Col.Vol.Sci.Pap. ICCAT, 49 (1) : 381-386 (1999)

NOTES ON ANALYSES OF SEX-RATIO-AT-SIZE (SRS) FOR SWORDFISH

V. R. Restrepo[1]

INTRODUCTION

The SCRS is ready to undertake the task of preparing sex-specific sets of catch and catch-rate data for swordfish, which should hopefully reduce or eliminate biases in the assessments that are due to sexual dimorphism. Much work has been conducted already on what methods can be used and on what are the patterns of the sex ratio-at-size (SRS) data for this species (see review by Mejuto et al. 1997). The objective of this document is to state some concerns that could be taken into consideration when making decisions about sex-specific analyses.

AN IDEAL SITUATION

Conceptually, obtaining sex-ratio data is no different from, say, obtaining length frequencies. The accuracy of a catch-at-size data set depends on the representativeness of length-frequency distributions, i.e. the sample sizes and coverage across different strata (gear/fleet, spatio-temporal). Thus, if we had sex information for every size measurement, the decisions to be made regarding substitutions, etc., for the sex data would be identical to those made for size-specific data. Unfortunately, this is not the case for historical (early on, sex specific data were not deemed of great importance in sampling programs) and technical reasons (determining the sex of an individual is more difficult than measuring or weighting it). Being far from an ideal situation we are faced with the problem of having to make decisions about what factors (time, location, gear) are important and what factors are not. We are approaching this problem from two sides: Statistical tests to decide what data sets can be grouped or not, combined with models to obtain SRS patterns, and simply obtaining catch at age data using different procedures to see if different choices result in different outcomes.

STATISTICAL SIGNIFICANCE

Detecting differences using proportions is a difficult thing to do. The sample sizes used are important as indicated by Figure 1 (see Snedecor and Cochran 1980).

1 University of Miami, Cooperative Unit for Fisheries Education and Research, 4600 Rickenbacker Causeway, Miami, FL 33149, USA.

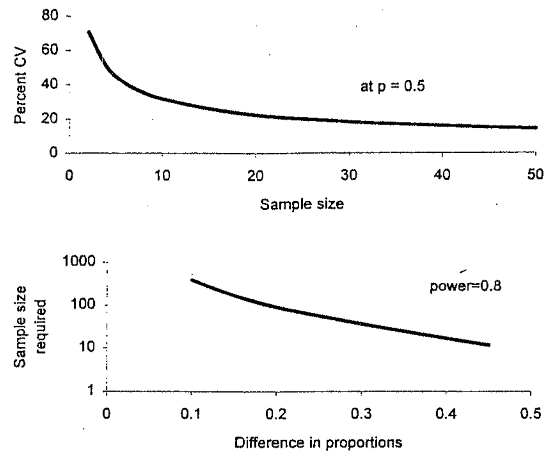

Figure 1. Top: Dependence of the percent coefficient of variation for a binomial proportion of 0.5 on sample size. Bottom: Sample size required to test for a given difference in proportion from a base of 0.5, at the 95% confidence level with a statistical power of 80%.

Thus, detecting a (two-sided) difference of 0.2 in the proportions from two independent samples requires sample sizes of about 90 fish in each sample when the proportion used as the basis for comparison equals 0.5

If we consider that the proportion of females is strongly linked to size, then detecting size-specific differences between two SRS data sets requires considerably more fish and corrections for conducting multiple tests.

MODELS

The sample size problem mentioned above is alleviated by using some type of model where size is used as a covariate or factor. Two such models that are being advocated are generalized linear models (GLiM, McCullagh and Nelder 1989) with size as a categorical variable, and generalized additive models (GAM, Hastie and Tibshirani 1990) with size as a continuous variable. In my view, both are useful approaches and both should be pursued, but it seems worthwhile to highlight some of the similarities and differences between the two.

Both GLiMs and GAMs can treat the proportions as being binomially distributed and weight them according to sample size. By default, the proportions (p) are subjected to a logit transformation, $\text{logit}(p) = \log\{p/(1-p)\}$. The logit transformation has the convenience that it has a direct interpretation as the logarithm of the odds in favor of a success (in this case the success means being a female swordfish).

