Journal of Econometrics 187 (2015) 275–292

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Diagnostic analysis and computational strategies for estimating discrete time duration models—A Monte Carlo study

Xianghong Li, Barry Smith *

York University, Canada

ARTICLE INFO

Article history:
Received 2 July 2009
Received in revised form
10 December 2014
Accepted 7 February 2015
Available online 12 March 2015

JEL classification:
C41
C14
C15

Keywords:
Discrete time duration model
Monte Carlo
Simulated annealing
Duration dependence
Unobserved heterogeneity

ABSTRACT

This paper uses Monte Carlo analysis to study important and contentious issues in estimating single-spell discrete time duration models. We find simulated annealing dominates gradient methods for recovering true models. We recommend a partially flexible step function for duration dependence combined with likelihood ratio tests for determining support points of unobserved heterogeneity. We find that ignoring time-changing features of explanatory variables introduces substantial biases in model coefficient and average partial effect estimates. These biases do not diminish as sample size increases.

1. Introduction

The estimation of single-spell duration models with unobserved heterogeneity has been an active area of research since Lancaster (1979), Nickell (1979), and Heckman and Flinn (1982). For many reasons, some of which are related to the complexity and numerical difficulty of the empirical models, several crucial issues remain unresolved. Researchers are primarily interested in estimating three objects: (i) duration dependence, (ii) coefficients of model covariates and, (iii) average partial effects (APE hereafter) of explanatory variable changes upon expected duration.

The duration dependence process is of interest for two reasons. First, it yields insights into the equilibrium of search behavior. Second, separate identification of duration dependence allows researchers to distinguish state dependence from unobserved heterogeneity (Heckman, 1981). In a policy setting, for example, the presence of substantial negative duration dependence, after accounting for unobserved heterogeneity, suggests that policy interventions should be made sooner rather than later in a spell of unemployment.

The coefficients of explanatory variables can have a structural interpretation and thus are also of interest in many policy contexts. For example, researchers may assess the roles played by increased unemployment duration and increased inflows to unemployment in generating cyclical unemployment. Finally, APE's have the advantage that they can be compared across studies even when they differ with respect to hazard function specification and time interval assumed (for example, monthly versus quarterly data). As such, researchers can summarize and compare empirical results for policy makers.

The literature, however, contains several unresolved questions and areas of contention. One unresolved question involves choosing the best numerical approach to optimizing the likelihood functions of the models. When a model incorporates the Heckman and Singer (1984) discrete specification of unobserved heterogeneity, the likelihood becomes a mixture of discrete distributions. Such models are known to present numerical optimization challenges because they may contain many local optima and flat regions. Will gradient-based search techniques with possible derivative enhancements (see, for example, Baker and Melino (2000, BM hereafter)) find the global optimum, or should the researcher use or combine other methods such as simulated annealing that are designed to search for a global optimum? The evidence in the literature appears to be mixed.

* Correspondence to: Department of Economics, York University, 4700 Keele Street, Vari Hall 1075, Toronto, ON M3J 1P3, Canada. Tel.: +1 416 736 2100; fax: +1 416 736 5987.

E-mail address: jbsmith@yorku.ca (B. Smith).

http://dx.doi.org/10.1016/j.jeconom.2015.02.024
0304-4076/

One area of contention involves how to model duration dependence. The very early studies used quite restrictive forms of duration dependence such as that implied by a Weibull hazard function. Others combined a flexible polynomial chosen by the data with a Heckman–Singer specification of unobserved heterogeneity (see, for example, Ham and Rea (1987), McCall (1996), and Ham et al. (1998)). Alternatively, Meyer (1990) suggested using a very flexible step function with different steps for each week of duration, but he could only estimate his model by assuming a restrictive form of unobserved heterogeneity. The polynomial versus step function issue remains an open question with potentially very serious estimation consequences.

Another area of contention involves model selection relating to the unobserved heterogeneity. The mixture nature of models with the Heckman–Singer specification means that special consideration should be given to the problem of selecting the number of mixture components.[1] Some researchers have chosen criteria that penalize overfitting. For example, Ham et al. (1998) and Baker and Melino (2000) suggest using more conservative criteria such as the Bayesian Information Criterion (BIC) and the Hannan–Quinn Information Criterion (HQIC), since they noted a tendency for their models to overestimate the contribution of unobserved heterogeneity and a corresponding instability in parameter estimates. On the other hand, Gaure et al. (2007) and Mroz and Zayats (2008) advocate more liberal criteria in adding additional points of support such as the Akaike Information Criterion (AIC) or the extremely aggressive approach of adding a point of support (equivalent to two parameters) whenever the increase in the log-likelihood exceeds 0.01. It is also not unusual for researchers to comment upon how difficult it is to estimate more than two support points thus apparently obviating some of the inference and selection issues. Further, several studies allow for time-changing covariates while many others keep covariates fixed over the spell. Finally, some researchers have focused upon coefficient parameter estimates while others stress APE's and some look at both.

