Exact Simulation of Nested Logit Draws*

Maximilian S. Maydanchik†

September 3, 2026

Abstract

Nested logit models are simple to estimate, but nested logit errors have long resisted closed-form simulation. Galichon (2022) resolves a longstanding conjecture to show that nested logit errors decompose into standard Gumbel and positive stable components. I combine this result with the Kanter (1975) representation of positive stable variables to demonstrate a method for simulating error vectors for arbitrary trees. The method is fast, exact, and scales well. Monte Carlo simulations and timing benchmarks demonstrate its advantages over existing methods employing numerical inversion or moment-matching. It is a practical tool for discrete choice applications requiring full vectors of correlated shocks.

JEL Codes: C15, C25, C63

Keywords: nested logit model, positive stable distribution, discrete choice, random utility models, Monte Carlo simulation

*I thank Felix Tintelnot for initially posing the question to me of how to practically simulate nested logit draws. I also thank David Bunch for suggesting that I turn the resulting computational note into this short paper and for other helpful comments. All code is available upon request.

†University of Chicago, Kenneth C. Griffin Department of Economics. mmaydanchik@uchicago.edu.

1 Introduction

In discrete choice models, a decision-maker faces a set of alternatives and picks the one that maximizes their objective. The utility of each alternative is modeled as having a random component to capture unobserved heterogeneity in preferences. One common specification is the nested logit, which allows for correlation in these random components across alternatives grouped into *nests* (McFadden, 1978). Typically, these nests encompass common attributes shared among alternatives, such as public/private transit modes or branded/generic products.

Surprisingly, until recently, it was not known how to simulate from arbitrary nested logit distributions in closed form.[1] For many applications, this was not a barrier because analytic expressions for choice probabilities, inclusive values, and aggregate welfare are available. These analytic expressions yield a tractable way to model correlated random utility components without additional, computationally demanding dimensions of heterogeneity.[2] Still, the ability to simulate from nested logit distributions is a useful tool for more complex applications requiring the full vector of utility shocks. Some examples include the maximum simulated likelihood procedure estimating the consumer search model in Kaye (2026), welfare calculations under income effects in Herriges and Kling (1999), the integrated assessment model in Ramea et al. (2018), or the analysis of finite-sample bias in spatial models by Dingel and Tintelnot (2026).

In this paper, I demonstrate a fast, exact method for simulating draws from nested logit distributions of arbitrary depth and structure using the Kanter (1975) representation. To my knowledge, no prior empirical work using the nested logit and simulating errors has employed this method, relying on numerical or distributional approximations instead.

The Kanter method is enabled by a novel representation result due to Galichon (2022). Cardell (1997) derives that a nested logit draw, ε, can be decomposed into two random variables: ν_λ, from a unique distribution $C(\lambda)$, and ϵ, which is standard Gumbel: $\varepsilon = \nu_\lambda + \lambda\epsilon \sim \text{Gumbel}(0, 1)$. Correlation between alternatives is generated by making ν_λ a common component of alternatives within a nest. Galichon (2022) proves that $C(\lambda)$ is in fact a transformation of the positive stable distribution via $\nu_\lambda = \lambda \log Z$, where $Z \sim \text{PS}(\lambda)$. The proof establishes their equivalence using the Laplace transform.

[1] In principle, if one needed to simulate draws, one could use a Markov chain Monte Carlo sampler for the joint generalized extreme value density, which is known, as suggested by McFadden (1999), but Herriges and Kling (1999) note that it is computationally intensive and therefore of limited practical use.

[2] When substitution patterns between products or pass-through of costs to prices are key objects, as may be the case in, e.g., merger simulation, the nested logit may still be too restrictive (Grigolon and Verboven, 2014; Miravete, Seim and Thurk, 2023).

To simulate draws, Galichon (2022) directs readers to Ridout (2009), who proposes an algorithm for the simulation of distributions by their Laplace transforms. However, closed form simulation of positive stable variables has long been known using Kanter's representation (Chambers, Mallows and Stuck, 1976; Stephenson, 2003).[3] Specifically, a positive stable distribution with parameter λ can be represented as

$$Z = \left(\frac{\sin((1-\lambda)U)}{W}\right)^{\frac{1-\lambda}{\lambda}} \frac{\sin(\lambda U)}{(\sin U)^{\frac{1}{\lambda}}} \sim PS(\lambda), \tag{1}$$

where $U \sim \text{Uniform}(0, \pi)$ and $W \sim \text{Exponential}(1)$. That is, Z is a simple combination of a uniform and exponential draw; ν_λ then follows directly. In the degenerate case of $\lambda = 1$, simulation of ν_λ can be avoided as this reduces to a multinomial logit.

2 Nested Logit Representation

Consider first the two-level nested logit, in which the first level contains $n = 1, ...N$ nests and the second level contains $j = 1, ..., J$ choices.[4] Figure 1, Panel (a) illustrates a simple example with two nests and two choices each. Call $n(j)$ the nest to which choice j belongs. Each nest has a parameter, $\lambda_n \in (0, 1]$, which controls the correlation between the random utility components of choices sharing that nest. Cardell (1997) shows that the random utility component of each choice can be represented as

$$\varepsilon_j = \nu_{n(j)} + \lambda_{n(j)}\epsilon_j \sim \text{Gumbel}(0, 1), \tag{2}$$

where there exists a unique distribution $C(\lambda)$ such that $\nu \sim C(\lambda)$, $\epsilon \sim \text{Gumbel}(0, 1)$, and ν and ϵ are independent.

