Simulation error and numerical instability in estimating random coefficient logit demand models

Daniel Brunner, Florian Heiss, André Romahn, Constantin Weiser

Journal of Econometrics 247 (2025) 105953

Contents lists available at ScienceDirect

Journal of Econometrics

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

Simulation error and numerical instability in estimating random coefficient logit demand models[☆]

Daniel Brunner[a], Florian Heiss[a], André Romahn[a,*], Constantin Weiser[b]

[a] Heinrich-Heine-Universität Düsseldorf, Germany
[b] Johannes Gutenberg Universität Mainz, Germany

ARTICLE INFO

JEL classification:
C13
C15
D12
L62

Keywords:
Structural demand estimation
Numerical integration

ABSTRACT

The nonlinear GMM-IV estimator of Berry, Levinsohn and Pakes (1995) can suffer from numerical instability resulting in a wide range of parameter estimates and economic implications. This has been reported to depend on technical details such as the choice of the optimization algorithm, starting values, and convergence criteria. We show that numerical approximation errors in the estimator's moment function are the main driver of this instability. With accurate approximation, the estimation approach is well-behaved. We provide a simple method to determine the required number of simulation draws.

1. Introduction

The seminal contribution of Berry et al. (1995; henceforth BLP) has become the workhorse for estimating differentiated product demand in Industrial Organization and its use has spread to other fields (see Table 1 of Berry and Haile (2016)). Consumers' preference heterogeneity depends on observed and unobserved attributes, or demographics, and consumer utility is driven by observed product characteristics and their interactions with demographics. Thereby, similar consumers make similar choices. This breaks the rigid functional form of the standard logit model and makes the relationship between prices, demand elasticities and the implied product markups more flexible.[1] The BLP model only requires market level data on quantities and prices, explicitly deals with endogenous product characteristics, typically price, and is parsimonious.[2]

Despite its many qualities, empirical practitioners have on occasion noted a lack of robustness in the model's estimates. In a detailed study that uses BLP's original data, Knittel and Metaxoglou (2014) demonstrate that changing starting values, optimization

☆ Financial support was provided by the DFG under the umbrella of the research group FOR 5392, Germany (project number 462020252). Computational support and infrastructure were provided by the "Centre for Information and Media Technology" (ZIM) at the University of Düsseldorf (Germany) and the Flemish Supercomputing Center (VSC) at the University of Leuven (Belgium). We would like to thank Richard Friberg, Mathias Reynaert, Joel Stiebale, Yutec Sun, Frank Verboven and seminar participants at EARIE 2017, the 2017 China Meeting of the Econometric Society, DICE, Télécom ParisTech, Tilburg University and the University of Cologne for helpful comments and suggestions. All remaining errors are our own.

* Corresponding author.
E-mail addresses: daniel.brunner@hhu.de (D. Brunner), florian.heiss@hhu.de (F. Heiss), romahn@hhu.de (A. Romahn), constantin.weiser@uni-mainz.de (C. Weiser).

[1] In the standard logit model, prices and markups of single-product firms are negatively related: high-priced products face highly elastic demand and thereby have lower markups than their low-priced rivals. Moreover, it is highly likely that the best substitute for any given product is the product with the largest market share.

[2] Software packages that implement the BLP estimation algorithm are freely available for R, https://cran.r-project.org/web/packages/BLPestimatoR/index.html, and for Python (see Conlon and Gortmaker (2020).

https://doi.org/10.1016/j.jeconom.2025.105953
Received 29 February 2020; Received in revised form 14 November 2024; Accepted 6 January 2025
Available online 27 January 2025
0304-4076/

algorithm and convergence criteria yields a wide range of coefficient estimates with widely varying economic implications.[3] We trace the source of such instabilities to numerical approximation errors in the model's market level share function, which must be solved by numerical integration. For practitioners, this emphasizes the importance of the estimator's asymptotic properties derived in Berry et al. (2004). With Monte Carlo simulation, the number of simulation draws must be proportional to the square of the number of products in the data to bound the effect of approximation errors in the aggregate share function. This result cannot be side stepped by adopting the alternative estimators of Dubé et al. (2012) or Grieco et al. (2023). The former is mathematically equivalent to BLP, but moves simulation error from the estimator's objective to constraints appended to the estimation. The latter estimator defaults to BLP in the absence of micro data, such as observed consumer level purchase decisions. We abstract from the availability of consumer-level data.

