Journal of Econometrics 185 (2015) 162–181

Contents lists available at ScienceDirect

Journal of Econometrics

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

Asymptotic theory for differentiated products demand models with many markets[☆]

Joachim Freyberger

Department of Economics, University of Wisconsin, Madison, United States

ARTICLE INFO

Article history:
Received 16 April 2012
Received in revised form
12 October 2014
Accepted 20 October 2014
Available online 11 November 2014

JEL classification:
C13
C15
L13

Keywords:
Demand estimation
Differentiated products
Many markets
Bias correction
Adjusted standard errors

ABSTRACT

This paper develops asymptotic theory for differentiated product demand models with a large number of markets T. It takes into account that the predicted market shares are approximated by Monte Carlo integration with R draws and that the observed market shares are approximated from a sample of N consumers. The estimated parameters are $\sqrt{T}$ consistent and asymptotically normal as long as R and N grow fast enough relative to T. Both approximations yield additional bias and variance terms in the asymptotic expansion. I propose a bias corrected estimator and a variance adjustment that takes the leading terms into account. Monte Carlo simulations show that these adjustments should be used in applications to avoid severe undercoverage caused by the approximation errors.

1. Introduction

Discrete choice models have been widely used in the empirical industrial economics literature to estimate demand for differentiated products. In these models, consumers in market t can typically choose one of J_t products or an outside option. The consumers choose the product that maximize their utility, which leads to an expression of the markets shares. The parameters of the utility function can then be estimated using observed markets shares and product characteristics.

The challenge of the asymptotic theory in these models is to deal with several approximation errors. First, in models with heterogeneous consumers, such as the model of Berry et al. (1995) (referred to as the BLP model), the market shares involve integrals over the distribution of random coefficients. When estimating the parameters of the model, these integrals cannot be calculated analytically and are usually approximated by Monte Carlo integration with R random draws from the distribution of the random coefficients.[1] Second, many data sets do not contain true market shares but approximations from a sample of N consumers. Consequently, asymptotic theory needs to allow for three sources of errors: the simulation error in approximating the shares predicted by the model, the sampling error in estimating the market shares, and the underlying model error.

The limiting distribution of the estimated parameters can be obtained by either letting the number of products, the number of markets, or both approach infinity. Since the asymptotic distribution of the estimator serves as an approximation of its unknown finite sample distribution, it depends on the particular data set which approximation is most suitable. While in some cases using an approximation where the number of products approaches infinity is appropriate (as in Berry et al., 1995), in other cases the number of markets is a lot larger than the number of products (e.g. Nevo, 2001). As shown in this paper the asymptotic properties of the estimator differ a lot depending on which approximation is

☆ I thank an associate editor and three referees for helpful comments which have greatly improved the results and the exposition of the paper. I am grateful to Ivan Canay, Joel Horowitz, and Elie Tamer for helpful discussions and suggestions. I have also received valuable feedback from Mike Abito, Steven Berry, Mark Chicu, Roland Eisenhuth, Amit Gandhi, Aviv Nevo, Henrique de Oliveira, and Ketan Patel.

E-mail address: jfreyberger@ssc.wisc.edu.

[1] Although the focus lies on the effect of using Monte Carlo integration to approximate integrals, polynomial-based quadrature rules are also discussed. See Remark 4 for further details.

http://dx.doi.org/10.1016/j.jeconom.2014.10.009
0304-4076/

used. It is therefore important that both approximations are well understood.

Berry et al. (2004) provide asymptotic theory for a *large number of products in one market*. In their paper all market shares go to 0 at the rate $1/J$ and a necessary condition for asymptotic normality is that J^2/R and J^2/N are bounded. In this case, they find that their estimator $\hat{\theta}$ of the parameter vector θ_0 satisfies

$$\sqrt{J}\left(\hat{\theta}-\theta_0\right) \xrightarrow{d} N\left(0, V_{GMM}+\lambda_1 V_{MC}+\lambda_2 V_{SM}\right),$$

where $\lambda_1 = \lim_{J,R\to\infty} J^2/R$ and $\lambda_2 = \lim_{J,N\to\infty} J^2/N$. Here V_{GMM} denotes the variance of the estimator when the integral is calculated exactly and the true market shares are observed. V_{MC} and V_{SM} are additional variance terms due to the approximation error of the integrals and the market shares, respectively. Hence, if J^2/R or J^2/N are bounded away from 0, the asymptotic distribution of the estimated parameter vector is centered at 0 but Monte Carlo integration and market share approximation leads to a larger variance.