When GLiMs are used with size as a factor, they lead to tests that do not depend on a functional relationship between p and size. (Note that one could also use GLiMs with size as a continuous variable and assume, for instance, a linear or cubic relationship between p and size). GAMs, on the other hand, fit a nonlinear relationship between p and size, whose curvature depends on the degrees of freedom specified for fitting. Thus, a GLiM will use up one parameter per size category while a GAM will use about 3 or 4 (or more, depending on user specifications) regardless of the number of sizes in the data set. This difference leaves more degrees of freedom for GAMs to test for second and third-order interactions, depending on the amount of data available.

At the same time, the predictions from a GAM will generally follow a smooth pattern over size and may therefore miss abrupt changes in the female proportions over contiguous sizes that the GliMs will not miss. However, in many cases where sample sizes are large, female proportions seem to change more or less smoothly over size such that both types of models will give similar results (e.g. Figure 2) with GAMs generally using up fewer degrees of freedom.

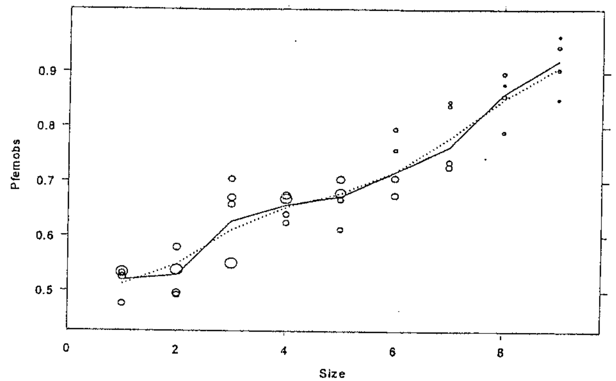

Figure 2: Female proportions of swordfish at size (size 1 = 120 cm) for Spanish longlines in the feeding area, 1993. Circles: observations by quarter with size of symbols proportional to sample size. Solid line: GLiM prediction (d.f. = 8, residual deviance = 47.2). Dotted line: GAM prediction (d.f. = 3, residual deviance = 51.7)

VISUAL DATA EXAMINATION

Before using GAMs or GliMs to test for significant factors, it is very useful to conduct visual examination of the data. This has been usually done in the past by simply plotting the proportions at size for different strata. However, because of the sample size considerations mentioned earlier on, it is desirable to also provide some idea of uncertainty in any visual comparisons. One possibility is to compute approximate confidence limits for the proportions as (Collet 1991):

$P(\text{lower}) = f \,/\, \{f+(n-f+1)\, F_{2(n-f+1),2f}(\alpha/2)\}$, and

$P(\text{upper}) = (f+1) \,/\, \{f+1+(n-f) \,/\, F_{2(f+1),2(n-f)}(\alpha/2)\}$

where

f = number of females,

n = total number of fish

$F_{v1,v2}(\alpha/2)$ = upper $(100\alpha/2)\%$ point of the F-distribution with $v1$ and $v2$ degrees of freedom.

As an example, the data from Japanese longlines in 1995, transition area, are plotted in Figure 3. Note the wide confidence intervals (95% level), even though the average number of fish sampled per size class was greater than 40.

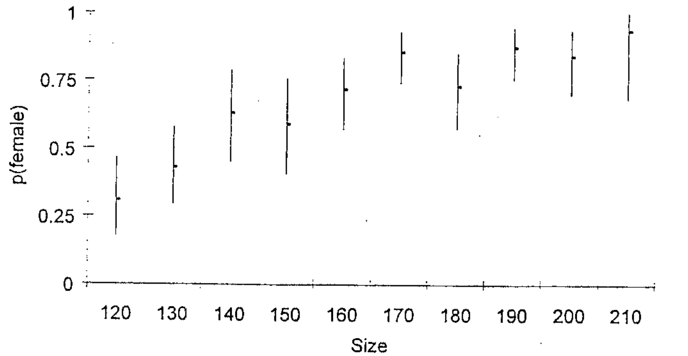

Figure 3. Approximate 95% confidence intervals for female proportion at size data from Japanese longlines in the transition area (1995).