Section 2 briefly reviews the BLP estimation approach, while Section 3 outlines our computational experiments and the results. In Section 4, we provide practitioners with a method to determine a suitable number of simulation draws that is straightforward to implement.[4]

2. The BLP model

The econometrician observes prices, aggregate product-level quantities and some product attributes for $j = 0, 1, \dots, J$ differentiated products in $m = 1, \dots, M$ markets. $j = 0$ denotes the outside option. To reduce notation, we suppress the market index. The indirect utility consumer i derives from purchasing product j is typically specified as follows.

$$\begin{aligned} u_{ij} &= \delta_j + \mu_{ij} + \varepsilon_{ij} \\ \delta_j &= x_j\beta - \alpha p_j + \xi_j, \text{ and} \\ \mu_{ij} &= p_j\sigma_p\nu_{ip} + \sum_{k=1}^{K} x_{jk}\sigma_k\nu_{ik}, \ \ \nu_{ip}, \nu_{ik} \sim NID(0,1) \end{aligned} \tag{1}$$

δ_j and μ_{ij} are the mean utility and consumer-specific deviations from mean utility, respectively.[5] x is a row vector of K directly observable characteristics, p denotes price, and ξ is a product attribute that is unobserved by the econometrician. As is standard, we assume that the unobserved consumer tastes for the K characteristics follow a normal distribution and are independent of each other, so that the distribution of the random coefficients is given by $\alpha_i \sim N(\alpha, \sigma_p^2)$ and $\beta_{ik} \sim N(\beta_k, \sigma_k^2)$.[6]

2.1. Model-implied aggregate market shares

Assuming that the ε_{ij} follow a Gumbel distribution, gives consumer-specific choice probabilities with the familiar logit functional form. Aggregating over consumers, we obtain the model-implied market shares. Let $\boldsymbol{\delta}$ denote the vector of mean utilities and define $\boldsymbol{\theta} = (\sigma_p, \sigma_1, \dots, \sigma_K)'$.[7]

$$\begin{aligned} s_j(\boldsymbol{\delta}, \boldsymbol{\theta}, \boldsymbol{\nu}) &= \int_{-\infty}^{\infty} \frac{\exp(\delta_j + \mu_{ij}(\boldsymbol{\theta}, \nu_i))}{1 + \sum_{\ell=1}^{J} \exp(\delta_\ell + \mu_{i\ell}(\boldsymbol{\theta}, \nu_i))} d\boldsymbol{\nu} \\ &\approx \sum_{r=1}^{R} w_r \frac{\exp(\delta_j + \mu_{rj}(\boldsymbol{\theta}, \nu_r))}{1 + \sum_{\ell=1}^{J} \exp(\delta_\ell + \mu_{r\ell}(\boldsymbol{\theta}, \nu_r))} \end{aligned} \tag{2}$$

The second line stresses that the solution to the aggregate share function is not available analytically and must be solved for numerically, instead. The ν_r values are associated with weights w_r.[8]

2.2. Instrumental variables and identification

The unobserved characteristic or structural error term ξ_j is a vertical product attribute. Consumers always prefer more of it. Both firms and consumers observe $\boldsymbol{\xi} = (\xi_1, \dots, \xi_J)'$; the econometrician does not, which gives $Cov(p, \xi) > 0$. The preference parameters are estimated consistently by imposing the standard GMM-IV moment restriction. Let z_j denote a row vector of $L \geq K$ relevant and

[3] Pál and Sándor (2023) compare the ability of various optimization algorithms to obtain the estimator's global minimum. Replicating Knittel and Metaxoglou (2014) with high numerical accuracy, we do not find the choice of algorithm to matter for the US automobile data used in BLP, however. The only notable difference is computational time.

[4] Heiss and Winschel (2008) and Judd and Skrainka (2011) demonstrate the potential of numerical quadrature methods to attain very high accuracy at low computational cost. Computing error bounds for these methods, however, is model-specific and rather cumbersome.

[5] The additively separable specification for δ is standard, but can be extended to accommodate higher-order terms for x and p.

[6] The model can accommodate preference correlations. We also abstract from the possibility of mis-specifying the distributional form of preference heterogeneity.

[7] Nevo (2000) shows that the linearly entering parameters $(\alpha, \beta)'$ can be concentrated out.

[8] In standard Monte Carlo integration, the ν_r are obtained from pseudo-random draws from the standard normal distribution. The weights are simply $1/R$ for all r. In quasi-Monte Carlo integration or quadrature approaches, the ν_r and associated w_r are precomputed and not pseudorandom but the solution to an optimization problem. We take these procedures for granted and do not provide further details.

valid instrumental variables.

$$E\left[\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\nu})\right] = E\left[\frac{1}{J}\sum_{j=1}^{J} z_j \xi_j(\boldsymbol{\theta},\boldsymbol{\nu})\right] = 0. \tag{3}$$

Let $\boldsymbol{\theta}^*$ denote the true population preference parameters. Given a suitable weighting matrix W, we obtain a consistent and, as Berry et al. (2004) prove, asymptotically normally distributed estimator of $\boldsymbol{\theta}^*$ by minimizing the GMM-IV objective function.[9]

2.3. Inverting the aggregate market shares

The aggregate shares, (2), match their empirically observed counterparts, which we denote by $\mathcal{S}$. BLP prove the existence of the unique inverse $\boldsymbol{\delta}_{match} = s^{-1}(\mathcal{S};\boldsymbol{\theta},\boldsymbol{\nu})$.[10] The corresponding estimates of the structural error terms and linearly entering parameters $\hat{\alpha}$ and $\hat{\boldsymbol{\beta}}$ are obtained by two-stage least squares regression.

