how preservation rates vary over time, however. In cases where data on fossil taxa include information on occurrences within stratigraphic

TABLE 1. Summary of the principal quantities used in this paper.

Quantity	Explanation
$p$	Origination rate per lineage-million-years (Lmy)
$q$	Extinction rate per Lmy
$r$	Preservation rate per Lmy
$P_B$	Probability of preservation before a point in time, for taxa extant at that time
$P_A$	Probability of preservation after a point in time, for taxa extant at that time
$P_{\rightarrow ij}$	Expected probability that last occurrence is in stratigraphic interval $j$ , given that first occurrence is in interval $i$
$P_{\leftarrow ij}$	Expected probability that first occurrence is in interval $i$ , given that last occurrence is in interval $j$
$X_{ij}$	Observed number of taxa with first occurrence in interval $i$ and last occurrence in interval $j$
$F_{\rightarrow ij}$	Observed proportion of taxa with last occurrence in interval $j$ of those with first occurrence in interval $i$
$F_{\leftarrow ij}$	Observed proportion of taxa with first occurrence in interval $i$ of those with last occurrence in interval $j$
$X_{bt}$	Observed number of taxa crossing bottom interval boundary and last appearing within interval
$X_{ft}$	Observed number of taxa first appearing within interval and crossing top interval boundary
$X_{ftl}$	Observed number of taxa with both first and last appearance within interval
$X_{bt}$	Observed number of taxa crossing both bottom and top interval boundaries

ranges, preservation rates can be estimated empirically, but much of the compiled information on stratigraphic ranges consists of only first and last appearances. A second, and in some ways more important, goal of this paper is therefore to estimate the temporal pattern of preservation from data on first and last appearances only. The method developed to achieve this goal will be tested with simulated data, with empirical data that include internal occurrences within stratigraphic ranges, and with empirical data that are artificially degraded.

Although the need to overcome the effects of variable preservation is obvious, true variation in taxonomic rates is distorted by incompleteness even if the quality of the fossil record is uniform. Variation in extinction rate is smeared backward in time to affect prior, apparent rates of extinction, and variation in origination has the same effect forward in time. This is a quite general problem (Foote 2000a), of which the Signor-Lipps effect is perhaps the most familiar example (Signor and Lipps 1982; Meldahl 1990; Rampino and Adler 1998). Equalizing the quality of the record by subsampling a uniform quota of data from each interval (Alroy 1996, 1998; Alroy et al. 2000, 2001) may help remove distortions in apparent taxonomic rates that are caused by

variation in the quality of the record, but we still require a way to remove those distortions that would exist even if the quality of the fossil record were uniform. The approach developed herein helps to meet this need.

Although the purpose here is mainly methodological, some preliminary results will be presented to illustrate the approach, to show cases in which the apparent record of taxonomic rates may be misleading, and to show cases in which the record may be reliable essentially at face value. Throughout this paper I will use *apparent rates* to refer to rates of origination and extinction calculated from fossil data taken at face value, *true rates* to refer to the underlying parameters we would like to estimate, and *estimated rates* to refer to the values of these parameters inferred from the modified survivorship analysis.

### Inferring Origination and Extinction Rates When Preservation Rates Are Known

All taxa known to exist during an interval of time fall into four exclusive categories (Table 1, Appendix 1) (Foote 2000a): (1) taxa with first appearance before the interval and last appearance within the interval; (2) taxa with first appearance within the interval and last appearance afterwards; (3) taxa with first and last appearance within the interval; (4) taxa

with first appearance before the interval and last appearance afterwards. Using  $b$  and  $t$  to refer to the crossing of bottom and top interval boundaries and  $F$  and  $L$  to refer to first and last appearance within the interval, the observed numbers of taxa falling into these categories are denoted  $X_{bL}$ ,  $X_{FL}$ ,  $X_{FL}$ , and  $X_{bt}$ . The relative numbers of these four classes of taxa are determined, in addition to interval length, by five quantities (Pease 1985; Foote 2000a): (1) the rate of origination  $p$  per lineage-million-years (Lmy) during the interval; (2) the rate of extinction  $q$  per Lmy during the interval; (3) the rate of preservation  $r$  per Lmy during the interval; (4) the probability  $P_B$  that a taxon extant at the start of the interval is preserved sometime before the interval, which is a function of all prior origination and preservation rates; (5) the probability  $P_A$  that a taxon extant at the end of the interval is preserved sometime after the interval, which is a function of all subsequent extinction and preservation rates. (In fact, it is always the product of rate and interval length that matters; doubling the interval length while leaving rates unchanged, for example, has the same effect as doubling  $p$ ,  $q$ , and  $r$  while leaving interval length unchanged.) Rates are assumed herein to be constant during an operational interval such as an epoch, age, or zone, but they may vary from one interval to the next. As I have done previously, I will use *preservation* as a shorthand way of indicating the complex processes leading up to the presence of a fossil taxon in a data set, including initial preservation of the taxon, survival and exposure of strata containing the taxon, sampling, identification, publication, and so on.

The estimation of origination and extinction rates, as well as preservation rates when these are not estimated or assumed a priori, depends on modifying standard forward and backward survivorship equations (Kendall 1948; Pease 1985; Raup 1985; Pearson 1992) to accommodate incompleteness and temporal variation in rates. Specifically, given an arbitrary time series of values of  $p$ ,  $q$ ,  $r$ ,  $P_B$ , and  $P_A$ , two sets of relevant probabilities can be determined (Appendix 2): (1)  $P_{\rightarrow ij}$  (“ $P$ -forward”), the probability that a taxon has last appearance in interval  $j$  given that its first appearance

is in interval  $i$ ; (2)  $P_{\leftarrow ij}$  (“ $P$ -backward”), the probability that a taxon has first appearance in interval  $i$  given that its last appearance is in interval  $j$ . *First appearance* and *last appearance* refer to the intervals of first and last preservation, not to the true times of origination and extinction and not to finely resolved levels within the interval. Let  $X_{ij}$  be the observed number of taxa with first appearance in interval  $i$  and last appearance in interval  $j$ , and express these numbers as the following relative frequencies:  $F_{\rightarrow ij}$  (“ $F$ -forward”) is the proportion of taxa having last appearance in  $j$ , of those that have first appearance in  $i$ ; and  $F_{\leftarrow ij}$  (“ $F$ -backward”) is the proportion of taxa having first appearance in  $i$ , of those that have last appearance in  $j$ . Because the expected temporal pattern of first and last appearances can be predicted given the temporal pattern of taxonomic and preservational rates, inferring the “best” values for these rates should be a simple matter of determining the corresponding patterns in  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  that most closely match the observed patterns in  $F_{\rightarrow ij}$  and  $F_{\leftarrow ij}$ . For this purpose I will minimize the sum of absolute log deviations,

$$\sum_{i=1}^n \sum_{j=i}^n |\ln(P_{\rightarrow ij}/F_{\rightarrow ij})| + \sum_{j=1}^n \sum_{i=1}^j |\ln(P_{\leftarrow ij}/F_{\leftarrow ij})|.$$

(Note that the probabilities and proportions are differential rather than cumulative.) This is in essence the same approach that has been used to estimate taxonomic rates from cohort survivorship data (Raup 1978; Jones and Nicol 1986; Foote 1988; Horowitz et al. 1985; Baumiller 1993); the only difference is that incomplete preservation and variability of taxonomic and preservational rates are explicitly included in this survivorship model. Experiments with simulated data, similar to those presented below, suggest that maximizing the support function (Edwards 1992),

$$\sum_{i=1}^n \sum_{j=i}^n X_{ij} \ln(P_{\rightarrow ij}) + \sum_{j=1}^n \sum_{i=1}^j X_{ij} \ln(P_{\leftarrow ij}),$$

may not yield accurate parameter estimates.

The probabilities  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  are similar in some ways to expressions developed by Koch and Morgan (1988). The principal difference is that Koch and Morgan’s approach implicitly

assumes that all taxa in reality range through the entire time span of interest, with patterns of first and last appearance determined solely by preservation and sampling. The survivorship analysis developed herein also differs from the method of confidence intervals on stratigraphic ranges (Shaw 1964; Strauss and Sadler 1989; Marshall 1997; Weiss and Marshall 1999) and from the assessment of likelihoods of true durations given observed ranges (Wagner 2000). Whereas these approaches seek estimates of true times of origination and/or extinction for specific taxa that are preserved, the survivorship analysis seeks estimates of taxonomic rates for the entire statistical pool of taxa, preserved and unpreserved, under the assumption that all taxa are characterized by the same rates of evolution and preservation (an assumption that will be relaxed below). Preserved taxa are biased toward longer true durations, so simply correcting their ranges may underestimate taxonomic rates (although it should still permit temporal patterns to be inferred). Moreover, some confidence-interval approaches require a prior assumption of extinction rate, i.e., average taxon duration (Marshall 1997: pp. 171–172; Weiss and Marshall 1999), but extinction rate is often one of the unknown quantities that we hope to estimate, and one that cannot be assumed to be constant. As shown below, the survivorship analysis also differs from most previous approaches insofar as it enables the estimation of rates of preservation from first and last appearances only, when there is no prior information on preservation.

Sean R. Connolly (personal communication 2001) has kindly reminded me that the problem of estimating origination, extinction, and sampling probabilities has been approached via mark-recapture (M-R) models that bear some similarities to the method developed herein, particularly insofar as they attempt to fit expected values of  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  to observed values of  $F_{\rightarrow ij}$  and  $F_{\leftarrow ij}$  (Nichols and Pollock 1983; Conroy and Nichols 1984; Nichols et al. 1986; Connolly and Miller 2001 [this volume]). Many of the M-R models require data on multiple occurrences, but the band-recovery model is suited for data that include only first and last occurrences (Nichols and Pollock 1983;

Nichols et al. 1986). Although a complete exploration of the relative strengths of M-R methods and the method developed herein is beyond the scope of this paper, a few salient differences should be noted: (1) Taxonomic turnover is assumed in the M-R models to occur only between intervals, whereas origination and extinction within intervals are herein treated as important factors. (2) As a result, single-interval taxa, which prove to be quite informative herein, have been ignored in M-R studies. (3) Separate M-R analyses are applied to estimate different parameters from occurrence data, whereas all parameters are estimated simultaneously herein. (4) The parameters estimated by some M-R models are functions of two or more of the fundamental parameters estimated herein.

If there are  $n$  stratigraphic intervals, the parameter space consists of  $n$  values of  $p$ ,  $n$  values of  $q$ ,  $n$  values of  $r$ ,  $P_B$  for interval 1, and  $P_A$  for interval  $n$ . For this section of the paper, I am assuming that the values of  $r$  are known; this assumption will be relaxed in the next section. When there are no data collected before interval 1,  $P_{B1}$  is constrained to equal zero, and likewise for  $P_{An}$  if there are no data after interval  $n$ . Ideally, we would like to find the values of  $p_1, \dots, p_n, q_1, \dots, q_n, P_{B1},$  and  $P_{An}$  that yield the single best match between expected probabilities and observed frequencies. There are two difficulties, however. First, a space of such high dimensionality is difficult to search exhaustively. Second, there are multiple optima with very similar goodness-of-fit. I have used the method of simulated annealing (Press et al. 1992: p. 444) to explore the space broadly without immediately seeking the nearest optimum, and I have explored the space starting from a large number of randomly chosen initial configurations, using the mean and standard deviation of the resulting solutions to estimate the parameters and their analytical standard errors (see Appendix 3 for details). This standard error captures only the uncertainty attributable to the fitting procedure. In estimating temporal variation in rates from empirical data later in this study, I will also use bootstrap resampling of data to take into consideration the additional uncertainty attributable to noise in the data.

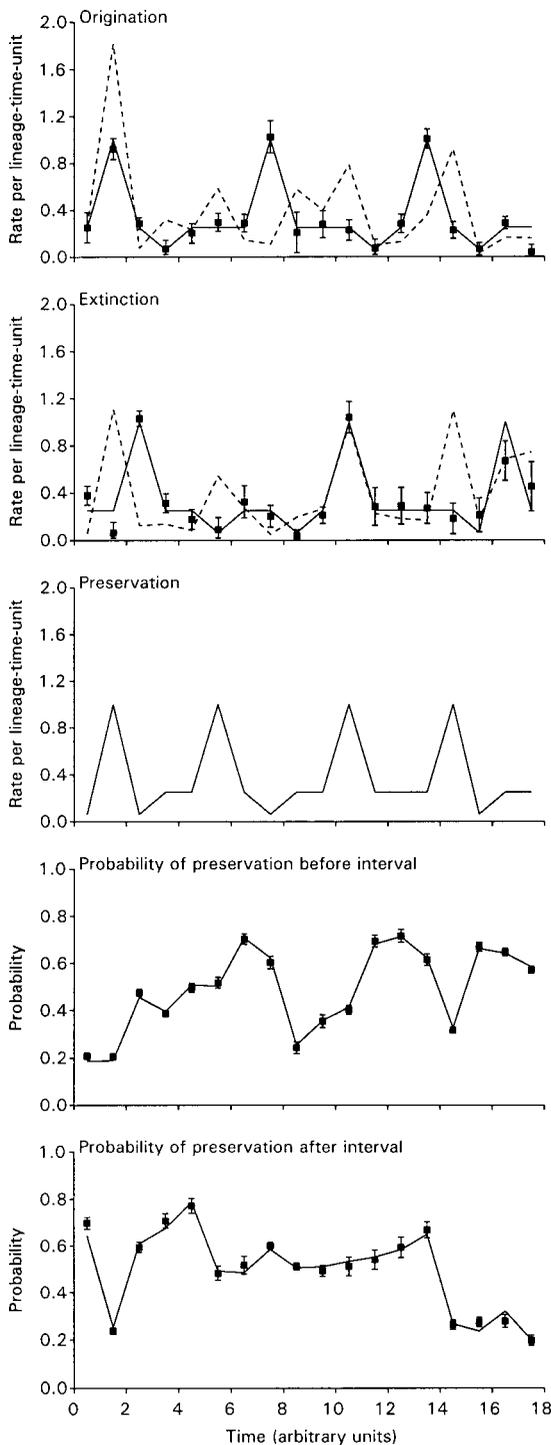


FIGURE 1. Estimation of origination and extinction rates from first and last appearances of simulated range data when preservation rates are known. See text for description of simulation. Initial number of lineages is 5000; total number produced and preserved is 26,647. First and last intervals are omitted, as these cannot be constrained owing to edge effects. Solid line in each panel shows true values of parameters. Broken lines in

Figure 1 shows results of an experiment with simulated data. Rates of origination, extinction, and preservation were varied in a pulsed pattern, in which a background level was punctuated with rates four times as high and one fourth as high, and in which the higher and lower values of  $p$ ,  $q$ , and  $r$  sometimes coincided and sometimes were out of phase. Each time interval was broken into 100 equal time steps, and branching and extinction were simulated using the standard MBL approach (Raup et al. 1973). Incomplete preservation was then simulated by dropping fossil occurrences at random along each lineage's duration, at a rate corresponding to the value of  $r$  for the relevant interval. These occurrences were then used to determine the preserved first and last appearance of each lineage. The broken lines in the top two panels show the apparent rate of origination and extinction when the preserved record is taken at face value, and the solid lines show the true values. Clearly, the apparent values in this simulation reflect variation in preservation rate more than variation in origination and extinction rates. We are fortunate enough to know the preservation rates in this particular case, so we need only to explore the parameter space of  $p_i$  and  $q_i$  and to find the time series of these rates that implies the best match between the expected  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  and the observed values of  $F_{\rightarrow ij}$  and  $F_{\leftarrow ij}$ , given our prior knowledge of preservation rates. Because there are no occurrences outside the time span of the simulation,  $P_{B1}$  and  $P_{A1}$  are known to be zero.

The solution yields the estimates of origination and extinction rates shown as closed squares on Figure 1, which are generally in good, albeit not perfect, agreement with the

←

top two panels show apparent origination and extinction rates, with record taken at face value, calculated as  $p = -\ln[X_{bt}/(X_{Ft} + X_{bt})]/t$  and  $q = -\ln[X_{bt}/(X_{bt} + X_{bl})]/t$  (Foote 1999, 2000a). Solid squares and error bars show mean and standard deviation of 200 estimates, each starting from a different random configuration in the parameter space. Note that apparent taxonomic rates tend to follow preservation rates, whereas estimated rates agree with true rates. Please note that solid line (true values) does not connect the solid squares (estimated values); it merely appears to do so because the estimates are so close to the true values.

true values, as are the probabilities of preservation before and after each interval. (I should emphasize that the solid lines are not simply connecting the rate estimates; rather, the solid lines are the true rates, to which the rate estimates are extremely close.) Because the assumptions of the model are met in this case (namely, that rates are stochastically constant within an interval of time and that all taxa are governed by the same rates), the disagreement between true and estimated values reflects noise in the data and uncertainty in the optimization procedure. There is no clear evidence for an obvious bias, i.e., a systematic over- or underestimation of rates. Note that probabilities of preservation before and after an interval are estimated with greater accuracy and precision than are rates for single intervals. This is because these probabilities depend on many rates and therefore can benefit from compensation of errors in the estimation of individual origination and extinction rates.