Another important aspect to visualize before jumping into model-based analyses is to determine how balanced the data are in terms of representativeness in different strata. This type of tabulation is routinely done by the SCRS in standardization of catch rates and should be performed here as well.

THE IMPORTANCE OF SEASON

It has been mentioned by other authors (see Mejuto et al. 1997) that seasonal effects appear to be negligible compared to those of other factors. However, as sex ratios are theoretically affected by many factors, one should expect to see some seasonal differences. To illustrate this point, I examined the Spanish longline data in the feeding area, which is the subset of available data with the most observations. I limited the analyses to sizes (10 cm intervals) and quarters for which there were at least 20 observations and ignored sizes greater than 200 cm LJFL because the data were sparse. The following table summarizes the results of the highest-order model with significant results that was examined:

	D.F. (approx.)	Residual deviance	Explained deviance	Percent Explained	Prob(Chi2)
NULL		3124.6			
s(Size)	3	861.1	2263.4	72.4	0.0E+00
Year	7	773.4	87.7	2.8	3.6E-16
Quarter	3	648.3	125.2	4.0	5.9E-27
s(Size):Year	20	594.9	53.4	1.7	7.2E-05
Year:Quarter	20	455.2	139.7	4.5	5.8E-20
s(Size):Year:Quarter	62	323.1	132.1	4.2	5.5E-07
Total:	115	323.1	2801.5	89.7	

Here, s(Size) means a smooth nonlinear function of size, and it explains over 70% of the variation in the data. However, other terms, including interactions, were significant even though they explained a relatively small proportion of the residual deviance. Model diagnostics and partial fits of the various main factors (in logit space) are shown in Figure 4. These plots suggest a good model fit to the data and there is no compelling reason to return to a simpler model with fewer terms.

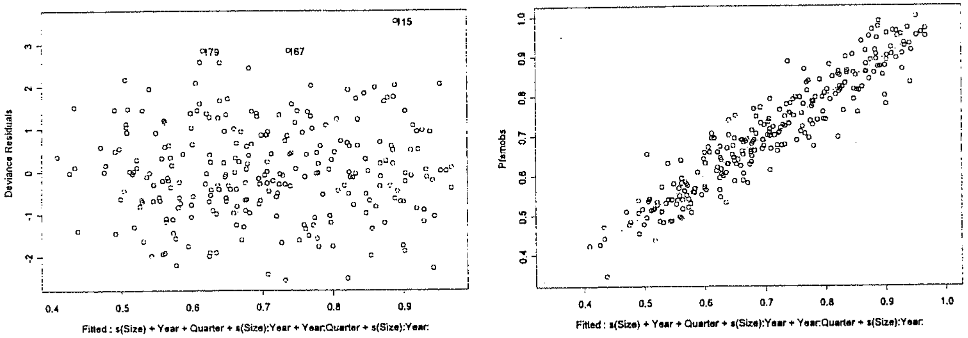

Figure 4. Diagnostics and partial fit plots for the generalized additive model fit to the [Spain-Longline-Feeding] data set using Size, Year and Quarter.