In this paper, we use Monte Carlo (MC) analysis to study, understand, and in many cases, resolve these issues. First, we find that simulated annealing (SA hereafter) performs better, and usually significantly better, than gradient methods in all of the models we estimate. Our experience indicates that it is essential to use SA to estimate all model parameters simultaneously as opposed to combining SA with other gradient methods (see, for example Gaure et al., 2007). Second, we find that a step function is the best way of specifying the unknown form of duration dependence. Although numerical problems were often observed when gradient methods were used to estimate models combining a step function specification of duration dependence with a Heckman–Singer specification of heterogeneity, in almost all cases, SA converges to a proper optimum in such models. Moreover, and in the spirit of variable bandwidth in local regression (for example, Fan and Gijbels (1992)), we find that a partially flexible step function specification, in which each step reflects at least 3% of observed transitions behaves very well with all sample sizes. On the other hand, fixed polynomial specifications of duration dependence, in general, are not sufficiently flexible. We show that a fixed third-order polynomial specification is the true source of the biases reported by BM.

Third, in terms of unobserved heterogeneity specification, we find that using BIC often leads to too few support points, and that relying on AIC, or the even more aggressive criterion of log likelihood improvement exceeding 0.01 usually results in too many support points. Further, we find that once we use an appropriate step function to model duration dependence, a likelihood ratio test (with two degrees of freedom) is the best approach for determining the unobserved heterogeneity distribution.

Finally, we find that ignoring any time-changing feature of explanatory variables leads to major consequences beyond substantial bias in the coefficient estimates for these variables. While APE estimates appear to be more robust to misspecification than estimates of duration dependence or coefficients on the explanatory variables, substantial biases can arise in all parameter estimates including APE when a time-changing variable is fixed at its observed value at the beginning of the spell. These biases do not diminish as sample size increases. This finding is particularly important because time-changing variables enter almost all duration models.

The remainder of this paper is organized as follows: Section 2 reviews computational strategies used to estimate duration models. Section 3 describes the discrete time duration models and data generation processes (DGP's) investigated in this study. Section 4 provides MC evidence on our base case, the main DGP used by BM. In Section 5, we explore three models with added complexity to our base case in unobserved heterogeneity. Section 6 introduces a DGP with both a randomly assigned treatment and a time-changing variable. This DGP captures most of the complexity of single-spell models appearing in applied work. In Section 7, we combine the evidence from the full spectrum of experiments, comment on a few important issues, and offer our conclusions.

2. Computational strategies

We consider single-spell duration models that use the discrete Heckman–Singer specification (1984) of unobserved heterogeneity. Most researchers use Newton-type derivative-based algorithms to optimize numerical (log) likelihood functions. For many nonlinear problems, these algorithms are relatively easy to implement and often converge quickly and reliably. However, the semiparametric mixture models resulting from the Heckman–Singer specification of unobserved heterogeneity usually have multiple local optima, and these models can behave in complicated ways at the boundaries of the parameter space. Derivative-based algorithms provide few safeguards against terminating at inferior local optima. Further, restarting with different initial parameter values often does not eliminate the problem.

Most Monte Carlo studies have explored alternative optimization algorithms as a means to improving the chance of converging to a global optimum.[2] Additional algorithms that have been used to estimate mixture models, in general, and duration models with the Heckman–Singer specification, in particular, include: Expectation Maximization (EM, as in Dempster et al. (1977)), Gateaux (directional) derivatives (Lindsay, 1983), and Simulated Annealing (SA, e.g. Kirkpatrick et al., 1983). We evaluated the performance of these three algorithms as well as three commonly used derivative-based algorithms: Sequential Quadratic Programming (SQP) by Klaus Schittkowski[3] and the DFP and GRADX routines in GQOPT (Goldfeld and Quandt).[4] Overall, we find that derivative-based algorithms often terminate suboptimally at a boundary of the parameter space even when the sample is large. The EM algorithm is very sensitive to starting values and converges slowly. In addition, we find that the Gateaux derivatives were not helpful in maximizing the discrete time duration models in this study.

1 As we note below, this is a non-standard testing situation since in such a test one parameter will not be identified under the null hypothesis of no additional support points.

2 There is a broad literature containing theoretical results for estimating duration models and many empirical studies in which this methodology is applied. However, there are few Monte Carlo studies in the literature. Besides the three recent studies discussed in the introduction, BM, Gaure et al. (2007), and Mroz and Zayats (2008), there are only a few earlier Monte Carlo studies, and they use relatively small samples (for example, Heckman and Singer (1984), Ridder (1987), and Huh and Sickles (1994)).

3 http://www.ai7.uni-bayreuth.de/software.html.

4 http://www.quandt.com/req.html.