The ability to infer true taxonomic rates when rates of preservation are known is certainly encouraging in its own right. There are cases in which we have some empirical estimate of how the quality of the record has varied over time (Raymond and Metz 1995; Alroy 1996, 1998; Miller and Foote 1996; Markwick 1998; Johnson and McCormick 1999; Alroy et al. 2000; Peters and Foote 2001; Smith 2001), or in which we can use principles of stratigraphy to model this variation (Holland 1995, 2000; Brett 1998; Holland and Patzkowsky 1999). Nevertheless, many comprehensive compilations of data, those that have permitted the analysis of large-scale origination and extinction patterns, consist only of first and last occurrences without detailed information on internal occurrences that would permit an obvious estimate of the quality of preservation (e.g., Sepkoski 1992, 1996). The task to which I will turn for the remainder of this paper, therefore, is to infer taxonomic rates from first and last appearance data when we have no prior knowledge of preservation rates.

### Inferring Origination and Extinction Rates When Preservation Rates Are Not Known

Considering the survivorship probabilities  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  (Appendix 2), preservation rates

do not play a special role that would require us to start with these rates and to estimate origination and extinction rates from this starting point. Following the optimization procedure illustrated in Figure 1, we could just as easily estimate rates of preservation if we had prior knowledge of true origination and/or extinction rates. In fact, we could even estimate all rates—origination, extinction, and preservation—given no prior knowledge of any of them. This is illustrated in Figure 2, which depicts an analysis of the same simulated data used for Figure 1. The only difference here is that no knowledge of preservation rates was used in the fitting procedure. Rather, origination, extinction, and preservation rates were all treated as unknown parameters that were fitted to the data simultaneously (Appendix 3). (Of course, the number of parameters to be estimated must be less than the number of observations, i.e., the number of non-zero elements in the observed  $F_{\rightarrow ij}$  matrix, which at most is equal to  $(n^2 + n)/2$ . With  $3n + 2$  parameters to be estimated, this means that there must be at least six intervals; more intervals could be required if the  $F_{\rightarrow ij}$  matrix is sparse, and, in any event, it is desirable when possible to use more intervals to obtain better constraints on estimated rates.)

A few salient points emerge from this exercise. First, and rather trivially, the uncertainty in estimated origination and extinction rates is greater when there is no prior knowledge of preservation rates. Second and most important, *it is possible to estimate preservation rates from data that consist only of first and last appearances*. Third, origination rates are estimated more poorly near the beginning of the time series and extinction rates more poorly near the end. As discussed below and in Appendix 1, this is related to the fact that the observed numbers of taxa falling into the four fundamental categories in effect constrain  $P_B$  and  $P_{A'}$ , and therefore the sum of  $p_i$  and  $r_i$  before each interval and of  $q_i$  and  $r_i$  after each interval (Foote 2000a: eqs. 24a, 25a). With fewer values of  $p_i$  contributing to each value of  $P_B$  near the beginning, there are fewer constraints; and similarly for later values of  $q_i$  and  $P_{A'}$ .

The estimation of preservation rates and

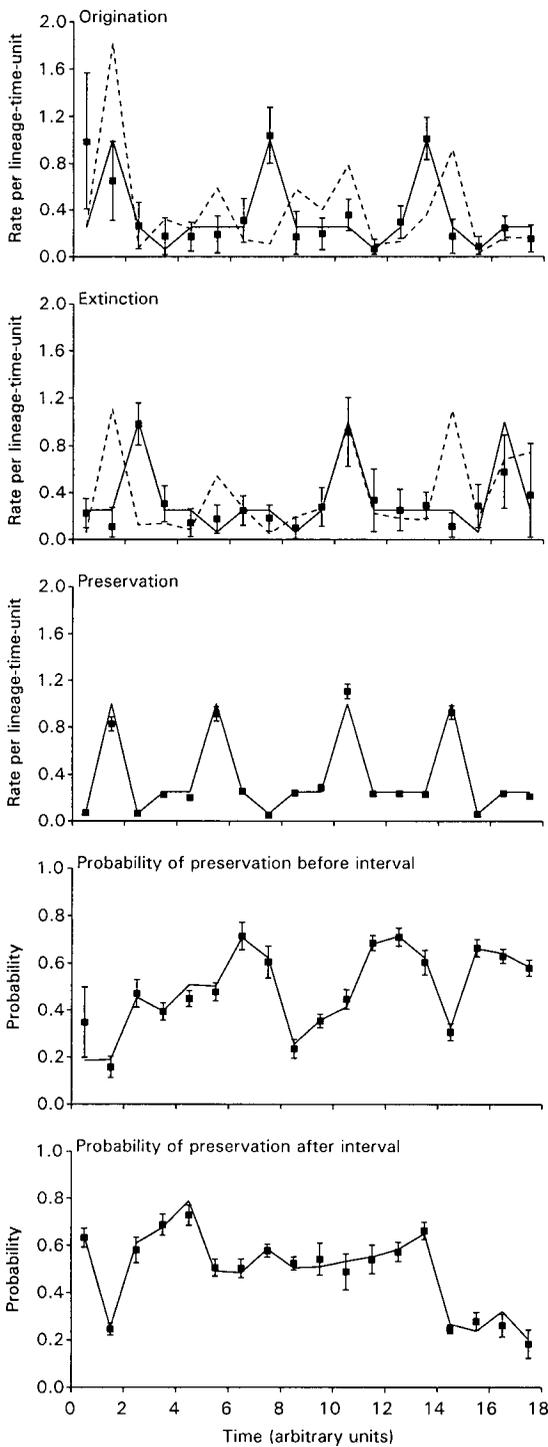


FIGURE 2. Estimation of origination, extinction, and preservation rates from first and last appearances of the same simulated data as in Figure 1. In contrast to Figure 1, no knowledge of preservation rates is assumed here. Although there is greater uncertainty in the estimation of taxonomic rates when preservation rates are not known, all three rates, as well as the preservation probabilities that they determine, are fairly well estimated from first and last appearance data alone.

true times of origination and extinction has generally relied on having internal occurrences within stratigraphic ranges (Paul 1982, 1998; Strauss and Sadler 1989; Solow and Smith 1997). There are exceptions, but these generally rely, as do some estimates based on internal occurrences, on the assumption of constant rates of origination, extinction, and preservation (Foote and Raup 1996; Foote 1997; Wagner 1997; Cheetham and Jackson 1998; Foote and Sepkoski 1999). The ability to infer temporally varying preservation rates from first and last appearances only may thus seem like good luck at best, black magic at worst. It is therefore worth considering why, intuitively, we should expect first/last appearance data to carry substantial information on temporal variation in the quality of preservation. Consider the four fundamental classes of taxa within an interval,  $X_{bL}$ ,  $X_{Ft}$ ,  $X_{FL}$ , and  $X_{bt}$ , and assume for the moment that all taxa with last appearances in the interval in fact existed before the interval, i.e., that all single-interval taxa are artifacts of incomplete preservation before and after the interval. Then it is clear that the proportion of last appearances that belong to taxa known before the interval, namely  $X_{bL}/(X_{bL} + X_{FL})$ , must be equal to  $P_B$ , the probability of preservation before the interval. But if some or all of the taxa with both first and last appearance within the interval were confined to the interval in reality, then  $P_B$  must be higher than implied by the initial assumption that all single-interval taxa are artifacts. Therefore,  $X_{bL}/(X_{bL} + X_{FL})$  must set an absolute lower bound on  $P_B$ , a bound that constrains the sum of origination and preservation rates preceding the interval (Foote 2000a: eq. 24a). This makes sense intuitively; it is also easily shown algebraically (Appendix 1). Likewise, the proportion of first appearances that belong to taxa known after the interval,  $X_{Ft}/(X_{Ft} + X_{FL})$ , gives a lower bound on  $P_A$  and thus on the sum of extinction and preservation rates following the interval (Foote 2000a: eq. 25a). Because a time series of these proportions implies a time series of overlapping sums, some constraint on the individual elements of the sums is obtained. I know of no general upper bound on  $P_B$  and  $P_A$  that is implied by the X-counts for a single interval; in

the absence of additional information, the upper bounds on these probabilities must be taken as unity.

The constraints just discussed concern taxonomic rates and preservation outside a stratigraphic interval. What constraints do the observed data for an interval place on the rates of interest within the interval itself? Because there are only four observed quantities ( $X_{bL}$ ,  $X_{Fl}$ ,  $X_{FL}$ , and  $X_{bl}$ ) but five unknown quantities on which these depend ( $p$ ,  $q$ ,  $r$ ,  $P_B$ , and  $P_A$ ), there is no unique solution for a single interval. The solution space is nevertheless finite (see Appendix 1 for an example). If  $P_B$  and  $P_A$  are both constrained to equal their lower bounds, then the preservation rate within the interval is maximal and the origination and extinction rates must both be zero. As  $P_B$  and  $P_A$  increase, the values of  $p$  and  $q$  required to explain the observations increase, while the required value of  $r$  decreases. To see why varying  $P_B$  and  $P_B$  from their minimal to their maximal values should have this effect, consider the role of single-interval taxa. If the probability of preservation outside the interval were minimal, then these singletons would most likely extend beyond the interval in reality, but simply wouldn't be preserved outside it. This would imply a lower level of origination and extinction within the interval than would be suggested by the number of first and last appearances, and it would also imply a higher rate of preservation within the interval, because these taxa all had to be preserved therein. Conversely, if the probability of preservation were maximal outside the interval, then the singletons would have been preserved outside it if they had in fact extended beyond it. We can therefore conclude that the number of singletons is due to high rates of both origination and extinction within the interval.

The foregoing discussion addresses why we should expect data on first and last appearances to bear any information on preservation. The stated constraints in practice are based on observed values that have error, especially when cast in the form of ratios, so the optimization procedure does not attempt to constrain the solution to satisfy all constraints in an absolute sense. Rather, it seeks the solution

that involves the least deviation between expected and observed patterns of first and last occurrence, as explained above.

For each interval, we have constraints on the sum of prior origination and preservation rates, on the sum of subsequent extinction and preservation rates, and on the range of values of origination, extinction, and preservation rate within the interval. Given readily available stratigraphic range data, we can estimate temporal patterns in taxonomic rates and rates of preservation. Any additional constraints that can be imposed, such as prior knowledge of some or all preservation rates or even their relative ordering, will increase statistical power and reduce uncertainty in the parameters that are estimated, but these additional constraints are by no means necessary. All that we really need are first and last appearances, provided that there are enough of them to limit noise and enough stratigraphic intervals so that there are more observed counts of  $F_{\rightarrow ij}$  and  $F_{\leftarrow ij}$  than there are parameters to be estimated.

#### Additional Tests of the Approach

*Simulated Data.*—There are indefinitely many scenarios of temporal variation in origination, extinction, and preservation that could be explored. Perhaps the principal question is whether spurious variation in origination and extinction rates, caused in turn by variation in the quality of the record, can be detected and corrected. The most direct way to explore this is to simulate data in which the true rates of origination and extinction are constant, while preservation rate varies. This is illustrated in Figure 3, which shows the results of a simulation much like that of Figures 1 and 2, except that taxonomic rates are constant and only rates of preservation are pulsed. In addition to showing edge effects, with origination artificially high near the start and extinction artificially high near the end, the apparent taxonomic rates go up and down with preservation rate. Inverting the simulated data, while ignoring any knowledge of preservation, yields estimated rates that tend to be equal to their true values, within uncertainty, and generally closer to their true values than to their apparent values. Even though the

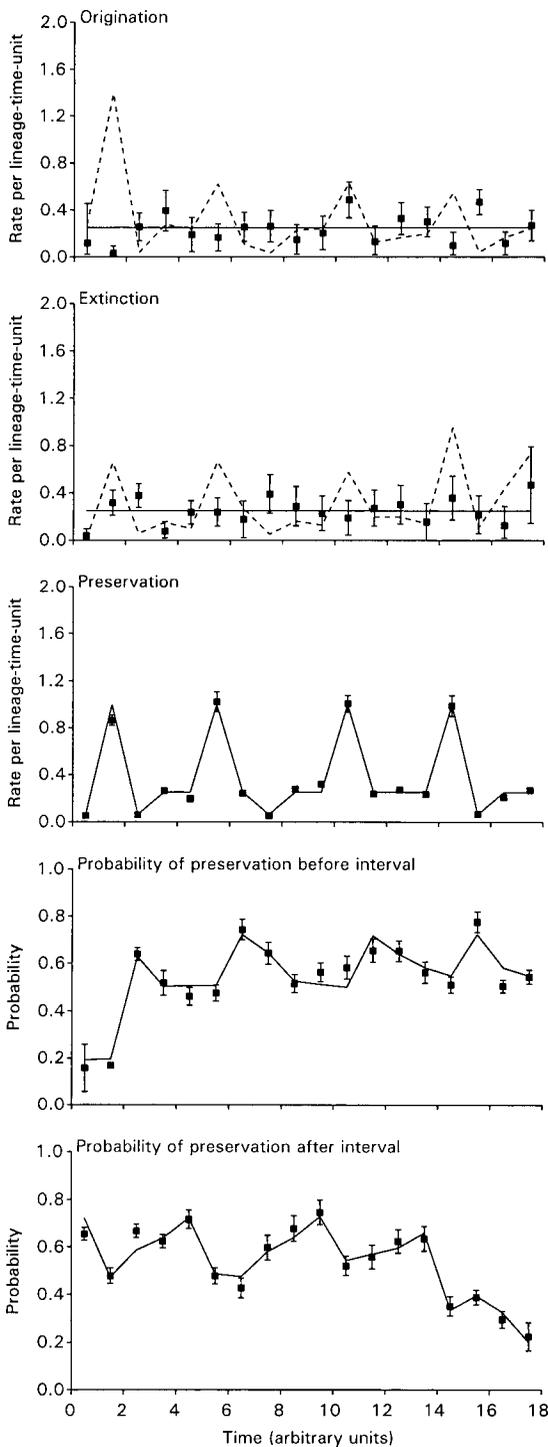


FIGURE 3. Estimation of origination, extinction, and preservation rates from first and last appearances of simulated data generated with constant taxonomic rates and variable preservation rate. See Figure 1 for explanation of lines and symbols. Starting diversity is 5000; total number of lineages produced and preserved is 16,221. Apparent rates of origination and extinction

fossil record is conspiring to trick us into thinking that origination and extinction rates are fluctuating substantially, the details of the distribution of first and last appearances allow us to infer that it is preservation that varies and that the true pattern of origination and extinction is much closer to constant than would be suggested by the apparent rates.

A different problem is posed in the case of a constant rate of preservation. Does the attempt to fit rates to data trick us into thinking that the pattern of first and last appearances is attributable in part to variation in preservation when it is in fact due only to variation in origination and extinction? Figure 4 illustrates a simulation that explores this question. Here, origination and extinction rates are pulsed as in Figures 1 and 2, whereas preservation rate is constant. The results are encouraging, even if less so than those of Figure 3. Although the estimate of preservation sometimes deviates significantly from its true value, this estimate varies within rather narrow bounds. Moreover, estimated origination and extinction rates generally agree with their true values, within error, and much of the temporal pattern is captured. Note that the two worst estimates of preservation rate occur about two-thirds of the way into the time series, where a peak in extinction is followed shortly by a peak in origination (both in the true rates and in the apparent rates). This is not surprising, because one possible way of explaining such a pattern in apparent peaks would be that they are both artifacts of the same intervening intervals of poor preservation. This does not happen if the peak in origination precedes the peak in extinction, as it does about halfway into the time series. A feature of the estimates in Figure 4 that is found in some other simulations (not presented) is the tendency for the estimated values of taxonomic rates to smooth out the true variation, especially near the ends of the time series where there are fewer constraints. Nevertheless, large upward or downward excursions

←

vary spuriously with preservation rate, but this spurious variation is largely eliminated and true rates are reasonably well estimated.

