Journal of Econometrics 205 (2018) 177–203

Contents lists available at ScienceDirect

Journal of Econometrics

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

Generalized indirect inference for discrete choice models[☆,☆☆]

Marianne Bruins[a,*], James A. Duffy[b], Michael P. Keane[c], Anthony A. Smith Jr.[d]

[a] *Nuffield College and Department of Economics, University of Oxford, United Kingdom*
[b] *Corpus Christi College and Department of Economics, University of Oxford, United Kingdom*
[c] *School of Economics, UNSW Business School, Australia*
[d] *Department of Economics, Yale University and National Bureau of Economic Research, United States*

ARTICLE INFO

Article history:
Available online 7 April 2018

ABSTRACT

This paper develops and implements a practical simulation-based method for estimating dynamic discrete choice models. The method, which can accommodate lagged dependent variables, serially correlated errors, unobserved variables, and many alternatives, builds on the ideas of indirect inference. The main difficulty in implementing indirect inference in discrete choice models is that the objective surface is a step function, rendering gradient-based optimization methods useless. To overcome this obstacle, this paper shows how to smooth the objective surface. The key idea is to use a smoothed function of the latent utilities as the dependent variable in the auxiliary model. As the smoothing parameter goes to zero, this function delivers the discrete choice implied by the latent utilities, thereby guaranteeing consistency. We establish conditions on the smoothing such that our estimator enjoys the same limiting distribution as the indirect inference estimator, while at the same time ensuring that the smoothing facilitates the convergence of gradient-based optimization methods. A set of Monte Carlo experiments shows that the method is fast, robust, and nearly as efficient as maximum likelihood when the auxiliary model is sufficiently rich.

1. Introduction

Many economic models have the features that (i) given knowledge of the model parameters, it is easy to simulate data from the model, but (ii) estimation of the model parameters is extremely difficult. Models with discrete outcomes or mixed discrete/continuous outcomes commonly fall into this category. A good example is the multinomial probit (MNP), in which an agent chooses from among several discrete alternatives the one with the highest utility. Simulation of data from the model is trivial: simply draw utilities for each alternative, and assign to each agent the alternative that gives them the greatest utility. But estimation of the MNP, via either maximum likelihood (ML) or the method of moments (MOM), is quite difficult.

☆ The authors thank Debopam Battacharya, Martin Browning and Liang Chen for helpful comments; and especially Tim Neal who provided superb research assistance. Keane's work on this project has been funded by the Australian Research Council under grant FL110100247. The manuscript was prepared with LyX 2.1.4 and JabRef 2.7b.

☆☆ An earlier version of this paper was circulated as the unpublished manuscript Keane and Smith (2003). That paper proposed the method of generalized indirect inference (GII), but did not formally analyze its asymptotic or computational properties. The present work, under the same title but with two additional authors (Bruins and Duffy), rigorously establishes the asymptotic and computational properties of GII. It is thus intended to subsume the 2003 manuscript. Notably, the availability of the 2003 manuscript allowed GII to be used in numerous applied studies (see Section 3.3), even though the statistical foundations of the method had not been firmly established. The present paper provides these foundations and fills this gap in the literature.

* Corresponding author.
E-mail address: marianne.bruins@economics.ox.ac.uk (M. Bruins).

https://doi.org/10.1016/j.jeconom.2018.03.010
0304-4076/

The source of the difficulty in estimating the MNP, as with many other discrete choice models, is that, from the perspective of the econometrician, the probability an agent chooses a particular alternative is a high-dimensional integral over multiple stochastic terms (unobserved by the econometrician) that affect the utilities the agent assigns to each alternative. These probability expressions must be evaluated many times in order to estimate the model by ML or MOM. For many years econometricians worked on developing fast simulation methods to evaluate choice probabilities in discrete choice models (see Lerman and Manski, 1981). It was only with the development of fast and accurate smooth probability simulators that ML or MOM-based estimation in these models became practical (see McFadden, 1989, and Keane, 1994).