$$\delta_{j,match}\left(\mathcal{S};\hat{\boldsymbol{\theta}},\boldsymbol{\nu}\right) = x_j\hat{\boldsymbol{\beta}} + \hat{\alpha}p_j + \xi_{j,match}(\boldsymbol{\theta},\boldsymbol{\nu}) \quad \text{for all } j \tag{4}$$

2.4. Propagation of simulation error

To obtain the model's error term, the model-implied aggregate market shares must be computed using (2). Berry et al. (2004) show that at this point, simulation error propagates in the sample moments and thereby enters the identifying moment restriction, (3),

$$\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\nu}_R) \approx \frac{1}{J}\sum_{j=1}^{J} z_j \left(\xi^*_{match}(\boldsymbol{\theta}) - \left[\left.\frac{\partial s^*(\boldsymbol{\xi},\boldsymbol{\theta}^*)}{\partial \boldsymbol{\xi}'}\right|_{\xi^*}\right]^{-1} e(\boldsymbol{\theta},\boldsymbol{\nu}_R)\right), \tag{5}$$

where all terms with an asterisk are functions of the population of consumers and $\boldsymbol{\nu}_R$ is the simulated sample of R consumers.[11] $\boldsymbol{\theta}^*$ is the true vector of random coefficient standard deviations and $e(\boldsymbol{\theta},\boldsymbol{\nu}_R)$ is the vector of simulation errors that is caused by deviations between the consumer population and the simulated sample of consumers. These random errors affect the shape and location of the estimator's objective function, so that there can be several values $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$ for which the objective function attains a local minimum.

This distortionary effect of simulation error grows with the number of products in the market. As $J \to \infty$, the entries in the Jacobian of the model-implied market shares tend to zero as long as all products are (weak) substitutes and thereby the inverse of the Jacobian tends to infinity. To bound simulation error, Berry et al. (2004) prove that $R \propto J^2$. This nonlinear effect of simulation error also differentiates the BLP setting from that in McFadden (1989). There, simulation errors enter the objective linearly and thereby cancel out on average, which preserves asymptotic consistency even when the number of simulation draws is fixed.

3. Computational experiments

We adopt the data generating process (DGP) of Reynaert and Verboven (2014) to obtain synthetic differentiated product market data. By repeatedly estimating the BLP model for a given dataset with varying sets and numbers of simulation draws, we can vary the accuracy with which the model's aggregate share function, (2), is solved and thereby trace out how simulation error affects estimation outcomes. In our application, using standard Monte Carlo simulation has the advantage that the draws are pseudo random and dispersion measures for estimation outcomes are straightforward to obtain.[12]

3.1. Data generating process

To be brief, we only present the main points of interest of the DGP.[13] A single dataset consists of $M = 25$ markets, populated with $J = 10$ single-product firms each.[14] Consumer utility is specified as follows.

$$\begin{aligned} u_{ijm} &= \beta_0 + \beta_i x_{jm} - \alpha p_{jm} + \xi_{jm} + \varepsilon_{ijm} \\ \beta_i &= \beta_1 + \sigma v_i, \;\; v_i \sim N(0,1) \\ x_{jm} &\sim U(1,2) \end{aligned} \tag{6}$$

[9] $\xi(\boldsymbol{\theta},\boldsymbol{\nu})$ and Z are the vertically stacked market-specific structural error terms and instrumental variable matrices. The point estimate is obtained by $\hat{\boldsymbol{\theta}} = \arg\min_{\theta} \xi(\boldsymbol{\theta},\boldsymbol{\nu}_R)' Z W Z' \xi(\boldsymbol{\theta},\boldsymbol{\nu}_R)$.

[10] The inverse can be computed with a fixed point that satisfies the global contraction property: $\delta_{j,\text{iter}+1} = \delta_{j,\text{iter}} + \log(\mathcal{S}_j) - \log\left(s_j(\boldsymbol{\delta}_{\text{iter}};\boldsymbol{\theta},\boldsymbol{\nu})\right)$ for all j. The iteration continues until successive iterates become sufficiently close. This convergence criterion is set by the econometrician.

[11] The sample is either obtained using simulation under a distributional assumption or by randomly drawing from the population of consumers.

[12] Quasi-Monte Carlo simulation and quadrature have fixed draws or nodes. Obtaining error bounds for these more efficient approaches is cumbersome and model specific.

[13] The details are identical to those presented in Reynaert and Verboven (2014) pages 87–88.

[14] While $J = 10$ may seem a small number of products, what matters for the propagation of simulation error is the magnitude of market shares. In the generated data, the share of the outside option is between 0.81 and 0.95, which reduces the level of shares and thereby magnifies the effect of simulation error. Setting $J = 10$ also gives a manageable computational burden.