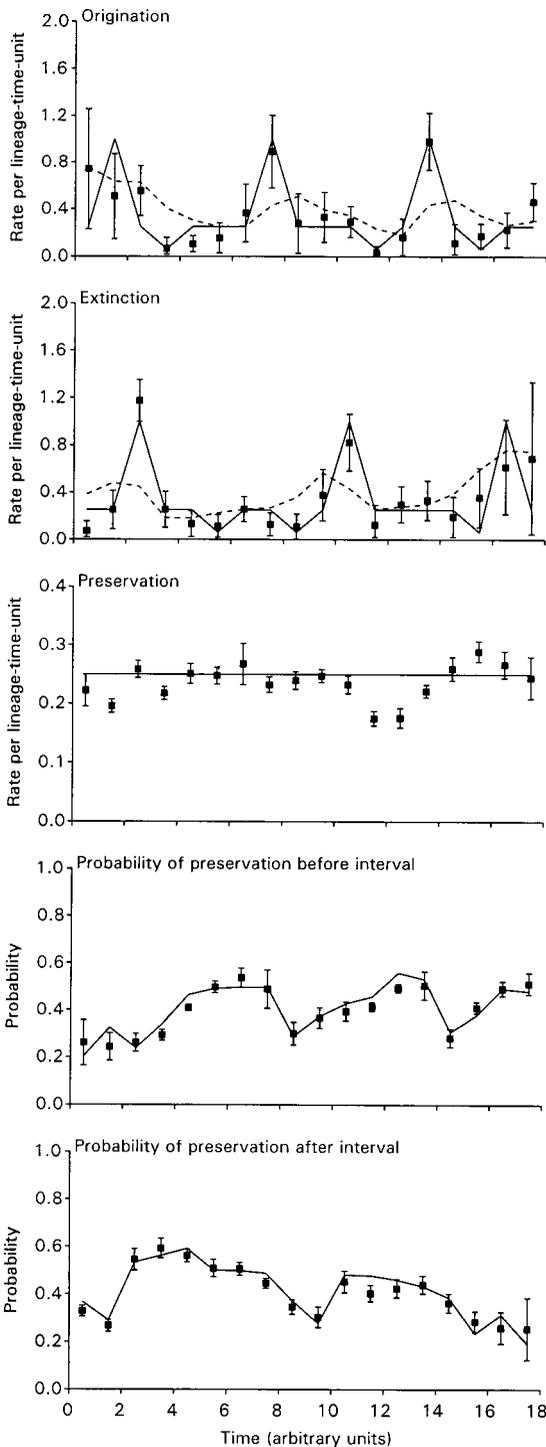


FIGURE 4. Estimation of origination, extinction, and preservation rates from first and last appearances of simulated data generated with constant preservation rate and variable taxonomic rates. See Figure 1 for explanation of lines and symbols. Starting diversity is 5000; total number of lineages produced and preserved is 20,757. Note that, when taxonomic rates truly vary, the

tend to be captured well. Whether there is sufficient reason to invoke a model with varying preservation rates, rather than a simpler model with constant rates, is, in part, a statistical question that can be addressed with empirical data. This will be touched on below.

*Comparing Independent Empirical Estimates of Preservation.*—To test the method empirically, I analyzed data on Ordovician marine animal genera kindly supplied by Arnold I. Miller (personal communication 2000). These data contain internal occurrences in addition to first and last appearances. The occurrences represent trilobites, brachiopods, bivalves, gastropods, and monoplacophorans (for details, see Miller and Mao 1995, 1998; Miller 1997a,b). Except for the Llanvirn and Llandeilo, which were left undivided, I analyzed data at the sub-series level of resolution. All told, 8528 genus occurrences are adequately resolved stratigraphically, and 983 genera have both first and last appearances properly resolved. In addition to estimating preservation rates based on the first and last appearances alone, I measured preservation directly from the internal occurrences in two ways: (1) The proportion of through-ranging taxa that are actually sampled in an interval of length  $t$  is expected to be equal to  $1 - e^{-rt}$ , so I estimated  $r$  for each interval as  $-\ln(1 - X_{bt,samp}/X_{bt})/t$ , where  $X_{bt,samp}$  is the number of genera that cross both interval boundaries and are sampled within the interval (see Foote 2000a for a similar analysis on an earlier version of these data). (2) Rather than scoring each through-ranging genus as either present or absent, I tabulated the number of occurrences of through-ranging genera, dividing this in turn by the number of such genera and by interval length. For a given interval, these two measures of preservation are based only on occurrences of genera that extend both before and after the interval. The rate estimates based on fitting  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  to  $F_{\rightarrow ij}$  and  $F_{\leftarrow ij}$  in contrast,

←

apparent taxonomic rates are distorted even if preservation is constant. Although some of the true variation in taxonomic rates is blurred, the estimated rates are generally close to the true values. See text for further discussion.

use only those occurrences that define the beginning or end of the stratigraphic range of a genus. The two direct estimates are therefore logically independent of the estimate derived from the fitting procedure.

Preservation rates estimated from the survivorship approach are reasonably well correlated with those estimated directly from the internal occurrences (Fig. 5). The same is true if rates are expressed per stratigraphic interval rather than per Myr, i.e., if intervals are treated as having unit length (results not presented). Rate estimates for the intervals in the upper right of Figure 5 are the most uncertain. As we would expect given the difficulty of constraining rates near the edges, these are the first and last intervals for which estimates are possible, the upper Tremadoc and the middle Ashgill. (Because the data contain no occurrences outside the Ordovician, rates for the lower Tremadoc and upper Ashgill cannot be constrained.) The points in Figure 5A fall above the line of 1:1 correlation. The higher absolute values of  $r$  as estimated from the internal occurrence data may well reflect taxonomically heterogeneous preservation potential. If there are different preservational classes with different proportions of taxa preserved, the bulk value of  $X_{bt,samp}/X_{bt}$  will be biased toward the better-preserved groups, because, all else equal, these will tend to have longer stratigraphic ranges and therefore will be disproportionately represented in  $X_{bt}$ . It is also possible that  $r$  is underestimated by the optimization approach because of taxonomic heterogeneity in rates of origination and extinction (see below). The important point here, however, is that the temporal pattern in preservation rate is similar whether we estimate it from data on internal occurrences or from first and last appearances only.

A similar analysis was performed on marine data that are being assembled as part of the Paleobiology Database (PD) (Alroy et al. 2001) (Fig. 6). These data consist of a broad sample of taxonomic lists with stratigraphic, lithologic, geographic, and other information. The subset of data analyzed here, current as of 30 August 2000, consists of genus-level occurrences of anthozoans, brachiopods, bivalves, cephalopods, gastropods, echinoderms, and

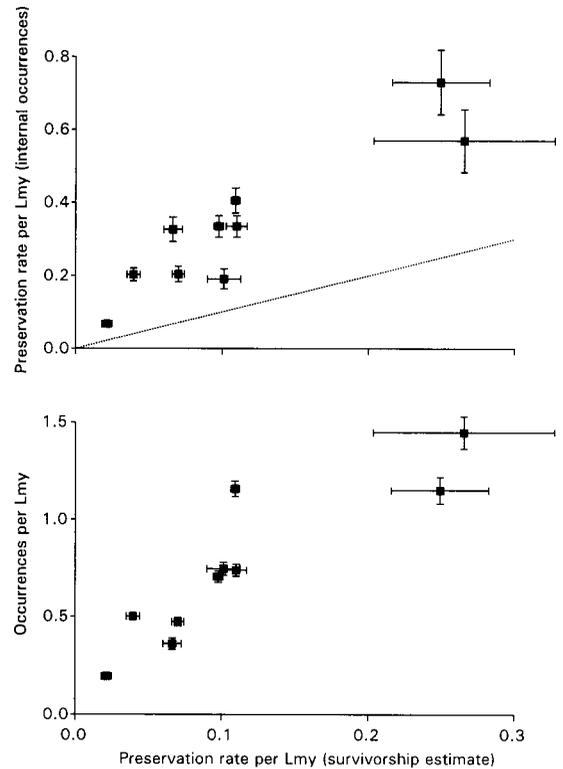


FIGURE 5. Comparison between Ordovician preservation rates estimated by survivorship analysis of first and last occurrences (abscissa) and quality of preservation measured from internal occurrences (ordinate). Data are from Arnold I. Miller (personal communication 2000; see text). The timescale is that of the Geological Society of America (Palmer and Geissman 1999), with modifications from Bowring and Erwin (1998). Each point represents a single stratigraphic interval. Fitted rates and their error bars are the mean and standard deviation of 200 optimizations starting from different random configurations. Error bars on preservation rate  $r$  per Lmy (top panel) are approximated by calculating standard error of the proportion  $R$  of through-ranging genera preserved ( $R = X_{bt,samp}/X_{bt}$ ) as  $SE(R) = \sqrt{R(1 - R)/X_{bt}}$ , adding this to  $R$ , and recalculating  $r$ . Error bars on occurrences per Lmy (bottom panel) are approximated by calculating standard error of number of occurrences,  $N_{occ}$ , as  $\sqrt{N_{occ}}$ . Dotted line in top panel shows 1:1 correlation. Conventional measures of preservation based on internal occurrences are positively correlated with rate estimates based on survivorship analysis of range endpoints only, although survivorship estimates are systematically lower (top panel). See text for further discussion.

trilobites from Europe, North America, North Africa, and the Middle East. I have used the PD's stratigraphic subdivision, which lumps and divides conventional stratigraphic stages and series into bins of approximately 10-Myr duration. I estimated preservation rates for

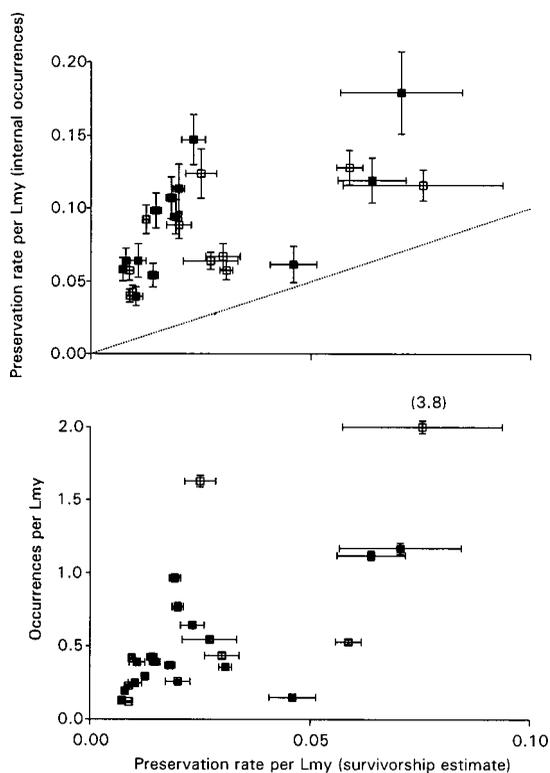


FIGURE 6. Comparison between Phanerozoic preservation rates estimated from first and last occurrences (abscissa) and quality of preservation measured from internal occurrences (ordinate). Data and timescale are from the Paleobiology Database (see text). Closed squares = Paleozoic intervals; open squares = post-Paleozoic intervals. See Figure 5 for explanation. Open square at top of bottom panel is plotted below its true value of 3.8 occurrences per Lmy.

mid-Ordovician through mid-Carboniferous intervals and for mid-Jurassic through Oligocene intervals. The Paleozoic part of the data includes 20,516 occurrences of 2106 genera; the post-Paleozoic part includes 37,722 occurrences of 2011 genera. There is more uncertainty in the estimates of preservation derived from survivorship analysis for these data than for the Ordovician data (Fig. 6). This greater uncertainty is likely to reflect several factors. First, the data are sparser; for example, the Paleozoic part consists of roughly twice as many genera and occurrences as Miller's data, despite covering a span of time nearly four times as long. (Although there are substantially more data in the PD than suggested by the foregoing tabulation of occurrences, many occurrences are not yet properly resolved into

the operational stratigraphic units.) Second, the individual intervals are longer, which can make it more difficult to constrain rates (see Appendix 1 and "Discussion and Conclusions"). Third, the PD data include a wider range of classes; they therefore deviate more from the assumption that all taxa are governed by the same rates. Nevertheless, the preservation-rate estimates from survivorship analysis of the PD data correlate reasonably well with direct estimates based on internal occurrences. The poorer match in the lower panel of Figure 6 partly reflects heterogeneity among intervals in the nature of faunal lists. Intervals with many lists that contain the same taxa repeatedly will have substantially more occurrences without adding to the "effective rate" of preservation.

The concordance between different preservation-rate estimates (Figs. 5, 6) permits some confidence that, even when we leave the idealized world of simulated data, it is possible to estimate how the quality of the fossil record varies over time even if the only data we have are first and last appearances.

*Artificially Depauperate Sampling.*—An additional test of the method was performed on the Ordovician data by artificially degrading the quality of the record. For each interval other than the lower Tremadoc and upper Ashgill, an arbitrary seven out of eight occurrences were randomly discarded. Occurrence data were then converted to first and last appearances, apparent origination and extinction rates were calculated, and true origination, extinction, and preservation rates were fitted to the data using only first and last appearances, without exploiting prior knowledge of sampling. As expected, the artificial holes in sampling lead to apparent origination and extinction rates that are lower than the apparent values based on all the data (Table 2). The survivorship analysis effectively detects and corrects these spuriously low rates for the most part, leading to estimated preservation rates lower than those obtained for all the data, and to estimated true origination and extinction rates that are essentially the same as those obtained when all the data are analyzed (Table 2). Incidentally, this exercise provides an example of a nonlinear relationship be-

TABLE 2. Analysis of Ordovician data with artificially depauperate sampling. For each interval indicated, one out of every eight occurrences was randomly retained and the rest were omitted. No such knowledge was used in subsequent rate estimation, which depends only on first and last appearances. Rates (per interval) given as mean  $\pm$  standard deviation, based on bootstrap resampling of genus ranges for apparent rates and bootstrap resampling and survivorship analysis for estimated rates.

Interval	Data	$X_{BL}$	$X_{FI}$	$X_{FL}$	$X_{BT}$	$r_{estimated}$	$P_{apparent}$	$P_{estimated}$	$q_{apparent}$	$q_{estimated}$
Upper Tremadoc	raw	30	50	38	71	0.59 $\pm$ 0.38	0.54 $\pm$ 0.73	1.0 $\pm$ 0.66	0.35 $\pm$ 0.058	0.37 $\pm$ 0.26
Lower Arenig	degraded	5	10	5	71	0.20 $\pm$ 0.24	0.13 $\pm$ 0.043	0.93 $\pm$ 0.61	0.074 $\pm$ 0.033	0.28 $\pm$ 0.22
Upper Arenig	raw	40	106	91	81	0.82 $\pm$ 0.36	0.83 $\pm$ 0.084	0.42 $\pm$ 0.48	0.40 $\pm$ 0.063	0.24 $\pm$ 0.28
Lower Llanvirn	degraded	13	42	21	81	0.40 $\pm$ 0.20	0.42 $\pm$ 0.068	0.53 $\pm$ 0.42	0.15 $\pm$ 0.041	0.13 $\pm$ 0.17
Upper Llandeilo	raw	19	15	13	168	0.18 $\pm$ 0.12	0.88 $\pm$ 0.21	0.21 $\pm$ 0.28	0.11 $\pm$ 0.026	0.22 $\pm$ 0.22
Lower Caradoc	degraded	1	4	4	168	0.054 $\pm$ 0.041	0.023 $\pm$ 0.013	0.21 $\pm$ 0.28	0.007 $\pm$ 0.007	0.18 $\pm$ 0.19
Upper Caradoc	raw	46	87	64	137	0.44 $\pm$ 0.14	0.49 $\pm$ 0.056	0.59 $\pm$ 0.38	0.29 $\pm$ 0.042	0.22 $\pm$ 0.24
Lower Ashgill	degraded	9	22	8	137	0.21 $\pm$ 0.15	0.15 $\pm$ 0.034	0.61 $\pm$ 0.30	0.061 $\pm$ 0.060	0.27 $\pm$ 0.21
Upper Ashgill	raw	39	52	37	185	0.27 $\pm$ 0.079	0.25 $\pm$ 0.031	0.084 $\pm$ 0.15	0.19 $\pm$ 0.033	0.18 $\pm$ 0.20
Lower Caradoc	degraded	9	11	4	185	0.12 $\pm$ 0.087	0.059 $\pm$ 0.017	0.14 $\pm$ 0.18	0.048 $\pm$ 0.015	0.19 $\pm$ 0.17
Middle Caradoc	raw	39	53	34	198	0.33 $\pm$ 0.10	0.24 $\pm$ 0.029	0.069 $\pm$ 0.11	0.18 $\pm$ 0.028	0.19 $\pm$ 0.19
Upper Caradoc	degraded	7	11	8	198	0.071 $\pm$ 0.026	0.057 $\pm$ 0.018	0.13 $\pm$ 0.14	0.034 $\pm$ 0.014	0.20 $\pm$ 0.16
Lower Caradoc	raw	39	36	23	212	0.31 $\pm$ 0.11	0.16 $\pm$ 0.023	0.042 $\pm$ 0.061	0.17 $\pm$ 0.028	0.046 $\pm$ 0.81
Middle Caradoc	degraded	5	15	5	212	0.092 $\pm$ 0.062	0.069 $\pm$ 0.018	0.035 $\pm$ 0.065	0.024 $\pm$ 0.012	0.074 $\pm$ 0.13
Upper Caradoc	raw	48	13	29	200	0.26 $\pm$ 0.10	0.064 $\pm$ 0.017	0.017 $\pm$ 0.010	0.21 $\pm$ 0.029	0.030 $\pm$ 0.054
Lower Caradoc	degraded	8	3	7	200	0.078 $\pm$ 0.054	0.014 $\pm$ 0.009	0.014 $\pm$ 0.025	0.039 $\pm$ 0.015	0.025 $\pm$ 0.059
Middle Ashgill	raw	30	4	5	183	0.15 $\pm$ 0.060	0.22 $\pm$ 0.011	0.045 $\pm$ 0.056	0.15 $\pm$ 0.030	0.23 $\pm$ 0.26
Upper Ashgill	degraded	6	1	1	183	0.075 $\pm$ 0.043	0.005 $\pm$ 0.006	0.13 $\pm$ 0.090	0.032 $\pm$ 0.013	0.27 $\pm$ 0.23
Middle Ashgill	raw	80	18	38	107	0.54 $\pm$ 0.21	0.15 $\pm$ 0.040	0.089 $\pm$ 0.11	0.56 $\pm$ 0.067	0.39 $\pm$ 0.53
Upper Ashgill	degraded	14	6	5	107	0.26 $\pm$ 0.24	0.055 $\pm$ 0.026	0.081 $\pm$ 0.089	0.12 $\pm$ 0.032	0.70 $\pm$ 0.53

tween the size of the sample and the effective rate of preservation per Lmy. In this case, reducing the number of occurrences in an interval to one-eighth its initial value results in a decrease in preservation rate of only about one-half on average.