A different approach to inference in discrete choice models is the method of "indirect inference". This approach (see Smith, 1990, 1993; Gourieroux et al., 1993; Gallant and Tauchen, 1996), circumvents the need to construct the choice probabilities implied by the economic model, because it is not based on the likelihood, or on moments based on choice frequencies. Rather, the idea of indirect inference (II) is to choose a statistical model that provides a rich description of the patterns in the data. This descriptive model is estimated on both the actual observed data and on simulated data from the economic model. Letting β denote the vector of parameters of the structural economic model, the II estimator is that $\hat{\beta}$ which makes the simulated data "look like" the actual data — in the sense (defined formally below) that the descriptive statistical model estimated on the simulated data "looks like" that same model estimated on the actual data. (The method of moments is thus a special case of II, in which the descriptive statistical model corresponds to a vector of moments.)

Indirect inference holds out the promise that it should be practical to estimate any economic model from which it is practical to simulate data, even if construction of the likelihood or population moments implied by the model is very difficult or impossible. But this promise has not been fully realized because of limitations in the II procedure itself. It is very difficult to apply II to models that include discrete (or mixed discrete/continuous) outcomes for the following reason: small changes in the structural parameters of such models will, in general, cause the data simulated from the model to change discretely. Such a discrete change causes the parameters of a descriptive model fit to the simulated data to jump discretely, and these discontinuities are inherited by the criterion function minimized by the II estimator.

Thus, given discrete (or discrete/continuous) outcomes, the II estimator cannot be implemented using gradient-based optimization methods. One instead faces the difficult computational task of optimizing a multidimensional step function using much slower derivative-free methods. This is very time-consuming and puts severe constraints on the size of the structural models that can be feasibly estimated. Furthermore, even if estimates can be obtained, one does not have derivatives available for calculating standard errors.

In this paper we propose a "generalized indirect inference" (GII) procedure to address this important problem (Section 3). The key idea is to generalize the original II method by applying two different descriptive statistical models to the simulated and actual data. As long as the two descriptive models share the same vector of pseudo-true parameter values (at least asymptotically), the GII estimator based on minimizing the distance between the two models is consistent, and will enjoy the same asymptotic distribution as the II estimator.

While the GII idea has wider applicability, here we focus on how it can be used to resolve the problem of non-smooth objective functions of II estimators in the case of discrete choice models. Specifically, the model we apply to the simulated data does not fit the discrete outcomes in that data. Rather, it fits a "smoothed" version of the simulated data, in which discrete choice indicators are replaced by smooth functions of the underlying continuous latent variables that determine the model's discrete outcomes. In contrast, the model we apply to the actual data is fit to observed discrete choices (obviously, the underlying latent variables that generate actual agents' observed choices are not seen by the econometrician).

As the latent variables that enter the descriptive model applied to the simulated data are smooth functions of the model parameters, the non-smooth objective function problem is obviously resolved. However, it remains to show that the GII estimator based on minimizing the distance between these two models is consistent and asymptotically normal. We show that, under certain conditions on the parameter regulating the smoothing, the GII estimator has the same limiting distribution as the II estimator, permitting inferences to be drawn in the usual manner (Section 4). Our theoretical analysis goes well beyond merely deriving the limiting distribution of the minimizer of the GII criterion function. Rather, in keeping with computational motivation of this paper, we show how the proposed smoothing facilitates the convergence of standard derivative-based optimizers, providing results for selected line-search and trust-region methods.

Finally, we conduct a set of Monte Carlo experiments to assess the performance of the GII estimator, in terms of bias, efficiency, and computation time, for a range of example models (Section 5). For models of only moderate complexity (i.e. on the order of 10 parameters), GII significantly outperforms conventional II (computed using the downhill simplex), in terms of both computation time and efficiency. We look at some cases where simulated maximum likelihood (SML) is also feasible, and show that efficiency losses relative to SML are small. We also show how judicious choice of the descriptive (or auxiliary) model is very important for the efficiency of the estimator. This is true not only here, but for II more generally.

Proofs of all theoretical results stated in the paper are given in the Supplementary Material.

2. The model

We first describe a class of (dynamic) discrete choice models, which motivate the estimation method developed in this paper. However, the ideas underlying the method could be applied to almost any conceivable model involving discrete outcomes, including models with mixed discrete/continuous outcomes (such as Model 5), and even models in which individuals' choices solve forward-looking dynamic programming problems.