The ε's follow a Gumbel distribution, the observable product characteristic is distributed uniformly on the interval $(1, 2)$, the demand unobservable is distributed normally with a unit variance and the true coefficient values are $\beta_0 = 2$, $\beta_1 = 2$, $\alpha = -2$ and $\sigma = 1.0$. Prices and shares are computed endogenously. The demand unobservable is positively correlated with a supply side unobservable and the single product characteristic enters marginal cost positively along with three exogenous cost shocks that are all uniformly distributed. Each firm sets its price in order to maximize profits. To tackle the endogeneity of price, the three exogenous cost shocks, their squares and their interactions with the observed product characteristic, x, are available as relevant and valid instruments.

3.2. Estimation setup

To obtain results, we generate 1000 synthetic datasets using the above DGP.[15] Each estimation run is a combination of dataset, starting value for σ, and one of six approaches that we use to solve the model's aggregate share function, (2), numerically. To trace out the effect of simulation error in a simple and transparent way, the first three approaches use standard Monte Carlo sampling with $R = 10$, $R = 100$ and $R = 1000$ simulation draws. Given the result by Berry et al. (2004) that $R \propto J = 10$ to bound the effect of simulation error, the econometrician faces the difficulty of not observing the factor of proportionality. Approaching this issue with a range of simulation draws that covers three orders of magnitude seems reasonable to us. In Section 4, we present a formal approach to determine the required number of simulation draws for a given data set. The remaining three numerical integration approaches employ variance reduction techniques to attain a given level of approximation quality with fewer draws than standard Monte Carlo sampling. These are modified latin hypercube sampling (MLHS; Hess et al. (2006)), Halton sequences, and Gauss–Legendre quadrature.

Each model estimation is performed in two stages. In the first, we use the exogenous cost shifter instruments while in the second stage, we employ approximately optimal instruments as in Reynaert and Verboven (2014). With this setup, we run a total of almost 1.2 million BLP model estimations.[16] To arrive at the final set of estimation outcomes, we impose three criteria. First, the nonlinear optimizer must return an exit flag that indicates a local minimum has been found.[17] Second, the magnitude of the gradient must not exceed $1e-4$ and third, the Hessian matrix at the candidate minimum must be positive definite.[18] This leaves a total of 596,773 minima in the first stage (99.6% of all estimation runs) and 573,300 minima in the second stage (95.7% of all estimations).

3.3. Results

Pooling all outcomes, Fig. 1 presents the estimates of the nonlinearly entering parameter σ that we obtain using the three standard Monte Carlo integration approaches. The bottom row of plots is closest to the estimation setup of Reynaert and Verboven (2014; see their Fig. 1) and thereby also serves as a comparison and sanity check. We obtain the same spike at zero in the first stage of the estimation, which disappears in the second stage. Moreover, in both stages, the largest mass of outcomes is centred around the true coefficient value of 1. We therefore clearly match the overall estimation results of Reynaert and Verboven (2014).

The top and middle rows of Fig. 1 showcases our main result. Propagation of simulation error in BLP model estimation leads to biased coefficient estimates. Especially for the least accurate approach, $R = 10$, outcomes are even worse in the second stage. Recall that the optimal instrument approach of Chamberlain (1987) requires a consistent first stage. With a biased first stage, outcomes in the second stage are not consistent. As integration accuracy increases, the propagation of simulation error is bounded and we again obtain consistent estimators. In what follows, we focus on the estimation outcomes of the second stage and report additional results for the first stage in the Appendix. To trace out how simulation error affects estimation outcomes within datasets, we use the Gauss–Legendre estimates for each given dataset as the reference outcomes for the less accurate integration approaches and average the outcomes over all datasets. With 101 nodes, the Gauss–Legendre approach is able to exactly integrate polynomial integrands up to degree 201. We therefore expect simulation error to be negligible in these estimation runs. Table 1 shows how the statistical significance as well as the bias and RMSE of the nonlinear coefficient change as integration accuracy increases. The effect on model-implied economic outcomes is measured using the own-price and own-characteristic elasticities of demand. For all outcome measures, we see that more accurate numerical integration yields substantially reduced bias and variability. Accurate numerical integration also raises the fraction of estimations that correctly reject the standard logit model against the true BLP model specification.

3.4. The shape of the estimator's objective function

What is driving the bias and increased variability of estimation outcomes? To illustrate, we trace out the estimator's objective function for one dataset using ten independently sampled sets of simulation draws with different numbers of draws and different sampling techniques.[19] Fig. 2 presents the outcomes. For readability, we normalize each of the ten objective functions to have a value of zero for $\sigma = 0$. Nonlinearly entering simulation errors warp the shape of the objective, such that for one set of draws there can

[15] For one of the 1000 datasets, computing equilibrium prices suffers from ill conditioning. We drop this dataset, so that all that follows is based on 999 generated datasets.

[16] With one dataset dropped, we estimate specification (6) exactly $999 * 100 * 6 * 2 = 1,198,800$ times.

[17] We use Matlab's fminunc. The corresponding exit flag has a value of 1.