### Preliminary Application to Empirical Data

Given the (qualified) success of the modified survivorship analysis when it was subjected to the foregoing tests, it seems reasonable to apply this approach to gain insight into empirical origination and extinction patterns. A thorough study of taxonomic rates through the entire Phanerozoic would involve a number of sensitivity analyses that are beyond the scope of this paper. Such analyses will therefore be deferred to a future contribution, but some preliminary results will be presented here. In addition to uncertainty in the temporal pattern that results from the optimization procedure itself, there is uncertainty due to noise in the data. I have combined both sources of uncertainty by repeating the optimization procedure many times, each time using a different bootstrap sample of the raw taxonomic ranges. The mean parameter estimates are generally the same with bootstrapping as without, but the bootstrapping leads to larger standard errors.

Analysis of the Upper Cambrian through Lower Carboniferous portion of Sepkoski's unpublished compendium of marine animal genera (Sepkoski 1996, 1997) and his published family compendium (Sepkoski 1992, personal communication 1998) shows some substantial discrepancies between rates taken at face value and rates estimated from the survivorship analysis, of which I will mention just a few examples. Similar results are found for genera and families, so only the genus analyses are presented here (Fig. 7). (1) The estimated rates suggest the possibility of an Upper Silurian and/or Lower Devonian peak in origination that is not present when the data are taken at face value, although the face-value rate is well within the uncertainty of the estimated rate. The magnitude of uncertainty does not allow this peak to be constrained to a single stage. The height of the peak partly reflects the fact that rates are expressed per

Myr, and that the Pridoli is a short interval. (2) Estimated origination is higher than apparent origination in the Upper Devonian. This is consistent with the forward smearing of origination rates that is expected to result from general incompleteness, and with data showing that some taxa once thought to have originated in the Carboniferous are now known from the Devonian (Lane et al. 1997; Webster et al. 1999). (3) Whereas apparent extinction rate shows a peak in the Frasnian stage of the Upper Devonian, the estimated rate suggests that this peak may in fact have been in the succeeding Famennian stage. Determining the number of Upper Devonian extinction peaks and isolating the interval of maximal severity have been persistent problems in the study of this biotic crisis (McGhee 1996). If the data are taken at face value, whether analyzed at the stage level as here or at a finer level of stratigraphic resolution (McGhee 1996), a peak in the (upper) Frasnian is indicated. The survivorship analysis suggests an alternative interpretation of the data, however. It could be that a substantial part of the difficulty in pinning down the extinction event reflects the backward smearing of the event caused by incomplete preservation. If so, this represents a case in which the Signor-Lipps effect is detected and overcome by the survivorship method.

It is also noteworthy that the well-known paucity of lowest Silurian faunas (e.g., Rong and Harper 1999) is visible as a low estimated preservation rate in Figure 7. One would expect the end-Ordovician extinction to be exaggerated by this drop in preservation, but the extinction event nevertheless appears to be a robust feature rather than an artifact. Although incomplete sampling in the Llandoverly implies that many Ashgillian last appearances belong to genera that truly survived into the Silurian, the solution also implies that many Caradocian last appearances belong to genera that became extinct in the Ashgill. This last implication is also inherent in the tabulation of sampling intensity and the rate estimates based on Miller's data (Table 2, Fig. 5). Consideration of the end-Ordovician extinction illustrates two general points. First, the relationship between apparent and true rates depends not on single excursions in the qual-

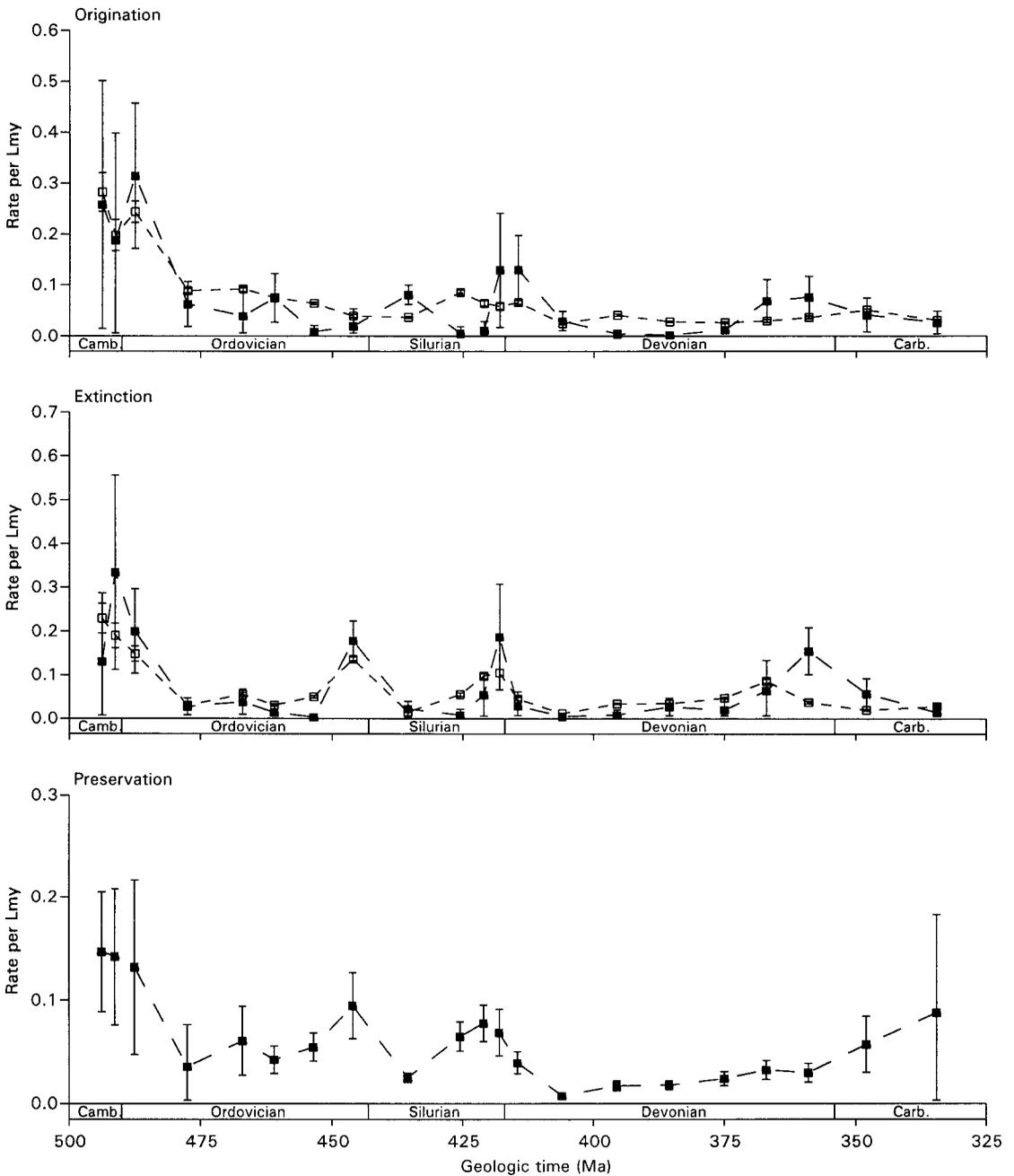


FIGURE 7. Rates of origination, extinction, and preservation estimated using first and last appearances from Sepkoski's unpublished data on marine animal and microfossil genera (Sepkoski 1996, 1997). In this and the next two figures, ranges are resolved to the level of Sepkoski's basic stratigraphic intervals, not divided into lower, middle, and upper parts (Sepkoski 1992: pp. 8–13), and the timescale is that of the Geological Society of America (Palmer and Geissman 1999), with modifications based on Bowring and Erwin 1998. A total of 11,668 genera were analyzed from the time span depicted in this figure. In this and the next two figures, open squares and short-dashed lines denote apparent rates, i.e., those with the record taken at face value, and solid squares and long-dashed lines denote rates estimated from survivorship analysis. Error bars on apparent rates are  $\pm$  one standard error, as determined from bootstrap resampling of genus ranges. These error bars are generally smaller than the plotted symbols. Error bars on estimated rates are also  $\pm$  one standard error, as determined from repeated bootstrap samples of genus ranges, each sample being subjected to the estimation procedure described herein (Appendix 3). See text for discussion.

ity of the record, but on the entire time series of taxonomic and preservational rates. Second, when apparent and estimated taxonomic rates both show a peak, chances are that the quality of preservation is also high, else the unusually high number of first or last appearances would not be observed. Origination in the Lower Ordovician and extinction in the Upper Silurian also demonstrate this point.

A striking feature of Figure 7 is that discrepancies between apparent and true taxonomic rates do not always reflect variation in the estimated quality of preservation in an obvious way. Variation in taxonomic rates itself distorts apparent rates, so it is of some interest to ask how estimated taxonomic rates would differ if we constrained preservation rates to be constant. As it turns out in this case, estimated taxonomic rates are rather similar in many respects whether we assume preservation is variable (Fig. 7) or constant at some value to be estimated (Fig. 8). This emphasizes the point that variable preservation is not the only factor that distorts apparent taxonomic rates. Figure 8 shows especially well the forward smearing of truly high origination rates to create apparently high rates afterwards (e.g., in the Lower and Middle Ordovician), and the backward smearing of truly high extinction rates (e.g., in the Upper Ordovician, Upper Silurian, and Upper Devonian).

The comparison between constant- and variable-preservation models suggests that, in this example, more of the variation in apparent taxonomic rates is statistically attributable to true variation in origination and extinction than to variation in preservation (see Foote 2000b for a similar result). It also raises the important question of whether we even need to take variable preservation into consideration. The mean and standard deviation of the sum of log deviations between observed data and expected probabilities, based on 200 independent bootstrap replicates, are  $215 \pm 11$  with the variable-rates model, and  $246 \pm 12$  with the constant-preservation model. Thus, there is only a 13% improvement in the goodness-of-fit when we take variable preservation into account, even though 22 parameters are added to the model. Considering log-transformed values of  $F_{\rightarrow ij}$ ,  $F_{\leftarrow ij}$ ,  $P_{\rightarrow ij}$ , and  $P_{\leftarrow ij}$ , the

coefficient of determination  $R^2$  for  $F_{\rightarrow ij}$  vs.  $P_{\rightarrow ij}$  is 0.68 for the constant-preservation model and 0.68 for the variable-preservation model. The corresponding values for  $R^2_{F_{\leftarrow ij}, P_{\leftarrow ij}}$  are 0.75 and 0.81. Although the standard statistical test for the significance of an increase in  $R^2$  with the addition of more parameters (Sokal and Rohlf 1981: p. 633) is not quite appropriate here, because the observed and predicted values are proportions with fixed sums, and although  $R^2$  is a different measure of fit than that actually optimized, the standard test, if applied, would show that the change in  $R^2$  is not statistically significant. From a purely statistical standpoint, there would be little need to invoke a complicated model with variable preservation if we had no prior reason to think that preservation in fact varied. Considering our prior knowledge (Figs. 5, 6), however, there is good reason, geological rather than statistical, for preferring the model with variable rates. It is also worth noting that fitting the data with a model that assumes complete preservation yields substantially worse fit (sum of log deviations  $345 \pm 14$ ,  $R^2_{F_{\rightarrow ij}, P_{\rightarrow ij}} = 0.33$ ,  $R^2_{F_{\leftarrow ij}, P_{\leftarrow ij}} = 0.55$ ), even though the number of parameters is only one less than the model of uniform but incomplete preservation. Thus, one could conceivably argue, on statistical grounds, that the temporal pattern of preservation is indistinguishable from uniform. It would be difficult, however, to support the notion that the record is sufficiently complete to be taken at face value.

Is it plausible to maintain that taxonomic rates were in fact constant through much of the Paleozoic, with apparent variation attributable solely to variation in the quality of preservation? Figure 9 shows a solution in which origination and extinction rates were each constrained to be constant at values to be determined. Estimated preservation rates bear some similarities to those of the variable-rates model, such as the dip in the Llandovery and the rise through the Devonian. This solution yields a mean log deviation of  $271 \pm 11$ , about 26% higher than the variable-rates model ( $R^2_{F_{\rightarrow ij}, P_{\rightarrow ij}} = 0.69$ ,  $R^2_{F_{\leftarrow ij}, P_{\leftarrow ij}} = 0.75$ ). This may not seem like a great difference in fit relative to the variable-rates model, but the constant-rate solution implies a logical inconsistency. If true taxonomic

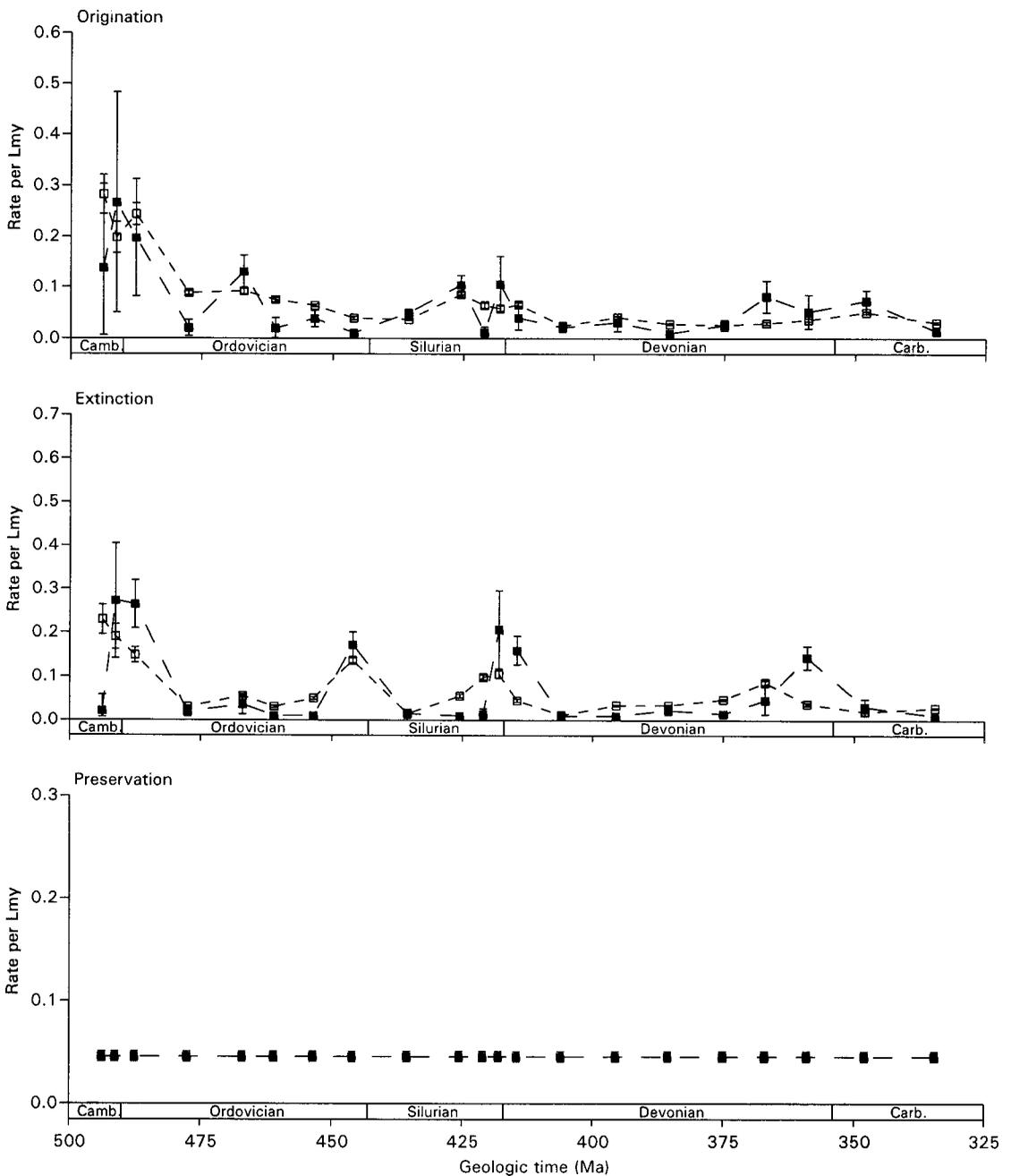


FIGURE 8. Rates of origination, extinction, and preservation estimated using the data of Figure 7 and assuming constant preservation rate. Results plotted using same conventions and at same scale as Figure 7. Note the similarity in taxonomic rates between this figure and Figure 7.

rates are constant, apparent rates should fluctuate around the true rates, with the same long-term average (Foote 2000a). Neither origination rate nor extinction rate shows this feature with the constant-rates solution. The contrast between different models clearly deserves more

attention, and it would be premature to discard the hypothesis that one or more of the rates was constant for much of the Paleozoic. For the moment, however, there are reasons to prefer the model in which origination, extinction, and preservation rates all vary.

