Spell durations and the impact of censoring

Michael A. Nolan

1 Department of Economics, University of St Andrews, Fife, KY16 9AL, Scotland, UK. (E-mail: man1@st-and.ac.uk)

First version received: September 1996/Final version received: June 2000

Abstract. A grouped hazard approach for analysing multiple-spell durations subject to censoring is applied to spells of absence from the workplace. We follow Barmby, Orme and Treble’s (1991) procedure for dealing with unobserved heterogeneity, but argue that their treatment of the observed discrete data, and the inherent censoring, is inappropriate and could lead to significant overestimation of duration dependence.

Key words: spell durations, censoring, Monte Carlo analysis

JEL classifications: C15, C41, J29

1. Introduction and background

Duration analysis is applied to several areas of economics – notably to data on unemployment spells, employment spells, strikes; and to spells of attendance at, and absence from, the workplace. The motivation for this paper is the duration analysis of absence from the workplace undertaken by Barmby, Orme and Treble (1991). Our approach and theirs have features in common; namely, a Weibull specification for the log-hazard, a view of weekends as acting as an independent censoring mechanism, an assumption that middle-censored spells are continuous and the application of corrections for unobserved individual-specific heterogeneity. However, while the earlier authors viewed the censoring impact of weekends as symmetric, we argue that inde-

Acknowledgements. This paper would not have been possible without the financial support of the Economic and Social Research Council. A particular debt of gratitude is owed to Chris Orme, and thanks are also due to John Hutton, Peter Smith, John Treble, Tim Barmby and three anonymous referees.
ependent left and right censoring cannot be regarded as similar phenomena. Moreover, we recognise the inherent grouped nature of the data by employing a grouped hazard specification and this contrasts with Barmby, Orme and Treble (1991), who presume continuous absence durations. To illustrate the differences in the two approaches, we report the results of some empirical analysis in which multiple-spell data are employed. The conclusions drawn are further supported by a Monte Carlo study which is summarised in Appendix 3.

Our specification is outlined in Section 2 and the resulting log-likelihood is constructed in Section 3. Two alternative approaches to the problem of unobserved heterogeneity are briefly summarised in Section 4. In Section 5, we outline the application of our specification to data concerning absence from the workplace. Empirical results are discussed in Section 6. Finally, Section 7 concludes.

2. Methodological issues

Barmby, Orme and Treble (1991) use an empirical setting where weekends have a censoring impact on absence duration. This is a form of independent censoring, because the timing of weekends is fixed – and thus independent of the underlying random variable from which durations are drawn. To illustrate why this is so, consider a simple hypothetical example within which there is an incidence of an underlying process of ρ spells per day, with weekends acting as an independent censoring mechanism. The assumption that ρ is fixed across the days of the week is moderate and affords tractability. We also assume independence between the onset of the process (sickness in our example) and the subsequent duration of that spell.

Of the 7ρ underlying spells yielded by the process in one week, all of the 5ρ which commence on a weekday result in an observed spell (the 4ρ spells starting on Tuesday to Friday do not concern us, except to note (for example) that an observed spell of t days beginning on Tuesday is underlain by a process duration, s, where (t − 1) < s ≤ t). However, a proportion (1 − ω2) of the ρ underlying spells which commence on the Sunday are so short that no corresponding spell is observed. Additionally, a larger proportion (1 − ω3) of the ρ underlying spells commencing on the Saturday do not continue until at least Monday – and, again, no spell is observed in these instances. Only [ρ(1 + ω2 + ω3)] underlying spells out of the 3ρ potential maximum actually result in an independently left-censored spell being observed. Consequently, of the spells observed to start on a Monday, fewest are underlain by a commencement of the underlying process on the previous Saturday; a greater number have a Sunday underlying start and still more begin on the Monday itself. For such spells, observed to commence on a Monday, the proportion which actually began on a Monday is therefore 1/(1 + ω2 + ω3) = π1. Similarly, the proportions which actually started on either Sunday or Saturday are ω2/(1 + ω2 + ω3) = π2 and ω3/(1 + ω2 + ω3) = π3, respectively.

Let us now assume that the onset of the underlying process is uniform across days of the week for a given individual (but may vary across individuals), and that the proportion of survivors depends upon individual characteristics. Consider the situation of a 5-day working week (Monday-Friday), where absences are only recorded in whole days, although the underlying