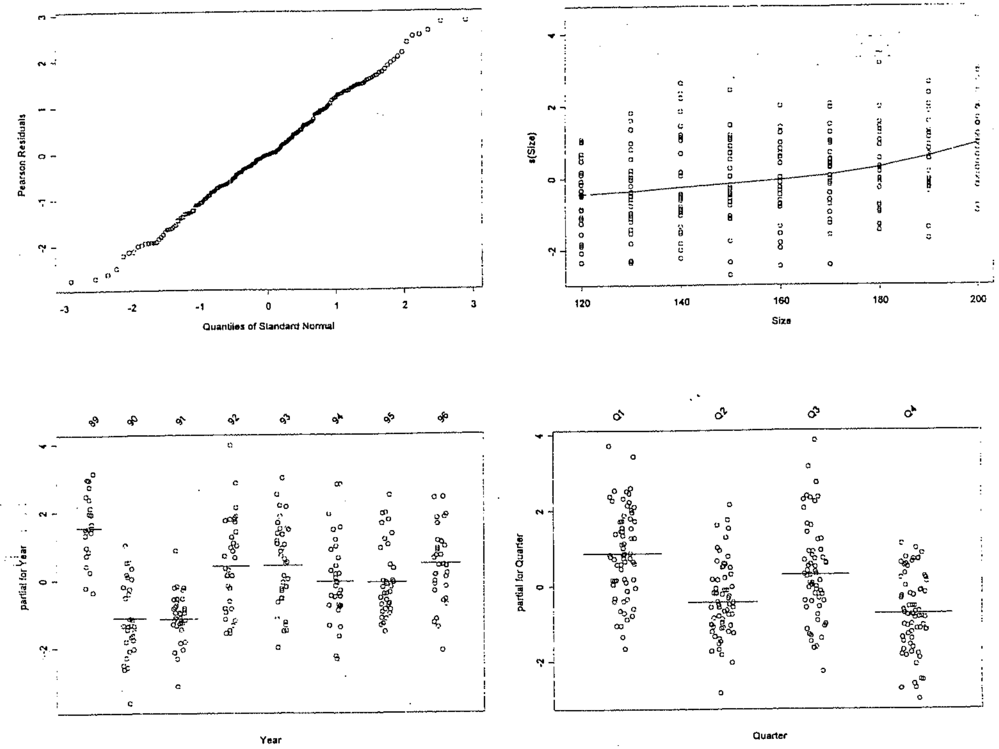

Figure 4 (cont.)

Despite the relatively small variation attributed to seasonal (year, quarter) effects, predictions from the model suggest that these may have an impact on the estimation of catch-at-size by sex (Figure 5), although I did not investigate this possibility any further.

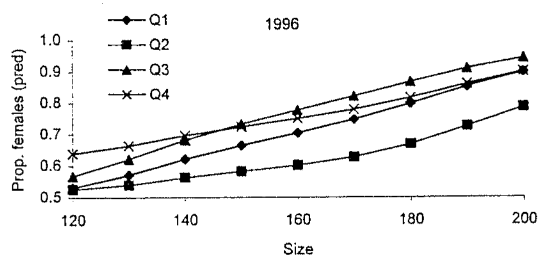

Figure 5. Model predictions for 1996 by quarter using a GAM (data: Spain Longline Feeding).

Many of the currently-available sex-specific data sets have insufficient sample sizes to detect seasonal patterns, but that is not to say that they do not exist. Figure 6 shows the model's predictions by year and quarter, with 95% confidence limits for the observed data.

THE FUTURE

The comments given above suggest that our ability to detect important features in the data is limited by the number of observations that are available. In this context, it is important to also think about the situation in the future in conjunction with sampling programs, which are not expected to intensify with respect to sex determination (if anything, they may stay level or scale down in some regions). For those data subsets that have a large number of observations, it is apparent that the proportion of females at size varies over fine scales (perhaps monthly, by 5-degree square, and by gear and/or depth). We could carry out very detailed analyses for those specific data sets including the fine-scale factors as explanatory variables and continue to use their predictions in some fashion to fill in gaps in the future. Alternatively, we may want very simple models with very few explanatory variables to make the predictions that will be used to fill in the gaps. The question is, which approach is better? Perhaps we could gain some insight by simulations. However, I would argue that, no matter which approach we follow, the safest thing to do is to collect more data intensively every so many years (3 or 5?) to check that we are on the right track.

As mentioned in the introduction, the current plan is to create catch-at-size (and catch-at-age) data sets starting from different points, and see if they differ in practical terms. This is in principle a good idea, particularly in creating a sense of comfort if two or more approaches give essentially the same "bottom-line" result. But what if they differ? And, what criteria – statistical or otherwise - will we use to determine if they differ? Ultimately, we may want to look at the effects that different catch-at-age data sets have on the results of age-structured assessments. My recommendation is that by thinking of such practical tests for the future, we don't loose sight of the need to continue to collect the basic data which are needed.

REFERENCES

Collet, D. Modelling Binary Data. 1991. Chapman and Hall, NY.

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized Additive Models. Chapman and Hall, London.

McCullagh, P., and J. A. Nelder. 1989. Generalized Linear Models (2nd ed.). Chapman and Hall, London.