### Relaxing the Assumption of Taxonomic Homogeneity

The approach to rate estimation outlined above assumes that all taxa are characterized by the same rates, an assumption that is obviously unrealistic. To explore the robustness of the method when this assumption is violated, I analyzed simulated data representing two different cases of taxonomic heterogeneity. In the first simulation, there are three groups with the same origination and extinction rates, and with preservation rates that have the same temporal pattern but are in the ratio of 1:2:4. When the simulated data are analyzed as if they were homogeneous, the origination and extinction rates are generally estimated accurately, and the temporal pattern of preservation is estimated reasonably well (Fig. 10).

In the second simulation, all three groups have the same preservation rate and the same temporal pattern of origination and extinction, whereas their taxonomic rates are in the ratio of 1:2:4. Although some estimated taxonomic rates are below the true values for all of the groups, the temporal pattern is captured accurately for the most part, especially the larger changes in rate (Fig. 11). The estimate of preservation rate shows a conspicuous feature: the temporal pattern is captured well, but the rates are systematically underestimated (cf. Figs. 5A, 6A). This may reflect the fact that higher taxonomic rates imply more taxa produced with shorter durations and thus a lower proportion of taxa preserved. The estimated preservation rate is something of an "effective rate" that is weighted by the relative completeness of each group (Foote and Raup 1996; Foote 1997).

These models of taxonomic heterogeneity are oversimplified, although not without foundation. For example, Raup and Boyajian (1988) showed that major biological groups, although having characteristically different average rates of extinction, nevertheless tend to exhibit extinction trajectories that track each other over the Phanerozoic (a pattern that could, admittedly, partly reflect variation in the quality of the fossil record). Likewise, animal classes show parallel changes in extinc-

tion rate from one paleoenvironmental setting to the next in the Paleozoic (Sepkoski 1987). The extent to which higher taxa manifest concordant temporal patterns in the quality of their fossil records depends in part on whether they inhabit similar environments and are reported with similar fidelity in the paleontological literature. In the Ordovician data analyzed above, the three principal groups—brachiopods, benthic mollusks, and trilobites—were deliberately chosen as the principal elements of Sepkoski's evolutionary faunas that are common in a wide range of benthic habitats and are commonly reported in faunal lists. It is perhaps not surprising, therefore, that they tend to show similar temporal patterns of completeness through the Ordovician (Fig. 12). Although I would not claim at this point that the similarity of preservational patterns is a general feature of the invertebrate fossil record, and although the foregoing analysis may be exploring heterogeneity at an unreasonably coarse scale, the Ordovician data at least leave open the possibility that the style of taxonomic heterogeneity modeled here may not be completely unrealistic. If we had strong prior reason to suspect the presence of different preservational categories that exhibit uncorrelated temporal patterns, it would make sense to analyze the data for these groups separately. This always involves a trade-off, because making groups more homogeneous also reduces their sample size and therefore the statistical power to estimate rates that are of interest. It is also worth noting that long-term secular changes in preservation, such as shifts in the proportion of calcitic versus aragonitic genera (Wilkinson 1979; Kidwell and Brenchley 1996; Harper et al. 1997; Carter et al. 1998; Stanley and Hardie 1998; Harper 2000), are expected to be less important in distorting taxonomic rates than is short-term variation (Foote 2000a). This is simply because a cohort of genera is affected by variation in preservation only in proportion to the number of those genera still extant at the time of the change. Regarding apparent taxonomic rates for Ordovician genera, for example, preservation in the Silurian is quite relevant, because many Ordovician genera survive into the Silurian,

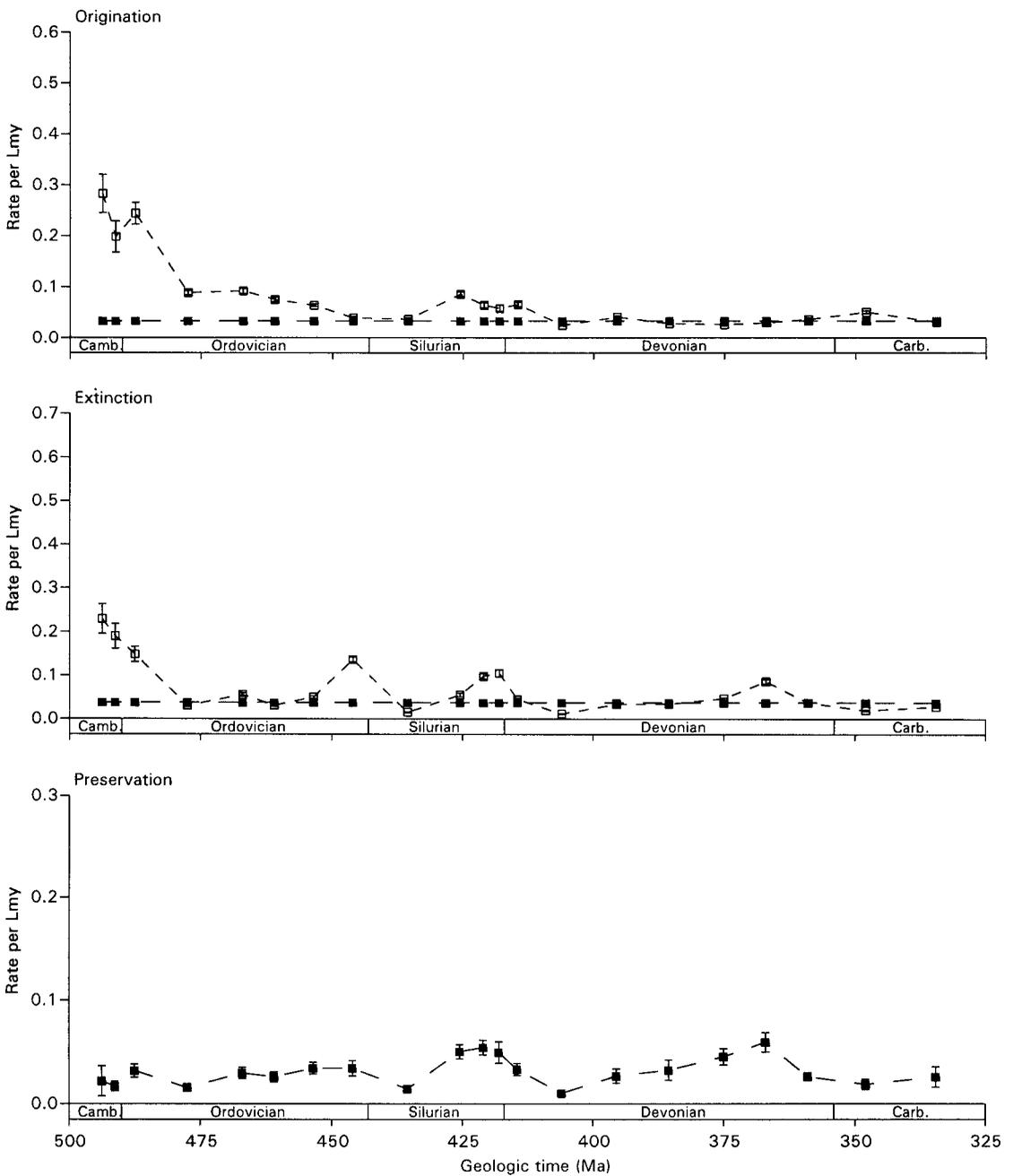


FIGURE 9. Rates of origination, extinction, and preservation estimated using the data of Figure 7 and assuming constant origination and extinction rates. Results plotted using same conventions and at same scale as Figure 7. Note that apparent taxonomic rates are generally higher than estimated rates, which is contrary to expectation if the constant-rate model were reasonable.

but preservation in the Mesozoic is relatively unimportant.

### Discussion and Conclusions

The modified survivorship analysis presented herein is aimed at estimating true rates

of origination, extinction, and preservation when one or more of these rates varies, because variation in any of the rates distorts the apparent record of origination and/or extinction. If prior knowledge of preservation is available or assumed, then this can be used to

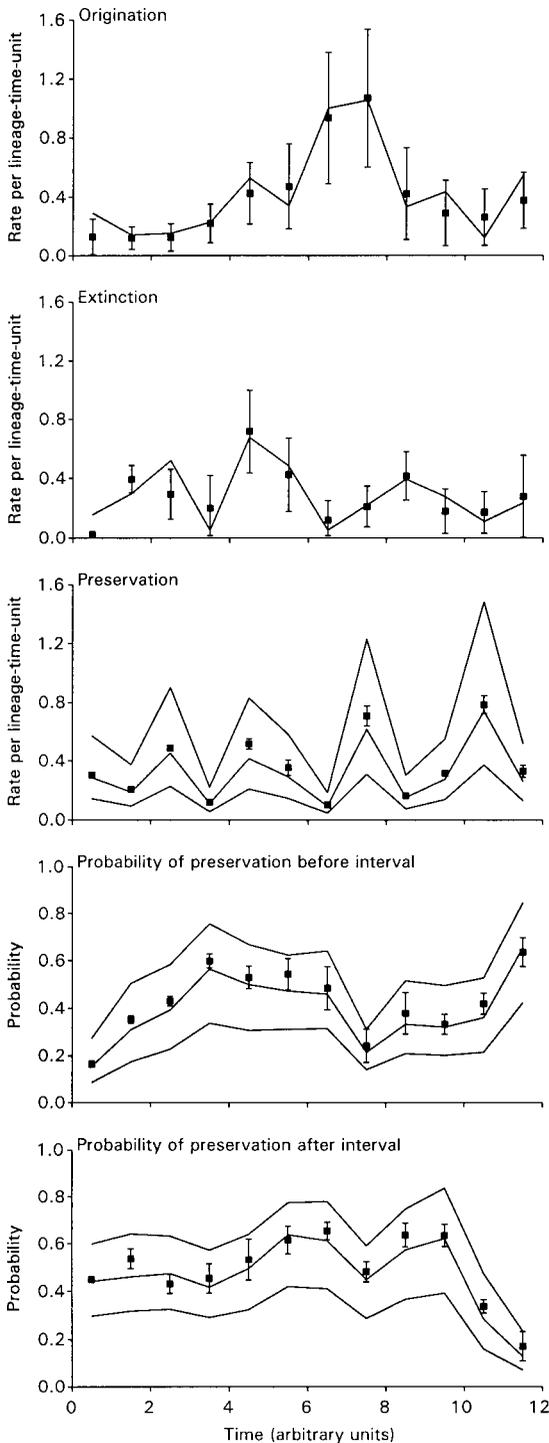


FIGURE 10. Estimation of origination, extinction, and preservation rates from first and last appearances of simulated data having taxonomic heterogeneity in preservation rate. No knowledge of this heterogeneity is used in the rate estimation. See Figure 1 for explanation of lines and symbols. Three solid lines in middle panel show preservation rates for three groups of taxa, which

obtain tighter constraints on estimated taxonomic rates. Even if preservation is known or assumed to be constant, or if it is in effect made constant by standardized sampling of available data, rates of origination and extinction cannot be taken at face value (cf. Alroy 1996, 1998; Johnson and McCormick 1999; Alroy et al. 2000). The approach developed herein should therefore be of general utility in the study of taxonomic rates. There are other important questions to which the estimation of preservation rates, once obtained, can be applied. Not least among these are phylogenetic issues (Wagner 1998, 1999; Smith 2000; Wagner and Sidor 2000). For example, stratocladistics uses taphonomic controls to assess whether a postulated lineage in a candidate evolutionary tree, where none is known from the fossil record, should be considered to contribute to stratigraphic parsimony debt (Fisher 1994; Clyde and Fisher 1997; Fox et al. 1999; Bodenbender and Fisher 2001). Knowing something of temporal variation in the quality of preservation can permit stratigraphic parsimony debt to be weighted on a continuous scale rather than incremented in discrete units.

The interpretation of preservation and completeness requires some comment. As pointed out earlier (Foote 1996, 1997, 2001b; Foote and Raup 1996), estimates of preservation such as those presented herein tend to reflect the quality of the fossil record where it exists rather than the overall completeness of the record. This can be thought of as local rather than global completeness. The approach developed herein leads to estimates of the overall probability of preservation before or after some point in time, for taxa extant at that time. This should strictly be interpreted to mean for taxa extant at that time and in the same regions that have left some fossil record at all. Global completeness estimates do not result from this

←

have rates that follow the same temporal pattern and are in the ratio of 1:2:4. Rates are drawn from an exponential distribution with stationary mean. Each group has starting diversity of 2000; numbers of lineages produced and preserved are 10,354, 16,805, and 22,003. Temporal patterns are detected despite heterogeneity in the data.

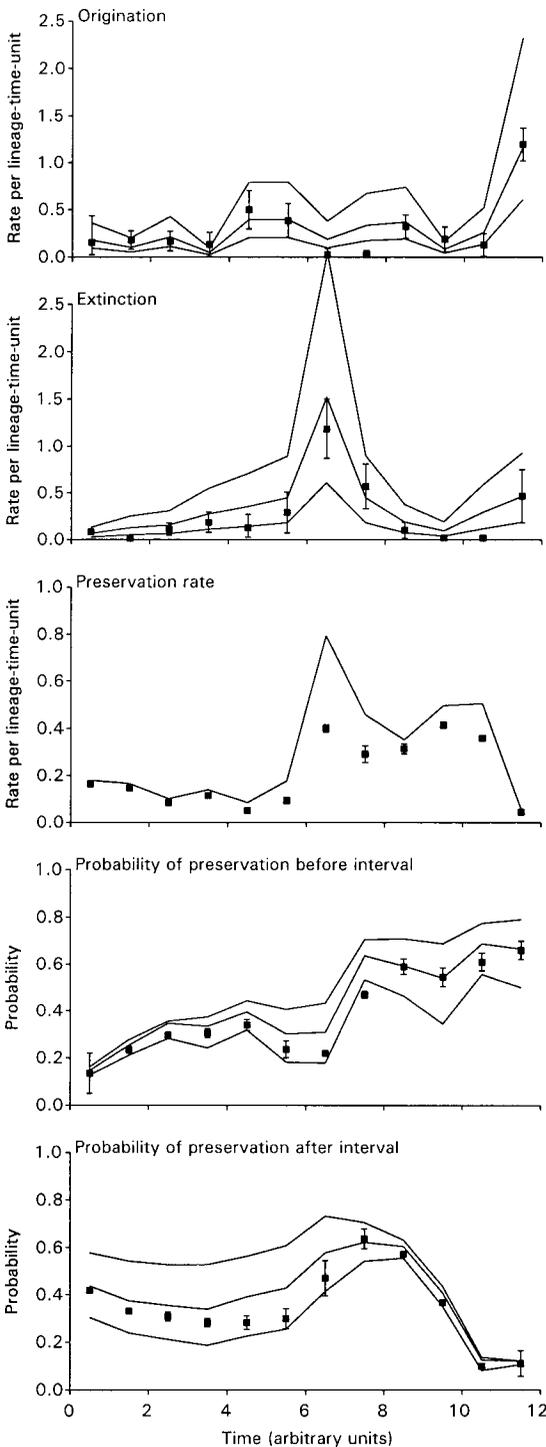


FIGURE 11. Estimation of origination, extinction, and preservation rates from first and last appearances of simulated data having taxonomic heterogeneity in taxonomic rates. No knowledge of this heterogeneity is used in the rate estimation. See Figure 1 for explanation of lines and symbols. Three solid lines in top two panels show taxonomic rates for three groups of taxa, which

approach, so it would not be appropriate, for example, to use the estimated probabilities of preservation to yield estimates of true, global taxonomic richness. For such an estimate we would need not only a measure of the intrinsic quality of the record where it exists, which is useful when considering origination and extinction measures that depend on relative numbers of taxa, but also some correction for the sheer amount of record preserved (Raup 1972, 1976; Alroy et al. 2001; Peters and Foote 2001; Smith 2001). Nor would it be appropriate to use estimates of intrinsic completeness as a straightforward measure of how overall paleontological completeness, a global measure, may have varied over geologic time (cf. Benton et al. 2000).