This paper is concerned with the asymptotic theory for a *small number of products, J_t, in a growing number of markets T*, which is the natural choice in many applications. However, this setup has not been considered in the literature so far. In Nevo (2001), for example, there are 1124 markets and 24 products (see also among others Kim, 2004, and Villas-Boas, 2007). Furthermore, in a similar (but more general) class of models, Berry and Haile (2014) provide nonparametric identification results for a large number of markets and a fixed number of products. These results can serve as a basis for nonparametric or semiparametric estimation of the model. Before such a flexible estimation procedure is developed, it is interesting to know how the commonly used fully parametric estimators behave. I prove consistency and asymptotic normality for these cases where T approaches infinity and $J_t \leq \bar{J}$ where $\bar{J}$ is fixed. The assumptions in the main part of the paper are stated for the BLP model, but the results in the appendix use higher level assumptions for a more general class of models. I find that the estimated parameters are $\sqrt{T}$ consistent and asymptotically normal as long as $\sqrt{T}/R$ and $\sqrt{T}/N$ are bounded. In this case, $\hat{\theta}$ satisfies

$$\sqrt{T}\left(\hat{\theta}-\theta_0\right) \xrightarrow{d} N\left(\tilde{\lambda}_1\mu_1+\tilde{\lambda}_2\mu_2, V_{GMM}\right)$$

where $\tilde{\lambda}_1 = \lim_{T,R\to\infty} \sqrt{T}/R$ and $\tilde{\lambda}_2 = \lim_{T,N\to\infty} \sqrt{T}/N$, again V_{GMM} is the variance of the estimator without approximation errors in the integrals and market shares, and μ_1 and μ_2 are constants. Hence, if $\sqrt{T}/R$ or $\sqrt{T}/N$ are bounded away from 0, Monte Carlo integration and market share approximation lead to an asymptotic distribution which is not centered at 0. The intuition for this result is that Monte Carlo integration (and similarly market share approximation) yields both additional bias as well as additional variance terms in the asymptotic expansion. The leading bias term is of order $O_p(\sqrt{T}/R)$, while the leading variance term is of order $O_p(1/\sqrt{R})$. Hence, the bias dominates the variance if R does not increase faster than T, and the variance does not affect the asymptotic distribution as long as $R \to \infty$. These results rely on using *different draws to approximate the integral in different markets*. If the same R draws are used in all markets one needs T/R to be bounded to obtain $\sqrt{T}$ consistency, which means that more draws are needed to approximate each integral relative to the number of markets.[2]

These results highlight that there are important differences between letting the number of products or the number of markets approach infinity. With a large number of products, it is important to correct the variance of the estimator due to the approximations. With a large number of markets and different draws in each market, the asymptotic distribution might not be centered at 0 and hence, a bias corrected estimator is needed. In both cases if R is too small, confidence intervals based on the usual asymptotic GMM distribution have the wrong size even asymptotically. Notice that when $\sqrt{T}/R$ and $\sqrt{T}/N$ converge to 0, the approximations do not affect the asymptotic distribution of the estimator. Contrary, in the setup where J approaches infinity, this result is only obtained when J^2/R and J^2/N converge to 0, which is a stronger requirement on the number of draws relative to the sample size.

The finite sample properties of the estimator depend on R and N due to both the additional bias and variance terms. I suggest an analytical bias correction which eliminates the leading bias term from the asymptotic distribution. I also show how one can easily incorporate the leading variance term when calculating standard errors. These two corrections greatly improve finite sample results. In particular, Monte Carlo simulations demonstrate that using a small number of simulation draws in comparison to the number of markets and using the usual GMM asymptotic distribution can yield distorted inference while the use of bias correction and adjusted standard errors leads to a considerably better performance.

These results might suggest that practitioners can simply use a very large number of draws and ignore Monte Carlo integration issues. Although this might be feasible depending on the model and computing resources available, in applications this is often not possible for several reasons. First, one does not know in advance how many draws suffice to obtain satisfactory results. As discussed in Section 4, the number of draws needed depends, among others, on the sample size, the number of random coefficients as well as unknown parameters, such as the variance of the random coefficients. Second, taking a very large number of draws is computationally very demanding because one needs to solve a complicated nonlinear optimization problem to estimate the parameters. The Monte Carlo results of the random coefficients logit model presented in Section 4 are based on a small number of products ($J = 4$) and six random coefficients to make the problem tractable. However, in the same setup as in Section 4 but with a sample size of $J = 24$ and $T = 1124$ (as in Nevo, 2001) it takes more than 24 h to minimize the objective function when $R = 2000$ and the starting values of the parameters are close to the true values.[3] Since we are dealing with a nonlinear optimization problem one needs to use several different starting values in applications. Taking a smaller number of draws considerably speeds up calculations. Third, even when taking the same draws for each product and each random coefficient, the number of draws needed is $T \times R$. In the previous example this means that 2248,000 draws are used to calculate the shares and the draws have to be stored before optimizing the function. As a consequence, more than 20 GB of memory is needed to run the program which is used to do the simulations in this paper. Finally, in case one wants to integrate over empirical distributions of demographic characteristics, R is constrained by the number of people in the database for a certain market.

The implication for empirical work with Monte Carlo integration or approximated market shares is that bias corrections and variance adjustments should be used. If the number of draws and the number of consumers is sufficiently large, the bias correction is close to 0 and the corrected standard errors will be very close to the

2 Remark 3 in Section 2.2 shows the differences in the asymptotic expansions when the same draws and different draws are used in each market. This section also highlights the differences between asymptotics in the number of markets and the number of products with and without different draws in different markets. See Remark 5 for details.

3 Computational details, including a description of the processors used, are presented in Section 4.