[18] Our results are robust to tightening the gradient criterion by an order of magnitude.

[19] The qualitative results are robust to choosing a different dataset.

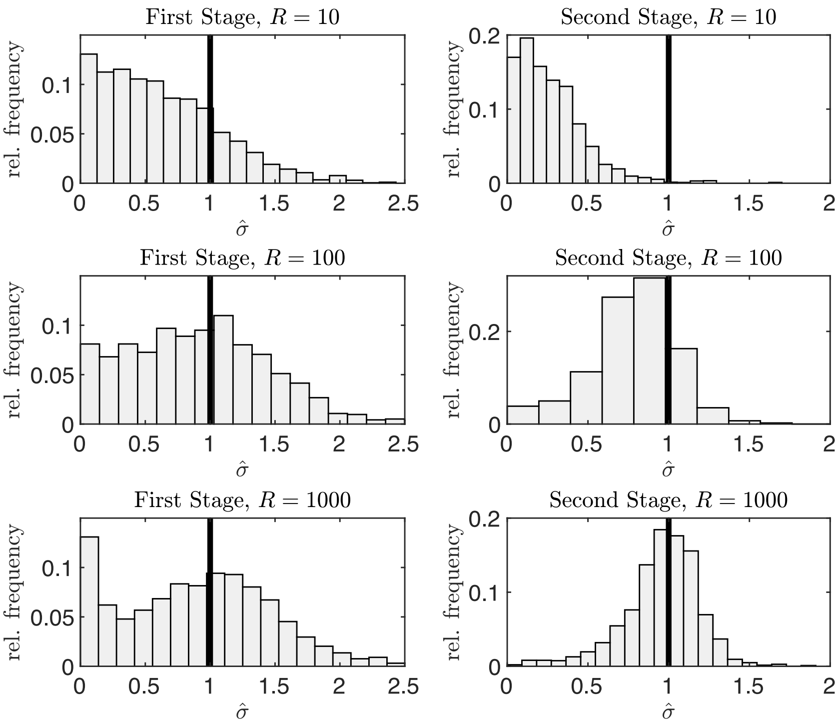

Fig. 1. The Effect of Simulation Error on the Distribution of $\hat{\sigma}$.
Note: All plots are based on estimation runs that return a local minimum of the GMM-IV objective function and that use the standard Monte Carlo (MC) sampling approach.

Table 1
The effect of simulation error on second-stage estimation outcomes.

	Wald Statistic, $H_0 : \hat{\sigma} = 0$			$\hat{\sigma}$ Within Datasets		Model-Implied $\hat{\eta}_p$		Model-Implied $\hat{\eta}_x$	
	Mean	Std. Dev.	Fraction H_1	Bias	RMSE	Bias	RMSE	Bias	RMSE
MC, $R = 10$	1685.7	31,977	0.8225	−0.7186	0.7685	0.3886	1.7950	−0.1550	0.8426
MC, $R = 100$	19.982	14.934	0.8091	−0.1851	0.2711	0.0803	1.0153	−0.0256	0.4856
MC, $R = 1000$	22.618	14.894	0.8561	−0.0221	0.1027	0.0101	0.3554	−0.0024	0.1700
MLHS, $R = 101$	22.543	14.810	0.8672	−0.0223	0.1016	0.0145	0.2852	−0.0068	0.1569
Halton, $R = 101$	22.691	14.714	0.8726	−0.0194	0.0958	0.0040	0.2652	0.0026	0.1404
Gauss–Legendre, $R = 101$	25.589	14.051	0.9526	–	–	–	–	–	–

Note: RMSE denotes root mean squared error and is computed using all bias numbers for a given numerical integration approach. MC stands for simple or standard Monte Carlo, where no variance reduction techniques are employed during sampling, and MLHS is the acronym for modified latin hypercube sampling, see Hess et al. (2006). We define the bias of the coefficient returned in estimation n as $\hat{\sigma}_n - \hat{\sigma}_n^{GL}$, where the GL superscript indicates that the reference estimate has been obtained using Gauss–Legendre quadrature. The bias and RMSE entries in the table are obtained by averaging these figures for all outcomes of a given numerical integration approach. This procedure takes account of the fact that each finite synthetic dataset delivers a set of estimates with different distances from the asymptotic outcome of returning the true model parameters. Gauss–Legendre quadrature in one dimension with 101 nodes is highly accurate and will exactly solve integrands of polynomial form for polynomials up to the 201st degree. We therefore expect the propagation of simulation error in these estimation runs to be negligible. The Gauss–Legendre estimates can therefore serve as a valid reference in these finite samples for tracing out how less accurate integration approaches affect BLP model estimates. η_p and η_x respectively denote the own-price and own-characteristic elasticity of demand. Thus, $\eta_p \equiv (p_j/s_j)[\partial s_j(\delta;\theta,\nu)/\partial p_j]$. For the elasticity bias and RMSE numbers, we compute the median deviation for each dataset as inputs. There tend to be a few outliers in each dataset, which mostly affects the RMSE numbers. Qualitatively, this choice does not affect the reported patterns.

be several local minima, as can be seen in the first panel. The locations of the minima also vary substantially between sets of draws with the same number of simulation draws. This effect vanishes with increasingly accurate numerical integration and the shape of the estimator's objective stabilizes, such that the outcomes for the ten sets of simulation draws cannot be visually distinguished and only a single minimum that is located very close to the true value of one remains.