The modified survivorship approach is not without its limitations, of which four deserve to be singled out:

1. Because many of the constraints placed on estimated rates by observed first and last appearances in effect depend on ratios among numbers of taxa in different categories ( $X_{bL}$ ,  $X_{FL}$ ,  $X_{FL}$ , and  $X_{bI}$ ), rather large numbers of taxa can be required to obtain stable rate estimates. At what point, if any, the uncertainty in the estimation procedure outweighs the inaccuracy inherent in taking rates at face value is a question that deserves further attention.

2. The assumption that rates are stochastically constant within an interval becomes less reasonable as interval length increases. Moreover, the observed ratios  $X_{bL}/(X_{bL} + X_{FL})$  and  $X_{FL}/(X_{FL} + X_{FL})$  provide ever looser constraints on  $P_B$  and  $P_A$  as interval length increases. It would therefore seem desirable to use the finest stratigraphic resolution available. This, however, generally leads to lower overall numbers of taxa used (as those with poorly resolved stratigraphic ranges may need to be omitted from analysis) and to lower numbers for each interval. Where to strike the balance

←

have rates that follow the same temporal pattern and are in the ratio of 1:2:4. Rates are drawn from an exponential distribution with stationary mean. Each group has starting diversity of 2000; numbers of lineages produced and preserved are 3501, 4990, and 7452. Temporal patterns in the data are shown in Figure 11. The three solid lines in the top two panels show taxonomic rates for three groups of taxa, which

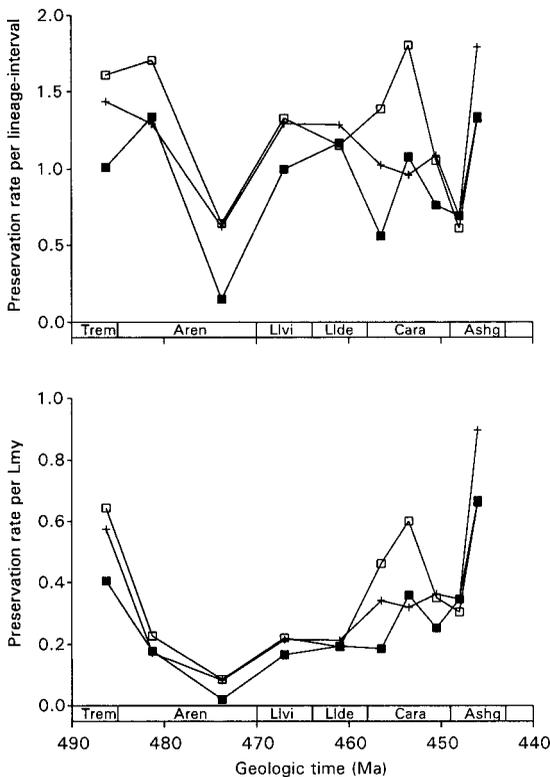


FIGURE 12. Temporal patterns of preservation for the three components of Miller's Ordovician data, measured directly from internal occurrences as  $r = -\ln(1 - X_{bt,samp}/X_{it})$  (see Figs. 5, 6). Open squares = brachiopods; closed squares = benthic mollusks; crosses = trilobites. Temporal patterns are largely in agreement, whether rate of preservation is expressed per stratigraphic interval (top panel) or per million years (bottom panel).

between the need for fine resolution and the need for stable proportions of taxa should also be examined further.

3. The assumption that all taxa are governed by the same rates of origination, extinction, and preservation is of course violated in real data. As discussed above, this may not be critical if different taxonomic groups exhibit similar temporal patterns, and there is some evidence that they may do so. Again, there is a trade-off between making groups more homogeneous and making them large enough to have sufficient data to permit analysis. Where the balance lies is an important question. One approach to addressing this would involve testing whether adding extra parameters (i.e., different rates of origination, extinction, and preservation for each group) improves the fit between model and data sufficiently to war-

rant a multigroup rather than a single-group model.

4. The survivorship approach does not correct for all possible distortions in data, only those that arise from incomplete and variable preservation coupled with variation in taxonomic rates from interval to interval. An especially important problem is that variation in taxonomic practice can affect apparent rates. For example, it is possible that taxonomic rates decline artificially toward the Recent, as fossil taxa are increasingly likely to be linked to extant genera, but it is also possible that these rates increase artificially toward the Recent, as morphological knowledge of extant groups may permit finer taxonomic partitioning of their fossil representatives (Raup 1979). Although the method presented herein can, in principle, detect and correct for the effects of increasing preservation probability toward the Recent and of nearly complete knowledge of the living fauna, it does not account for those aspects of the Pull of the Recent that reflect taxonomic practice. Likewise, a change in the breadth of the genus concept can cause a spurious change in average rates (Foote 1988), and a known extinction boundary can be exaggerated if different workers study the fauna on either side of the boundary and therefore are unlikely to identify lineages that cross the boundary (Whittington 1954; Fortey 1983, 1989). Although some temporal patterns of origination and extinction are sufficiently robust that they are detected with or without taxonomic standardization (Adrain and Westrop 2000), such standardization at least has the potential to help eliminate spurious variation in apparent rates (Culver et al. 1987).

One might add that an additional limitation of this approach is the mere fact that it involves a model with some simplified assumptions about rates of evolution and preservation. I would contend that the assessment of temporal variation in taxonomic rates that involves taking the record at face value involves a potentially more problematic, tacit assumption. This assumption is not that the quality of the record is constant, for even with constant preservation the true variation in origination and extinction leads to distorted taxonomic rates. Rather, it is that stratigraphic ranges are

so complete that they do not require correction. The preliminary analyses presented above suggest, reassuringly, that some variations in rate, such as high origination in the early Ordovician and high extinction in the late Ordovician, are so large that they will be detected whether one relies upon apparent or corrected rates, and whether one assumes constant or variable preservation. Thus, I do not mean to suggest that temporal patterns in taxonomic rates can never be taken at face value. A number of studies, including some of the author, have nevertheless sought to interpret variations in origination and extinction more subtle than those associated with the Ordovician radiation and the end-Ordovician mass extinction (Foote 2000b; Kirchner and Weil 2000). Whether these interpretations will withstand the scrutiny permitted by the modified survivorship approach is a question meriting further attention.

### Acknowledgments

For kindly allowing access to data, I thank Arnold I. Miller, the organizers of the Paleobiology Database (URL <http://flatpebble.nceas.ucsb.edu>), and the late J. J. Sepkoski Jr. I also thank D. C. Fisher, D. Jablonski, S. M. Kidwell, and J. S. Trefil for discussion, and S. R. Connolly, R. H. De Simone, S. M. Holland, A. I. Miller, M. E. J. Newman, S. E. Peters, and D. M. Raup for reading the manuscript. This work was supported by the National Science Foundation (grant EAR-9506568). This is Paleobiology Database publication number 3.

### Literature Cited

- Adrain, J. M., and S. R. Westrop. 2000. An empirical assessment of toxic paleobiology. *Science* 289:110–112.
- Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- . 1998. Equilibrium diversity dynamics in North American mammals. Pp. 233–287 in McKinney and Drake 1998.
- Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American mammalian evolution. In D. H. Erwin and S. L. Wing, eds. *Deep time: Paleobiology's perspective*. *Paleobiology* 26(Suppl. to No. 4):259–288.
- Alroy, J., C. R. Marshall, R. K. Bambach, K. Bezusko, M. Foote, F. T. Fürsich, T. A. Hansen, S. M. Holland, L. C. Ivany, D. Jablonski, D. K. Jacobs, D. C. Jones, M. A. Kosnik, S. Lidgard, S. Low, A. I. Miller, P. M. Novack-Gottshall, T. D. Olszewski, M. E. Patzkowsky, D. M. Raup, K. Roy, J. J. Sepkoski Jr., M. G. Sommers, P. J. Wagner, and A. Webber. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences USA* 98:6261–6266.
- Baumiller, T. K. 1993. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. *Paleobiology* 19:304–322.
- Benton, M. J., M. A. Wills, and R. Hitchin. 2000. The quality of the fossil record through time. *Nature* 403:534–537.
- Bodenbender, B. E., and D. C. Fisher. 2001. Stratocladistic analysis of blastoid phylogeny. *Journal of Paleontology* 75:351–369.
- Bowring, S. A., and D. H. Erwin. 1998. A new look at evolutionary rates in deep time: uniting paleontology and high-precision geochronology. *GSA Today* 8(9):1–8.
- Brett, C. E. 1998. Sequence stratigraphy, paleoecology, and evolution: biotic clues and responses to sea-level fluctuation. *Palaos* 13:241–262.
- Budd, A. F., and K. G. Johnson. 1999. Origination preceding extinction during late Cenozoic turnover of Caribbean reefs. *Paleobiology* 25:188–200.
- Carter, J. G., E. Barrera, and M. J. S. Tevesz. 1998. Thermal potentiation and mineralogical evolution in the Bivalvia (Mollusca). *Journal of Paleontology* 72:991–1010.
- Cheetham, A. H., and J. B. C. Jackson. 1998. The fossil record of cheilostome Bryozoa in the Neogene and Quaternary of tropical America: adequacy for phylogenetic and evolutionary studies. Pp. 227–242 in S. K. Donovan and C. R. C. Paul, eds. *The adequacy of the fossil record*. Wiley, Chichester, England.
- Clyde, W. C., and D. C. Fisher. 1997. Comparing the fit of stratigraphic and morphologic data in phylogenetic analysis. *Paleobiology* 23:1–19.
- Connolly, S. R., and A. I. Miller. 2001. Global Ordovician faunal transitions in the marine benthos. I. Proximal causes. *Paleobiology* 27:779–795 (this volume).
- Conroy, M. J., and J. D. Nichols. 1984. Testing for variation in taxonomic extinction probabilities: a suggested methodology and some results. *Paleobiology* 10:328–337.
- Culver, S. J., M. A. Buzas, and L. S. Collins. 1987. On the value of taxonomic standardization in evolutionary studies. *Paleobiology* 13:169–176.
- Edwards, A. W. F. 1992. *Likelihood*, expanded ed. Johns Hopkins University Press, Baltimore.
- Fisher, D. C. 1994. Stratocladistics: morphological and temporal patterns and their relation to phylogenetic process. Pp. 133–171 in L. Grande and O. Rieppel, eds. *Interpreting the hierarchy of nature*. Academic Press, San Diego.
- Foote, M. 1988. Survivorship analysis of Cambrian and Ordovician trilobites. *Paleobiology* 14:258–271.
- . 1996. On the probability of ancestors in the fossil record. *Paleobiology* 22:141–151.
- . 1997. Estimating taxonomic durations and preservation probability. *Paleobiology* 23:278–300.
- . 1999. Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. *Paleobiology Memoirs No. 1*. *Paleobiology* 25(Suppl. to No. 2):1–115.
- . 2000a. Origination and extinction components of taxonomic diversity: general problems. In D. H. Erwin and S. L. Wing, eds. *Deep time: Paleobiology's perspective*. *Paleobiology* 26(Suppl. to No. 4):74–102.
- . 2000b. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- . 2001a. Evolutionary rates and the age distributions of living and extinct taxa. Pp. 245–294 in J. B. C. Jackson, S. Lidgard, and F. K. McKinney, eds. *Evolutionary patterns: growth, form, and tempo in the fossil record*. University of Chicago Press, Chicago.
- . 2001b. Estimating completeness of the fossil record. Pp.

- 504–508 in D. E. G. Briggs and P. R. Crowther, eds. *Palaeobiology II*. Blackwell Scientific, Oxford.
- Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 22:121–140.
- Foote, M., and J. J. Sepkoski Jr. 1999. Absolute measures of the completeness of the fossil record. *Nature* 398:415–417.
- Fortey, R. A. 1983. Cambrian-Ordovician trilobites from the boundary beds in western Newfoundland and their phylogenetic significance. *Special Papers in Palaeontology* 30:179–211.
- . 1989. There are extinctions and extinctions: examples from the Lower Palaeozoic. *Philosophical Transactions of the Royal Society of London B* 325:327–355.
- Fox, D. L., D. C. Fisher, and L. R. Leighton. 1999. Reconstructing phylogeny with and without temporal data. *Science* 284:1816–1819.
- Harper, E. M. 2000. Are calcitic layers an effective adaptation against shell dissolution in the Bivalvia? *Journal of Zoology* 251:179–186.
- Harper, E. M., T. J. Palmer, and J. R. Alphey. 1997. Evolutionary response by bivalves to changing Phanerozoic sea-water chemistry. *Geological Magazine* 134:403–407.
- Holland, S. M. 1995. The stratigraphic distribution of fossils. *Paleobiology* 21:92–109.
- . 2000. The quality of the fossil record: a sequence-stratigraphic perspective. In D. H. Erwin and S. L. Wing, eds. *Deep time: Paleobiology's perspective*. *Paleobiology* 26(Suppl. to No. 4):148–168.
- Holland, S. M., and M. E. Patzkowsky. 1999. Models for simulating the fossil record. *Geology* 27:491–494.
- Horowitz, A. S., R. F. Blakely, and D. B. Macurda Jr. 1985. Taxonomic survivorship within the Blastozoa (Echinodermata). *Journal of Paleontology* 59:543–550.
- Johnson, K. G., and T. McCormick. 1999. The quantitative description of biotic change using palaeontological databases. Pp. 227–247 in D. A. T. Harper, ed. *Numerical palaeobiology*. Wiley, Chichester, England.
- Jones, D. S., and D. Nicol. 1986. Origination, survivorship, and extinction of rudist taxa. *Journal of Paleontology* 60:107–115.
- Kendall, D. G. 1948. On the generalized "birth-and-death" process. *Annals of Mathematical Statistics* 19:1–15.
- Kidwell, S. M., and P. J. Brenchley. 1996. Evolution of the fossil record: thickness trends in marine skeletal accumulations and their implications. Pp. 290–336 in D. Jablonski, D. H. Erwin, and J. H. Lipps, eds. *Evolutionary paleobiology*. University of Chicago Press; Chicago.
- Kirchner, J. W., and A. Weil. 2000. Delayed biological recovery from extinctions throughout the fossil record. *Nature* 404:177–180.
- Koch, C. F. 1991. Species extinctions across the Cretaceous-Tertiary boundary: observed patterns versus predicted sampling effects, stepwise or otherwise? *Historical Biology* 5:355–361.
- Koch, C. F., and J. P. Morgan. 1988. On the expected distribution of species' ranges. *Paleobiology* 14:126–138.
- Lane, N. G., J. A. Waters, and C. G. Maples. 1997. Echinoderm faunas of the Hongguleleng Formation, Late Devonian (Famennian), Xinjiang-Uygur Autonomous Region, People's Republic of China. *Paleontological Society Memoir* 47. *Journal of Paleontology* 71(Suppl. to No. 2).
- Markwick, P. J. 1998. Crocodilian diversity in space and time: the role of climate in paleoecology and its implications for understanding K/T extinctions. *Paleobiology* 24:470–497.
- Marshall, C. R. 1997. Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology* 23:165–173.
- McGhee, G. R. 1996. *The Late Devonian mass extinction*. Columbia University Press, New York.
- McKinney, M. L., and J. A. Drake, eds. 1998. *Biodiversity dynamics: turnover of populations, taxa, and communities*. Columbia University Press, New York.
- Meldahl, K. H. 1990. Sampling, species abundance, and the stratigraphic signature of mass extinction: a test using Holocene tidal flat molluscs. *Geology* 18:890–893.
- Miller, A. I. 1997a. Dissecting global diversity patterns: examples from the Ordovician Radiation. *Annual Review of Ecology and Systematics* 28:85–104.
- . 1997b. A new look at age and area: the geographic and environmental expansion of genera during the Ordovician Radiation. *Paleobiology* 23:410–419.
- . 1998. Biotic transitions in global marine diversity. *Science* 281:1157–1160.
- Miller, A. I., and M. Foote. 1996. Calibrating the Ordovician Radiation of marine life: implications for Phanerozoic diversity trends. *Paleobiology* 22:304–309.
- Miller, A. I., and S. G. Mao. 1995. Association of orogenic activity with the Ordovician Radiation of marine life. *Geology* 23:305–308.
- . 1998. Scales of diversification and the Ordovician Radiation. Pp. 288–310 in McKinney and Drake 1998.
- Newell, N. D. 1967. *Revolutions in the history of life*. Geological Society of America Special Paper 89:63–91.
- Nichols, J. D., and K. H. Pollock. 1983. Estimating taxonomic diversity, extinction rates, and speciation rates from fossil data using capture-recapture models. *Paleobiology* 9:150–163.
- Nichols, J. D., R. W. Morris, C. Brownie, and K. H. Pollock. 1986. Sources of variation in extinction rates, turnover, and diversity of marine invertebrate families during the Paleozoic. *Paleobiology* 12:421–432.
- Palmer, A. R., and J. Geissman. 1999. 1999 geologic time scale. Geological Society of America, Boulder, Colo.
- Paul, C. R. C. 1982. The adequacy of the fossil record. Pp. 75–117 in K. A. Joysey and A. E. Friday, eds. *Problems of phylogenetic reconstruction*. Academic Press, London.
- . 1998. Adequacy, completeness and the fossil record. Pp. 1–22 in S. K. Donovan and C. R. C. Paul, eds. *The Adequacy of the fossil record*. Wiley, Chichester, U.K.
- Pearson, P. N. 1992. Survivorship analysis of fossil taxa when real-time extinction rates vary: the Paleogene planktonic foraminifera. *Paleobiology* 18:115–131.
- Pease, C. M. 1985. Biases in the durations and diversities of fossil taxa. *Paleobiology* 11:272–292.
- Peters, S. E., and M. Foote. 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology* 27:583–601 (this volume).
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical recipes in C*, 2d ed. Cambridge University Press, Cambridge.
- Rampino, M. R., and A. C. Adler. 1998. Evidence for abrupt latest Permian mass extinction of foraminifera: Results of tests for the Signor-Lipps effect. *Geology* 26:415–418.
- Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. *Science* 177:1065–1071.
- . 1976. Species diversity in the Phanerozoic: an interpretation. *Paleobiology* 2:289–297.
- . 1978. Cohort analysis of generic survivorship. *Paleobiology* 4:1–15.
- . 1979. Biases in the fossil record of species and genera. *Bulletin of the Carnegie Museum of Natural History* 13:85–91.
- . 1985. Mathematical models of cladogenesis. *Paleobiology* 11:42–52.
- . 1986. Biological extinction in Earth history. *Science* 231:1528–1533.
- Raup, D. M., and G. E. Boyajian. 1988. Patterns of generic extinction in the fossil record. *Paleobiology* 14:109–125.
- Raup, D. M., and J. J. Sepkoski Jr. 1982. Mass extinctions in the marine fossil record. *Science* 215:1501–1503.