Mejuto, J., J. M. De La Serna, and B. García. 1997 (*in press*). Some considerations on the spatial and temporal variability in the sex ratio at size of the swordfish, *Xiphias gladius*. ICCAT SCRS/97/32.

Snedecor, G. W., and W. G. Cochran. 1980. Statistical Methods (7th ed.). Iowa State Univ. Press, Iowa.

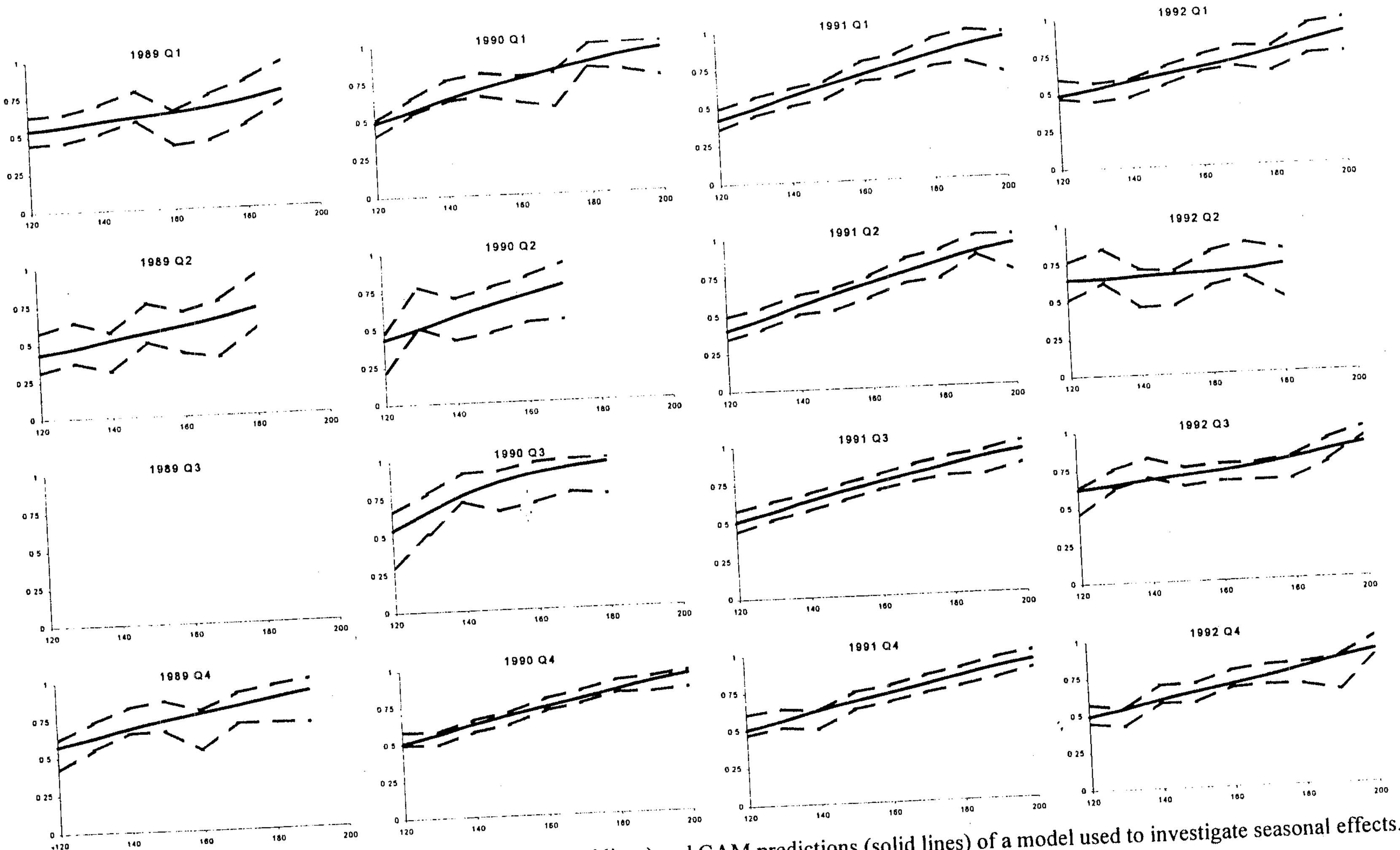

Figure 6. Confidence limits (95%) for the observations (dashed lines) and GAM predictions (solid lines) of a model used to investigate seasonal effects.