4. Determining the required number of simulation draws

Following Caflisch (1998), we use a simple procedure to determine the required number of simulation draws in BLP model estimations. We first define the Monte Carlo integration error in the GMM-IV objective function for a given vector of the structural

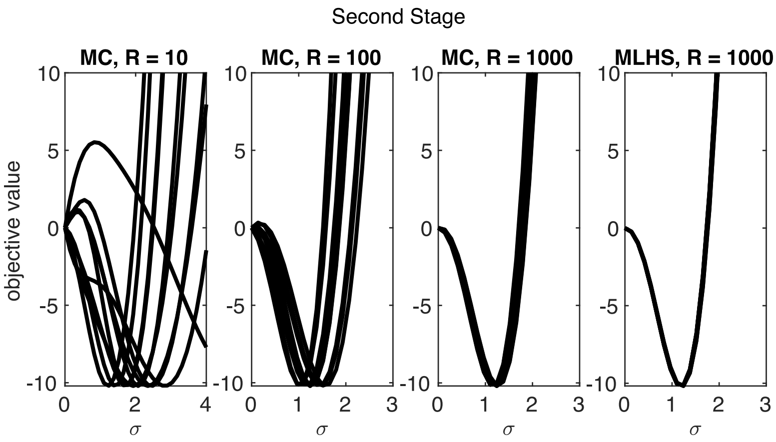

Fig. 2. The Effect of Simulation Error on the Shape of the Objective Function.
Note: For the 58[th] dataset, we trace out the objective function using evenly spaced values of σ. To reveal the effect of simulation error, we plot the resulting objectives for ten independently sampled sets of simulation draws. For readability, we subtract from each set of values the value of the objective at $\sigma = 0$, so that all traced out functions share the same vertical axis intercept. MC stands for standard Monte Carlo without variance reduction techniques, while MLHS is the modified latin hypercube sampling approach of Hess et al. (2006). The results are qualitatively identical and quantitatively close for other datasets.

parameters θ and a set of R simulation draws, which we denote by $\mathcal{J}(\theta, \nu_R) = \xi(\theta, \nu_R)^T ZWZ^T \xi(\theta, \nu_R)$.[20]

$$\epsilon_R(\theta) \equiv \mathcal{J}(\theta, \nu_\infty) - \mathcal{J}(\theta, \nu_R) \tag{7}$$

The bias of Monte Carlo integration is then $E(\epsilon_R)$, where the expectation is taken over independent sets of R draws, and its RMSE is $\sqrt{E(\epsilon_R^2)}$. Feller (1971) shows that for large R, the Monte Carlo integration error is approximately normal.

$$\epsilon_R(\theta) \approx \frac{\sigma_{\mathcal{J}}}{\sqrt{R}}\zeta, \ \zeta \sim N(0,1) \tag{8}$$

$\sigma_{\mathcal{J}}$ is the standard deviation of the BLP estimator's objective function that is unobserved by the econometrician. To obtain an estimate of the required number of simulation draws, the econometrician needs to (1) provide an estimate $of \sigma_{\mathcal{J}}$ and has to specify (2) an upper bound for the integration error, $\epsilon_R(\theta) < d$, as well as (3) the confidence level, c, at which the upper bound holds. The required number of simulation draws, $\underline{R}$, then solves the following equation.

$$Prob\left(\zeta < \frac{d}{\hat{\sigma}_{\mathcal{J}}}\sqrt{\underline{R}}\right) = c, \ \zeta \sim N(0,1) \tag{9}$$

Letting $z(c)$ denote the inverse of the standard normal density, we obtain the required number of draws at a given error target d and confidence level c.

$$\underline{R} = \frac{z(c)^2}{d^2}\hat{\sigma}_{\mathcal{J}}^2 \tag{10}$$

To estimate the unobserved $\sigma_{\mathcal{J}}$, we independently sample N_S sets of R simulation draws and compute the GMM-IV objective function value for each of these sets of draws. Let $\overline{\mathcal{J}}(\theta, R) = (1/N_S)\sum_{n=1}^{N_S} \mathcal{J}(\theta, \nu_{n,R})$ with some abuse of notation.

$$\hat{\sigma}_{\mathcal{J}} = \sqrt{R}\sqrt{N_S^{-1}\sum_{n=1}^{N_S}\left(\mathcal{J}(\theta, \nu_{n,R}) - \overline{\mathcal{J}}(\theta, R)\right)^2} \tag{11}$$

We perform this exercise for the 999 synthetic datasets used in our estimations by using 100 independently sampled sets of draws for standard Monte Carlo, MLHS and Halton sequences. We set an error bound of 0.1 and a confidence level of 99 percent.[21] Each dataset has its own unknown constant that shifts the objective function's variance and thereby also the required number of simulation draws. Table 2 presents the resulting distributions for $\underline{R}$.