- . 1984. Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences USA* 81:801–805.
- Raup, D. M., S. J. Gould, T. J. M. Schopf, and D. S. Simberloff. 1973. Stochastic models of phylogeny and the evolution of diversity. *Journal of Geology* 81:525–542.
- Raymond, A., and C. Metz. 1995. Laurussian land-plant diversity during the Silurian and Devonian: mass extinction, sampling bias, or both? *Paleobiology* 21:74–91.
- Rong, J.-Y., and D. A. T. Harper. 1999. Brachiopod survival and recovery from the latest Ordovician mass extinctions in South China. *Geological Journal* 34:321–348.
- Sepkoski, J. J., Jr. 1975. Stratigraphic biases in the analysis of taxonomic survivorship. *Paleobiology* 1:343–355.
- . 1987. Environmental trends in extinction during the Paleozoic. *Science* 235:64–66.
- . 1992. A compendium of fossil marine animal families, 2d ed. Milwaukee Public Museum Contributions in Biology and Geology 83:1–156.
- . 1996. Patterns of Phanerozoic extinctions: a perspective from global databases. Pp. 35–52 in O. H. Walliser, ed. *Global events and event stratigraphy*. Springer, Berlin.
- . 1997. Biodiversity: past, present, and future. *Journal of Paleontology* 71:533–539.
- . 1998. Rates of speciation in the fossil record. *Philosophical Transactions of the Royal Society of London B* 353:315–326.
- Shaw, A. B. 1964. *Time in stratigraphy*. McGraw-Hill, New York.
- Signor, P. W., III, and J. H. Lipps. 1982. Sampling bias, gradual extinction patterns and catastrophes in the fossil record. *Geological Society of America Special Paper* 190:291–296.
- Smith, A. B. 2000. Stratigraphy in phylogeny reconstruction. *Journal of Paleontology* 74:763–766.
- . 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. *Philosophical Transactions of the Royal Society of London B* 356:1–17.
- Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*, 2d ed. W. H. Freeman, San Francisco.
- Solow, A. R., and W. Smith. 1997. On fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 23:271–278.
- Stanley, S. M., and L. A. Hardie. 1998. Secular oscillations in the carbonate mineralogy of reef-building and sediment-producing organisms driven by tectonically forced shifts in seawater chemistry. *Palaeogeography, Palaeoclimatology, Palaeoecology* 144:3–19.
- Strauss, D., and P. M. Sadler. 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Mathematical Geology* 21:411–427.
- Van Valen, L. M. 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- . 1984. A resetting of Phanerozoic community evolution. *Nature* 307:50–52.
- Wagner, P. J. 1997. Patterns of morphologic diversification among the Rostroconchia. *Paleobiology* 23:115–150.
- . 1998. A likelihood approach for evaluating estimates of phylogenetic relationship among fossil taxa. *Paleobiology* 24:430–449.
- . 1999. The quality of the fossil record and the accuracy of phylogenetic inferences about sampling and diversity. *Systematic Biology* 49:65–86.
- . 2000. Likelihood tests of hypothesized durations: determining and accommodating biasing factors. *Paleobiology* 26:431–449.
- Wagner, P. J., and C. A. Sidor. 2000. Age rank/clade rank metrics—sampling, taxonomy, and the meaning of “stratigraphic consistency.” *Systematic Biology* 49:463–479.
- Webster, G. D., D. J. Hafley, D. B. Blake, and A. Glass. 1999. Crinoids and stelleroids (Echinodermata) from the Broken Rib Member, Dyer Formation (Late Devonian, Famennian) of the White River Plateau, Colorado. *Journal of Paleontology* 73:461–486.
- Weiss, R. E., and C. R. Marshall. 1999. The uncertainty in the true end point of a fossil’s stratigraphic range when stratigraphic sections are sampled discretely. *Mathematical Geology* 31:435–453.
- Whittington, H. B. 1954. Status of invertebrate paleontology, 1953, VI. Arthropoda: Trilobita. *Bulletin of the Museum of Comparative Zoology, Harvard College* 112:193–200.
- Wilkinson, B. H. 1979. Biomineralization, paleoceanography, and the evolution of calcareous marine organisms. *Geology* 7:524–527.

## Appendix 1

Constraints on  $P_{Bv}$ ,  $P_{Av}$ ,  $p$ ,  $q$ , and  $r$ 

Just as the observed taxa known to have existed during an interval of time belong to one of four categories with counts  $X_{BL}$ ,  $X_{FL}$ ,  $X_{FL}$ , and  $X_{BT}$ , all taxa that in reality existed during an interval, whether or not they are known, belong to analogous categories with counts  $N_{BL}$ ,  $N_{FL}$ ,  $N_{FL}$ , and  $N_{BT}$ .  $P_B$  was defined above as the probability that a taxon extant at the start of the interval is preserved sometime before it, and  $P_A$  as the probability that a taxon extant at the end of the interval is preserved sometime after it. In addition, there are four relevant probabilities of preservation during the interval. These depend on how much of the interval is truly spanned by the complete duration of the taxon, which in turn depends on the rate of origination and/or extinction during the interval (Foote 2000a). Let  $P_{D|BL}$  be the probability that a taxon is preserved sometime during the interval given that it truly originated before the interval and truly became extinct within it, i.e., that it truly crossed only the bottom interval boundary. Similarly, let  $P_{D|FL}$ ,  $P_{D|FL}$ , and  $P_{D|BT}$  be the probabilities of preservation for taxa truly crossing the top boundary only, truly confined to the interval, and truly crossing both boundaries. Then the expected  $X$ -values are related to the  $N$ -values as follows:

$$X_{BL} = N_{BL}P_B P_{D|BL} + N_{BT}P_B(1 - P_A)P_{D|BT}, \quad (1a)$$

$$X_{FL} = N_{FL}P_A P_{D|FL} + N_{BT}P_A(1 - P_B)P_{D|BT}, \quad (1b)$$

$$X_{FL} = N_{BL}(1 - P_B)P_{D|BL} + N_{FL}(1 - P_A)P_{D|FL} + N_{FL}P_{D|FL} + N_{BT}(1 - P_B)(1 - P_A)P_{D|BT}, \quad \text{and} \quad (1c)$$

$$X_{BT} = N_{BT}P_B P_A. \quad (1d)$$

These quantities are discussed in more detail elsewhere (Foote 2000a). What is important to consider here are the ratios

$$\frac{X_{BL}}{(X_{BL} + X_{FL})} \quad \text{and} \quad \frac{X_{FL}}{(X_{FL} + X_{FL})}.$$

Substituting from above yields

$$\begin{aligned} \frac{X_{BL}}{(X_{BL} + X_{FL})} &= P_B \{ [N_{BL}P_{D|BL} + N_{BT}P_{D|BT}(1 - P_A)] \\ &\div [N_{BL}P_{D|BL} + N_{FL}P_{D|FL}(1 - P_A) \\ &+ N_{FL}P_{D|FL} \\ &+ N_{BT}P_{D|BT}(1 - P_A)] \} \end{aligned} \quad (2a)$$

and

$$\begin{aligned} \frac{X_{Fl}}{(X_{Fl} + X_{FL})} &= P_A \{ [N_{Fl} P_{D|Fl} + N_{Bl} P_{D|Bl} (1 - P_B)] \\ &\div [N_{Fl} P_{D|Fl} + N_{Bl} P_{D|Bl} (1 - P_B) \\ &\quad + N_{FL} P_{D|FL} \\ &\quad + N_{Bl} P_{D|Bl} (1 - P_B)] \}. \end{aligned} \tag{2b}$$

Because the quantities in braces on the right-hand sides of these equations must be less than or equal to unity,  $P_B$  and  $P_A$  must be greater than or equal to the corresponding left-hand sides. Thus, the observed proportion of first or last appearances belonging to taxa known outside the interval provides a lower bound on the net probability of preservation outside the interval. How close these lower bounds are to the true values of  $P_B$  and  $P_A$  depends on the products  $pt$  and  $qt$  (Fig. 13). By substituting into these expressions, following the equations of Foote (2000a: eqs. 3–6, 26–29), it can be shown that, in the limit as  $pt$  goes to zero,  $X_{Bl}/(X_{Bl} + X_{FL}) = P_{B'}$ , and in the limit as  $qt$  goes to zero,  $X_{Fl}/(X_{Fl} + X_{FL}) = P_A$ .

Figure 14 illustrates the constraints on  $pt$ ,  $qt$ , and  $rt$  for a single interval that are implicit in  $X_{Bl}$ ,  $X_{Fl}$ ,  $X_{FL}$ , and  $X_{Bl}$ . There is assumed to be no error, so that the  $X$ -counts are equal to the ideal expectations. There are more parameters than observed numbers. We therefore cannot solve for  $p$ ,  $q$ ,  $r$ ,  $P_{B'}$ , and  $P_A$  uniquely, but we can constrain the range of possible solutions. We know that  $P_B$  must lie between  $X_{Bl}/(X_{Bl} + X_{FL})$  and unity, and  $P_A$  must lie between  $X_{Fl}/(X_{Fl} + X_{FL})$  and unity. The different solutions depicted correspond to different assumed values of  $P_B$  and  $P_A$ . If  $P_B$  takes on its minimal value, then  $pt = 0$  and  $rt$  is maximized. Similarly, if  $P_A$  takes on its minimal value, then  $qt = 0$  and  $rt$  is maximized. As  $P_B$  ( $P_A$ ) is assumed to increase to its maximal value of unity,  $pt$  ( $qt$ ) increases and  $rt$  decreases to a minimum.

Appendix 2

Modified Survivorship Equations

The goal is to develop expressions for the probability that a taxon's last appearance is in interval  $j$ , given that its first appearance is in interval  $i$ , and for the probability that its first appearance is in interval  $i$ , given that its last appearance is in interval  $j$ . Using  $\rightarrow$  and  $\leftarrow$  to indicate forward survivorship and backward survivorship (or prenaissance [Foote 2001a]), let us denote these probabilities  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$ .

Let there be  $n$  stratigraphic intervals with durations  $t_i$ . Assume that the rates of origination, extinction, and preservation per lineage-million-years (Lmy) are all constant within an interval, though they may vary among intervals. Denote these rates  $p_i$ ,  $q_i$ , and  $r_i$ . Previous applications of survivorship analysis to paleontological data have operationally placed all first appearances at an arbitrary point within the interval (Van Valen 1973; Sepkoski 1975; Raup 1978; Foote 1988) or ignored the interval of first appearance altogether (Foote 2001a), but, because of the special role played by single-interval taxa in the problem at hand, a more precise approach is required. Given only the fact that the first appearance of a lineage falls somewhere within an interval of time, we need to consider the probability distribution of precisely where within the interval this first appearance may fall, since this affects how far forward the lineage is likely to extend. Let  $z$  denote the position within the interval, from  $z = 0$  at the start to  $z = t_i$  at the end. By assumption, the rate of preservation within the interval is constant; the density  $f(z)$  of fossil occurrences within the interval, summed over all lineages, therefore follows the expected diversity trajectory. This density is equal to  $1/t_i$  if  $p_i = q_i$ ; i.e., it is uniform. If  $p_i \neq q_i$ , the density is proportional to standing diversity at time  $z$  relative to the integral of standing diversity over the whole interval:

$$f(z) = \frac{e^{(p_i - q_i)z}}{\int_0^{t_i} e^{(p_i - q_i)T} dT}, \tag{3a}$$

which is equal to

$$f(z) = \frac{(p_i - q_i)e^{(p_i - q_i)z}}{e^{(p_i - q_i)t_i} - 1}. \tag{3b}$$

Given that there is an occurrence at  $z$ , the probability that it is a first occurrence, i.e., the probability that there are no occurrences of the particular lineage before  $z$ , is given by:

$$\begin{aligned} \Pr\{\text{FO} | z\} &= \frac{p_i}{p_i + r_i} (1 - e^{-(p_i + r_i)z}) \\ &\quad + e^{-p_i z} e^{-r_i z} (1 - P_B). \end{aligned} \tag{4}$$

The first term of this expression gives the probability that the lineage originated between  $t = 0$  and  $t = z$  and was not preserved during this time, and the second term gives the probability that the lineage was already extant at  $t = 0$  and was not preserved before or after this time.