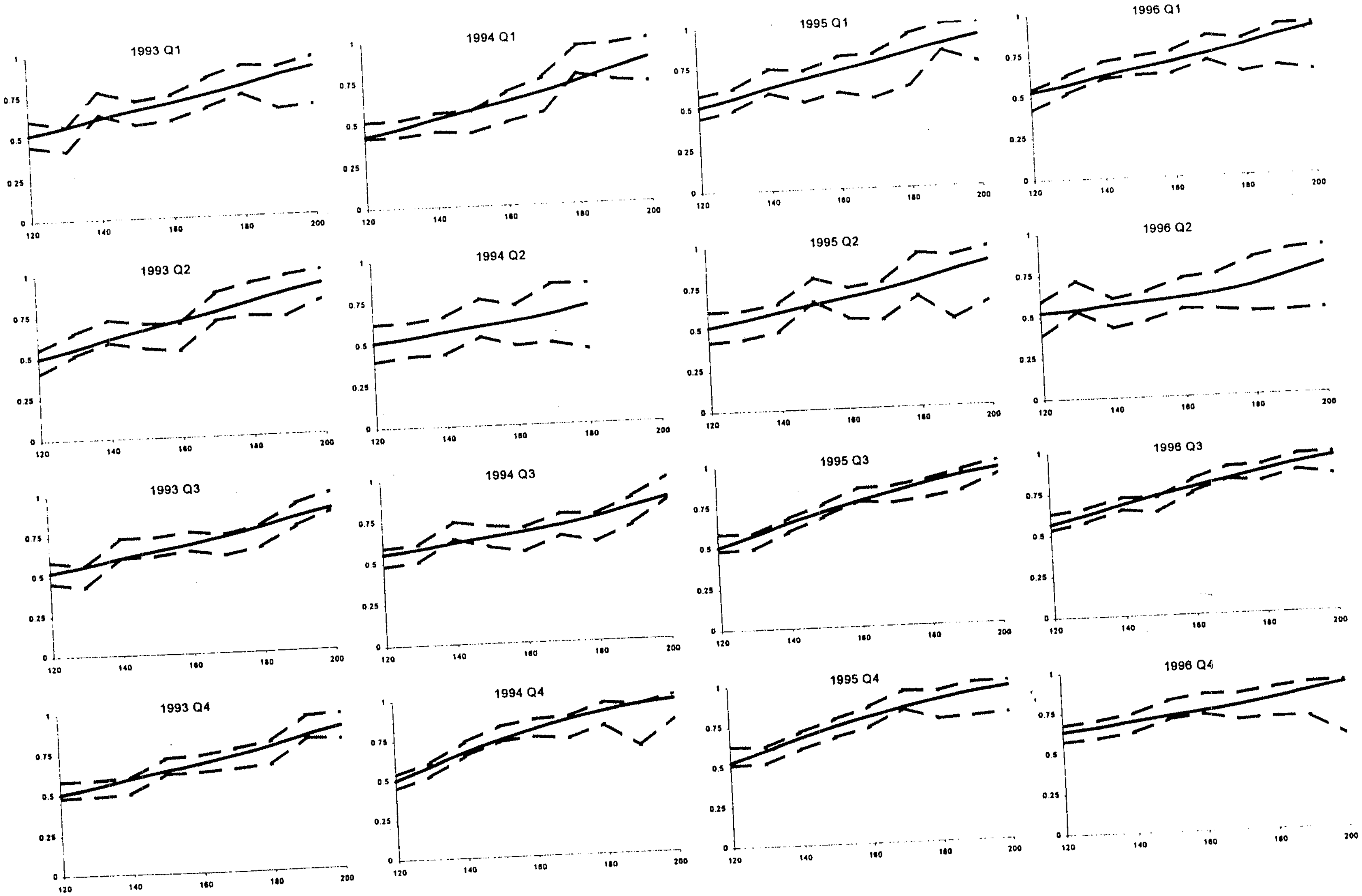

Figure 6 (cont.)