At the median, simple Monte Carlo sampling requires roughly seven times as many draws as does MLH sampling and about eight times as many as Halton sequences. The computational burden of shielding BLP model estimates from simulation error can therefore be effectively reduced by employing more efficient numerical integration approaches; a point also made by Judd and Skrainka (2011) and Heiss and Winschel (2008).

[20] We cannot define the error on the aggregate share function (2) directly, because it is always matched up to a given numerical precision by construction and therefore does not vary with independently drawn sets of ν_R. Instead, we need to define the error on one of the outcomes of the matching procedure: $\xi(\theta, \nu_R)$ does vary between sets of simulation draws, which in turn affects the estimator's objective function.

[21] With Gauss–Legendre quadrature, the mean objective function value over all estimation runs in the final sample is 7.34. An RMSE of 0.1 gives a relative error of roughly 1.4 percent.

Table 2
Estimating the required number of simulation draws, $\underline{R}$.

	Minimum	25th	median	mean	75th	Maximum
MC	1725	4438	5694	5830	6983	12,779
MLHS	130	597	818	921	1153	3024
Halton	138	512	716	808	1011	2914

Note: Based on the 999 synthetic data sets used in the BLP model estimations of Section 3. To allow the outcomes for Halton sampling to vary, we use non-overlapping sets of Halton sequences.

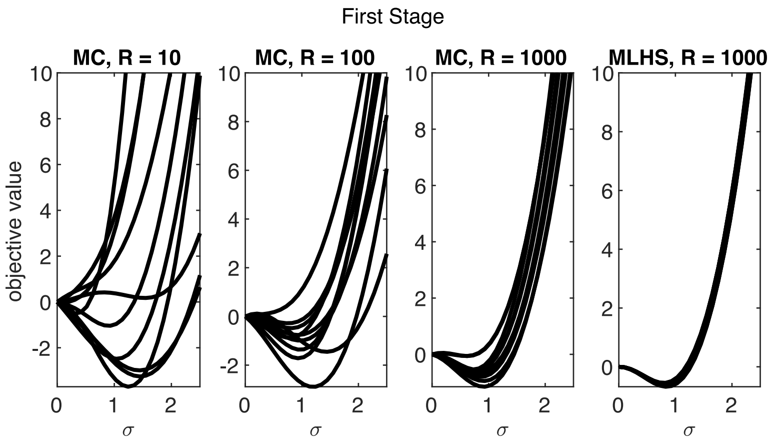

Fig. 3. The Effect of Simulation Error on the Shape of the Objective Function.
Note: For the 58th dataset, we trace out the objective function using evenly spaced values of σ. To reveal the effect of simulation error, we plot the resulting objectives for ten independently sampled sets of simulation draws. For readability, we subtract from each set of values the value of the objective at $\sigma = 0$, so that all traced out functions share the same vertical axis intercept. MC stands for standard Monte Carlo without variance reduction techniques, while MLHS is the modified latin hypercube sampling approach of Hess et al. (2006). The results are qualitatively identical and quantitatively close for other datasets.

5. Conclusion

Our findings demonstrate that bounding simulation error in BLP model estimations is of first-order concern for practitioners. Employing crude numerical integration techniques in the model's market-level share function warps the estimator's objective, such that several local minima emerge that yield biased coefficient estimates that vary substantially between independent estimation runs. Avoiding such brittle estimation outcomes is straightforward. We provide practitioners with a simple procedure to determine the required number of simulation draws for a given error target and confidence level that must be provided by the econometrician. This approach avoids making a substantial error in the choice of the number of simulation draws. Our findings also highlight the potential of widely used variance reduction techniques to reduce the computational burden of bounding simulation error in BLP model estimations.

Overall, with bounded simulation error and relevant and valid instruments, the BLP estimator is well behaved.

Appendix

We provide additional results for the first-stage estimation outcomes that use the cost shifter instruments. The results are very close to those of the second stage, which are presented in the main text. Table 3 presents the effect of simulation error on the statistical significance, bias and spread of the nonlinear parameter estimates as well as the resulting bias and spread of the model-implied economic outcomes, measured by the own-price and own-characteristic elasticities of demand.

Fig. 3 is the first-stage counterpart of Fig. 2. For the least accurate numerical integration approach, we can again clearly see that several local minima and a wide spread of minima appear. As integration accuracy increases, the objective's shape stabilizes and the resulting local minima can no longer be visually distinguished.

Table 3
The effect of simulation error on first-stage estimation outcomes.