Given  $f(z)$  and  $\Pr\{\text{FO} | z\}$ , the density function reflecting the probability that an occurrence is at  $z$  and is a first occurrence is given by  $g(z, \text{FO}) = f(z)\Pr\{\text{FO} | z\}$ . Integrating over all values of  $z$ , we obtain the overall probability that a randomly chosen occurrence from interval  $i$  is the first occurrence of some lineage:

$$\Pr\{\text{FO}\} = \int_0^{t_i} f(z)\Pr\{\text{FO} | z\} dz. \tag{5}$$

From the foregoing probabilities, we obtain the conditional density that an occurrence is at  $z$ , given that it is a first occurrence:

$$h(z | \text{FO}) = g(z, \text{FO}) / \Pr\{\text{FO}\}. \tag{6}$$

Given that the first occurrence is at  $z$ , the probability that there is an occurrence sometime after interval  $i$  is given by the product of probability of survival to the end of the interval,  $e^{-q_i(t_i - z)}$ , and the probability of preservation after the interval,  $P_{A_i}$ . Therefore, the probability that the last appearance is within the same interval as the first appearance, given that the first appearance is at  $z$ , is equal to:

$$\Pr\{\text{LO} = i | z\} = 1 - P_{A_i} e^{-q_i(t_i - z)}. \tag{7}$$

The overall probability that the last appearance falls within interval  $i$ , given that the first appearance falls within interval  $i$ , is the preceding probability integrated over all values of  $z$ :

$$\begin{aligned} \Pr\{\text{LO} = i | \text{FO} = i\} &= \int_0^{t_i} h(z | \text{FO}) \Pr\{\text{LO} = i | z\} dz. \end{aligned} \tag{8}$$

This expression is evaluated by substituting into the foregoing equations. If  $p_i = q_i$ , it is equal to:

$$\begin{aligned} P_{\rightarrow ii} &= 1 - \left\{ P_{A_i} e^{-p_i t_i} \left[ \frac{r_i (e^{p_i t_i} - 1) + p_i (e^{-r_i t_i} - 1)}{r_i (p_i + r_i)} \right. \right. \\ &\quad \left. \left. - \frac{(1 - P_{B_i})(e^{-r_i t_i} - 1)}{r_i} \right] \right\} \\ &\div \left\{ \frac{p_i}{p_i + r_i} \left[ t_i + \frac{e^{-(p_i + r_i)t_i} - 1}{(p_i + r_i)} \right] \right. \\ &\quad \left. - \frac{(1 - P_{B_i})(e^{-(p_i + r_i)t_i} - 1)}{p_i + r_i} \right\}. \end{aligned} \tag{9a}$$

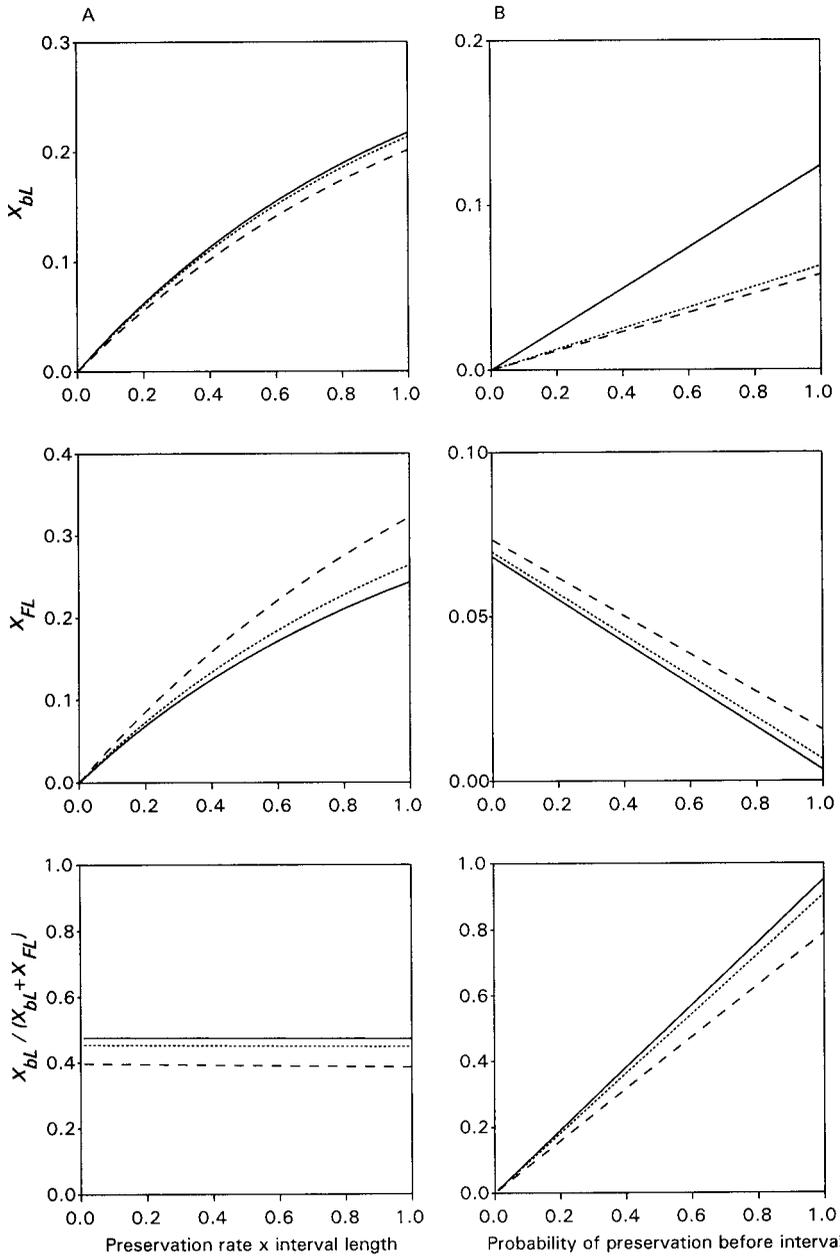


FIGURE 13. Expected number of observed taxa crossing bottom interval boundary only ( $X_{bL}$ , top panels) and expected number of observed taxa confined to the interval ( $X_{FL}$ , middle panels). For this figure  $p = q$ . Solid, dotted, and dashed curves show values corresponding to  $pt = qt = 0.1$ ,  $pt = qt = 0.2$ , and  $pt = qt = 0.5$ , respectively. A, Variation in  $rt$  with  $P_B = 0.5$  and  $P_A = 0.3$ . B, Variation in  $P_B$  with  $rt = 0.2$ . Although both categories of taxa increase with quality of preservation within the interval, they do so at nearly the same rate. The proportion of last appearances that belong to taxa known before the interval (bottom left panel) is therefore nearly independent of preservation rate within the interval. This proportion gives a lower bound on  $P_B$ , however, converging on  $P_B$  as  $pt$  approaches zero. Similar results obtain for  $X_{FL}$  and  $P_A$ .

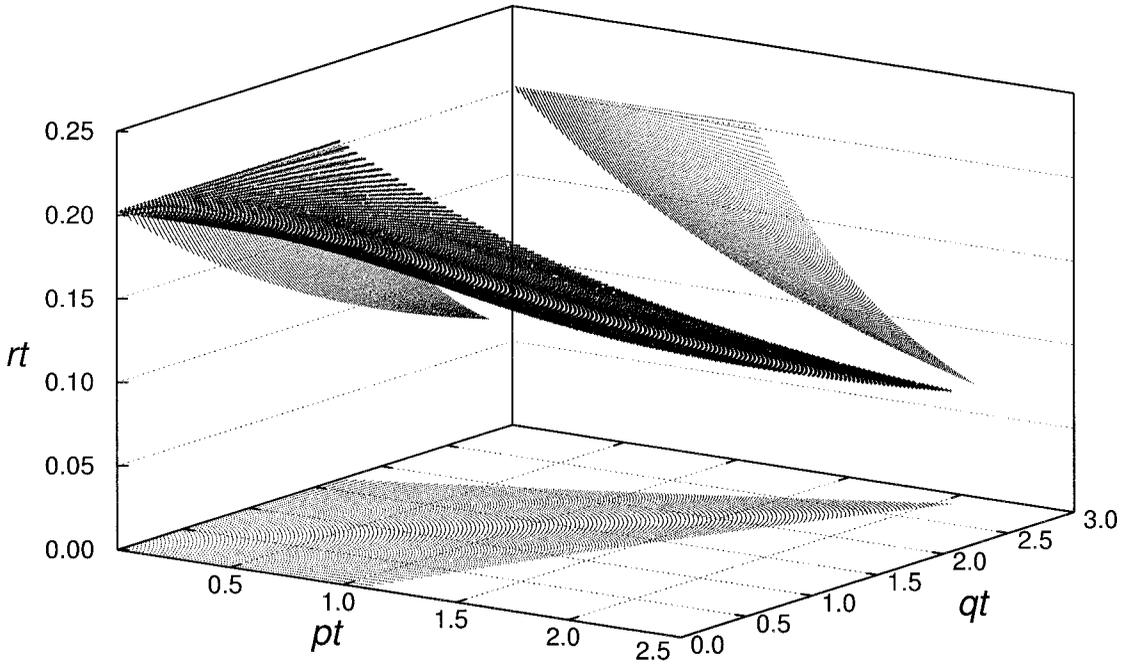


FIGURE 14. Values of  $pt$ ,  $qt$ , and  $rt$  for a single interval consistent with values of  $X_{bl}$ ,  $X_{Fl}$ ,  $X_{FL}$ , and  $X_{bl}$ . For this particular example,  $pt = 0.2$ ,  $qt = 0.15$ ,  $rt = 0.2$ ,  $P_B = 0.6$ , and  $P_A = 0.4$ ; this example illustrates the general form of the relationships among parameters.  $P_B$  and  $P_A$  are varied from their minimal values ( $X_{bl}/(X_{bl} + X_{FL}) = 0.543$  and  $X_{Fl}/(X_{Fl} + X_{FL}) = 0.372$ ) to their maximal values of unity. Higher values of  $pt$  and lower values of  $rt$  correspond to lower values of  $P_B$ . Higher values of  $rt$  and lower values of  $qt$  correspond to lower values of  $P_A$ . Lightly shaded areas show projections onto  $pt$ - $rt$ ,  $qt$ - $rt$ , and  $pt$ - $qt$  planes. See text for discussion.

If  $p_i \neq q_i$ , it is equal to:

$$P_{\rightarrow ii} = 1 - \left\{ P_{Ai} e^{-q_i t_i} \left[ \frac{r_i(e^{p_i t_i} - 1) + p_i(e^{-r_i t_i} - 1)}{r_i(p_i + r_i)} - \frac{(1 - P_{Bi})(e^{-r_i t_i} - 1)}{r_i} \right] + \left\{ \frac{p_i}{p_i + r_i} \left[ \frac{e^{(p_i - q_i)t_i} - 1}{p_i - q_i} + \frac{e^{-(q_i + r_i)t_i} - 1}{q_i + r_i} \right] - \frac{(1 - P_{Bi})(e^{-(q_i + r_i)t_i} - 1)}{q_i + r_i} \right\} \right\} \quad (9b)$$

For a lineage to have last appearance in interval  $j$ , where  $j > i$ , it must survive at least to the start of interval  $j$ ; the corresponding probability is  $e^{-\sum_{k=i+1}^j q_k t_k}$ . Either it can become extinct during interval  $j$  and be preserved during this interval (with probability  $(1 - e^{-q_j t_j})P_{D|jL}$ ), or it can survive to the end of interval  $j$ , be preserved during this interval, and not be preserved afterwards (with probability  $e^{-q_j t_j}P_{D|jL}(1 - P_{Aj})$ ). From these conditions, we can develop an expression for the probability that the last appearance will be in interval  $j$ , given that the first appearance is in interval  $i$  ( $j > i$ ) and given that it is preserved (with probability  $P_{Ai}$ ) sometime after interval  $i$ :

$$P_{\rightarrow ij} = \{(1 - P_{\rightarrow ii})e^{-\sum_{k=i+1}^j q_k t_k} [(1 - e^{-q_j t_j})P_{D|jL} + e^{-q_j t_j}P_{D|jL}(1 - P_{Aj})]\} / P_{Ai} \quad (10)$$

To obtain the probability that the first appearance is in interval  $i$ , given that the last appearance is in interval  $j$ , we follow

the same reasoning, simply switching the direction of time and interchanging origination rates with extinction rates, and  $P_B$  with  $P_A$  as necessary. For  $p_j = q_j$  and  $i = j$ , this yields:

$$P_{\leftarrow jj} = 1 - \left\{ P_{Bj} e^{-p_j t_j} \left[ \frac{r_j(e^{p_j t_j} - 1) + p_j(e^{-r_j t_j} - 1)}{r_j(p_j + r_j)} - \frac{(1 - P_{Aj})(e^{-r_j t_j} - 1)}{r_j} \right] + \left\{ \frac{p_j}{p_j + r_j} \left[ t_j + \frac{e^{-(p_j + r_j)t_j} - 1}{(p_j + r_j)} \right] - \frac{(1 - P_{Aj})(e^{-(p_j + r_j)t_j} - 1)}{p_j + r_j} \right\} \right\} \quad (11a)$$

If  $p_j \neq q_j$  and  $i = j$ , the probability is equal to

$$P_{\leftarrow jj} = 1 - \left\{ P_{Bj} e^{-p_j t_j} \left[ \frac{r_j(e^{q_j t_j} - 1) + q_j(e^{-r_j t_j} - 1)}{r_j(q_j + r_j)} - \frac{(1 - P_{Aj})(e^{-r_j t_j} - 1)}{r_j} \right] + \left\{ \frac{q_j}{q_j + r_j} \left[ \frac{e^{(q_j - p_j)t_j} - 1}{q_j - p_j} + \frac{e^{-(p_j + r_j)t_j} - 1}{p_j + r_j} \right] - \frac{(1 - P_{Aj})(e^{-(p_j + r_j)t_j} - 1)}{p_j + r_j} \right\} \right\} \quad (11b)$$

If  $i < j$ , the probability is given by

$$P_{\leftarrow ij} = \{(1 - P_{\leftarrow ij})e^{-\sum_{k=i+1}^j p_k t_k} \times [(1 - e^{-p_i t_i})P_{D|F_i} + e^{-p_i t_i}P_{D|B_i}(1 - P_{B_i})]\} \div P_{B_j} \quad (12)$$

### Appendix 3

#### Optimization Procedure

Given  $n$  stratigraphic intervals, there is a parameter space of  $3n + 2$  dimensions to be explored, corresponding to  $n$  values of  $p$ ,  $n$  values of  $q$ ,  $n$  values of  $r$ ,  $P_B$  for the first interval, and  $P_A$  for the last interval. Strictly speaking, only  $P_B$  and  $P_A$  are bounded (between 0 and 1);  $p$ ,  $q$ , and  $r$  have no theoretical upper bound. In practice, these parameters are assigned an arbitrary upper bound, generally 10 per Lmy, which is more than two orders of magnitude greater than their typical values. A candidate solution is a single point in this space. Each search through this space begins with one such point, chosen at random by picking a value for each parameter from a uniform distribution.

The optimization method of simulated annealing (Press et al. 1992: p. 444) consists of two main parts: (1) a "cooling" schedule of a number of progressively declining "temperature" levels, and (2) a protocol for varying the candidate solution within each level, evaluating the goodness-of-fit between the candidate solution and the observed data, and deciding whether to promote the candidate solution to the status of best solution.

For the problem at hand, the initial "temperature" is chosen to be of about the same order of magnitude as a typical change in goodness-of-fit with a slight perturbation of the initial candidate solution (see below), but the ultimate solution is not very sensitive to this starting condition. "Temperature" is decreased through 100 levels, each time declining by 10%. At each temperature level, the procedure cycles many times (generally 1000) through a perturbation of the current best solution. At each perturbation, a single parameter is chosen and is changed by a small, randomly chosen amount, generally between -10% and +10% of the current parameter value, to give a candidate solution. Thus, each candidate solution differs from the current best solution in only one of its parameters. (With 30 intervals,

for example, there would be 92 parameters, and so each one would be perturbed on average about 11 times at each temperature level, or 1100 times through the entire optimization procedure.) The predicted values of  $P_{\rightarrow ij}$  and  $P_{\leftarrow ij}$  corresponding to this candidate solution are tabulated, and the log deviation (described in the main text) between these probabilities and the observed  $F_{\rightarrow ij}$  and  $F_{\leftarrow ij}$  is calculated. If this log deviation is smaller than that corresponding to the best solution (i.e., if the goodness-of-fit is higher), then the candidate solution becomes the current best solution.

To avoid becoming quickly trapped on a local optimum, the simulated annealing algorithm allows movement from a better to a worse solution. A candidate solution that implies a decrease in the goodness-of-fit can become the current best solution with a probability that declines as the "temperature" declines. Specifically, if  $D_{best}$  is the log deviation for the current best solution,  $D_{cand}$  is the log deviation for the candidate solution,  $T$  is the temperature, and  $\Delta D$  is equal to  $D_{cand} - D_{best}$ , the probability of promoting the candidate solution to the status of current best solution is 1.0 if  $\Delta D$  is negative (so improvements in fit are always accepted), and this probability is otherwise equal to  $\exp(-\Delta D/T)$ . In the latter case, a uniform number on (0,1) is chosen at random; if it is less than  $\exp(-\Delta D/T)$ , the candidate solution is promoted to the current best solution. Thus, early on in the annealing schedule large decreases in goodness-of-fit are tolerated, whereas only small decreases are tolerated later. In addition to the current best solution, the solution corresponding to the lowest value of  $D$  encountered anywhere during the annealing procedure is recorded. At the end of the annealing schedule, this global best solution, which is often but not always the same as the current best solution, becomes the final solution:  $\hat{p}_1, \dots, \hat{p}_n, \hat{q}_1, \dots, \hat{q}_n, \hat{r}_1, \dots, \hat{r}_n, \hat{P}_{B1},$  and  $\hat{P}_{An}$ .

The entire annealing procedure is then repeated many times, each time with a different, randomly chosen starting solution. (In addition, when bootstrap resampling of data is used, each repetition of the annealing procedure uses a different bootstrap sample.) For each parameter, the mean and standard deviation are computed over all realizations of the annealing procedure. These means are the estimated rates discussed in the main body of this paper, and the standard deviations serve as standard errors of these estimates.