	Wald Statistic, $H_0 : \hat{\sigma} = 0$			$\hat{\sigma}$ Within Datasets		Model-Implied $\hat{\eta}_p$		Model-Implied $\hat{\eta}_x$	
	Mean	Std. Dev.	Fraction H_1	Bias	RMSE	Bias	RMSE	Bias	RMSE
MC, $R = 10$	2.4541	4.5810	0.0439	−0.2287	0.6633	0.1657	2.4051	−0.0493	1.1273
MC, $R = 100$	1.4142	1.5427	0.0125	0.0567	0.3673	0.0145	1.0930	−0.0007	0.5154
MC, $R = 1000$	1.3522	1.5772	0.0150	0.0482	0.1786	−0.0096	0.3683	0.0041	0.1740
MLHS, $R = 101$	1.3368	1.5562	0.0150	0.0485	0.1838	−0.0020	0.3047	−0.0018	0.1584
Halton, $R = 101$	1.3230	1.5795	0.0155	0.0296	0.1737	−0.0007	0.2974	0.0016	0.1495
Gauss–Legendre, $R = 101$	1.3266	1.5986	0.0150	–	–	–	–	–	–

Note: RMSE denotes root mean squared error and is computed using all bias numbers for a given numerical integration approach. MC stands for simple or standard Monte Carlo, where no variance reduction techniques are employed during sampling, and MLHS is the acronym for modified latin hypercube sampling, see Hess et al. (2006). We define the bias of the coefficient returned in estimation n as $\hat{\sigma}_n - \hat{\sigma}_n^{GL}$, where the GL superscript indicates that the reference estimate has been obtained using Gauss–Legendre quadrature. The bias and RMSE entries in the table are obtained by averaging these figures for all outcomes of a given numerical integration approach. This procedure takes account of the fact that each finite synthetic dataset delivers a set of estimates with different distances from the asymptotic outcome of returning the true model parameters. Gauss–Legendre quadrature in one dimension with 101 nodes is highly accurate and will exactly solve integrands of polynomial form for polynomials up to the 201st degree. We therefore expect the propagation of simulation error in these estimation runs to be negligible. The Gauss–Legendre estimates can therefore serve as a valid reference in these finite samples for tracing out how less accurate integration approaches affect BLP model estimates. η_p and η_x respectively denote the own-price and own-characteristic elasticity of demand. Thus, $\eta_p \equiv (p_j/s_j)[\partial s_j(\delta;\theta,\nu)/\partial p_j]$. For the elasticity bias and RMSE numbers, we compute the median deviation for each dataset as inputs. There tend to be a few outliers in each dataset, which mostly affects the RMSE numbers. Qualitatively, this choice does not affect the reported patterns.

References

Berry, Steven T., Haile, Philip, 2016. Identification in differentiated products markets. Annu. Rev. Econ. 8, 27–52.

Berry, Steven T., Levinsohn, James, Pakes, Ariel, 1995. Automobile prices in market equilibrium. Econometrica 63 (4), 841–890.

Berry, Steven T., Linton, Oliver, Pakes, Ariel, 2004. Limit theorems for estimating the parameters of differentiated product demand systems. Rev. Econ. Stud. 71 (3), 613–654.

Caflisch, Russell E., 1998. Monte Carlo and quasi-Monte Carlo methods. Acta Numer. 7, 1–49.

Chamberlain, Gary, 1987. Asymptotic efficiency in estimation with conditional moment restrictions. J. Econometrics 34, 305–334.

Conlon, Christopher, Gortmaker, Jeff, 2020. Best practices for differentiated products demand estimation with PyBLP. Rand J. Econ. 51 (4), 1108–1161.

Dubé, Jean-Pierre, Fox, Jeremy T., Su, Che-Lin, 2012. Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. Econometrica 80 (5), 2231–2267.

Feller, William, 1971. An Introduction To Probability Theory and Its Applications: I. Wiley.

Grieco, Paul L.E., Murry, Charles, Pinkse, Joris, Sagl, Stephan, 2023. Conformant and efficient estimation of discrete choice demand models. mimeo.

Heiss, Florian, Winschel, Viktor, 2008. Likelihood approximation by numerical integration on sparse grids. J. Econometrics 144 (1), 62–80.

Hess, Stephane, Train, Kenneth E., Polak, John W., 2006. On the use of a modified latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. Transp. Res. B 40, 147–163.

Judd, Kenneth L., Skrainka, Benjamin S., 2011. High performance quadrature rules: How numerical integration affects a popular model of product differentiation. available at SSRN, https://doi.org/10.2139/ssrn.1870703.

Knittel, Christopher R., Metaxoglou, Konstantinos, 2014. Estimation of random-coefficient demand models: Two empiricists' perspective. Rev. Econ. Stat. 96 (1), 34–59.

McFadden, Daniel, 1989. A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica 57 (5), 995–1026.

Nevo, Aviv, 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. J. Econ. Manag. Strat. 9 (4), 513–548.

Reynaert, Mathias, Verboven, Frank, 2014. Improving the performance of random coefficients demand models: The role of optimal instruments. J. Econometrics 179 (1), 83–98.