Cardell (1997) also shows that arbitrary tree structures can be recursively represented in this way. For example, Figure 1, Panel (b) illustrates a three-level nested logit with two nests at the first level, each having two subnests at the second level, and each of those having two choices at the third level. Represent the depth of the tree as $d = 0, ..., D$, where each nest $n(d)$ at depth $d = 1, ..., D-1$ has its own parameter, $\lambda_{n(d)}$, the alternatives exist at the bottom level $d = D$, and the root node, at $d = 0$, is assigned $\lambda_{n(0)} = 1$ by convention. Each choice j belongs to a sequence of nests $n_j(1), \ldots, n_j(D-1)$. In this

[3] In a parallel literature, Stephenson (2003) derives a positive stable representation for a related class of extreme value distributions, of which the nested logit with a single nest is a special case. The R package `evd` implements his simulation algorithm for this class of distributions using the Kanter method. However, it does not handle deeper trees and scales poorly in the number of alternatives (Townsend, 2025).

[4] This notation follows Galichon (2022).

case, the random utility component of each choice can be represented as (Galichon, 2022, Equation (11))

$$\varepsilon_j = \sum_{d=1}^{D-1} \Lambda_{n_j(d-1)} \nu_{n_j(d)} + \Lambda_{n_j(D-1)} \epsilon_j, \qquad \Lambda_{n_j(d)} = \prod_{t=0}^{d} \lambda_{n_j(t)}. \tag{3}$$

$\Lambda_{n_j(d)}$ is simply the product of nest parameters along the path from the root to depth d on the way to choice j.[5] For example, a specification of the three-level nested logit in Figure 1, Panel (b) if both nests had common parameters λ_1 and all subnests had common parameters λ_2 would look like $\varepsilon_j = \nu_1 + \lambda_1 \nu_2 + \lambda_1 \lambda_2 \epsilon_j$.

3 Simulation Using the Kanter Method

The Kanter method applied to the nested logit proceeds in two steps:

1. *Nests*: for each nest $n(d)$ at depth $d = 1, \ldots, D-1$ with parameter $\lambda = \lambda_{n(d)}$:
   - Draw N : $U \sim \text{Uniform}(0, \pi)$
   - Draw N : $W \sim \text{Exponential}(1)$
   - Form N : $\nu_{n(d)} = (1-\lambda)\Big[\log(\sin((1-\lambda)U)) - \log W\Big] + \lambda \log(\sin(\lambda U)) - \log(\sin U)$
2. *Choices*: for each choice $j = 1, \ldots, J$ belonging to nests $n_j(1), \ldots, n_j(D-1)$:
   - Draw J : $\epsilon_j \sim \text{Gumbel}(0, 1)$
   - Form D-1 : $\Lambda_{n_j(d)} = \prod_{t=0}^{d} \lambda_{n_j(t)}$ for $d = 0, \ldots, D-1$
   - Form J : $\varepsilon_j = \sum_{d=1}^{D-1} \Lambda_{n_j(d-1)} \nu_{n_j(d)} + \Lambda_{n_j(D-1)} \epsilon_j$

For each simulated individual, the total number of draws is only $J + 2N$—the number of choices plus twice the number of nests—which can be held fixed across candidate λ values. This eliminates resampling noise within estimation routines, requiring only cheap additional computations to update ν, Λ, and ε.

Figure 2 validates the method with a more complex tree structure using 200,000 draws of nested logit error vectors. Panel (a) displays the tree: it has depth 5, with 15 different nesting parameters and 25 alternatives, including a singleton attached directly to the root that functions as the commonly included "outside option" in discrete choice models. Panel (b) plots a histogram of all $200{,}000 \times 25 = 5$ million errors against the analytic Gumbel density, validating the implied Gumbel marginals. Panel (c) plots the simulated correlation

[5] The correlation between two choices j and k comes from the deepest nest that they have in common. Call that parent nest $p := n_j(d^+)$ where $d^+ = \max\{d : n_j(d) = n_k(d)\}$. Then, $\text{corr}(\varepsilon_j, \varepsilon_k) = 1 - (\Lambda_p)^2$. Since $\lambda_n \in (0, 1]$, this implies that as alternatives share deeper nests, their correlation increases.

matrix between alternatives. Panel (d) plots the difference between the implied analytic correlations and the 300 unique off-diagonal entries in the matrix. These simulated errors are consistent with Monte Carlo noise, which is on the order of $1/\sqrt{200,000} \approx 0.002$. Finally, because these pairwise errors are not independent, Panel (e) bootstraps the mean pairwise correlation error over 500 resamples of the same 200,000 vectors, and the resulting 95% confidence interval covers zero.

4 Comparison Across Methods

Having demonstrated the algorithm, I close by benchmarking the performance of the Kanter (1975) representation (hereafter, the *Exact* method) in speed and scaling against three alternatives from the literature. First, and already mentioned, the Ridout (2009) approach (hereafter, the *Laplace* method) numerically inverts the Laplace transform to approximate the positive stable CDF and density. Second and third, two methods use Cardell (1997)'s provided characteristic function for $C(\lambda)$. Most recently, Townsend (2025) employs the fast Fourier transform (hereafter, the *Fourier* method) to numerically approximate the density and integrate it into a CDF on a grid. Alternatively, Bunch and Rocke (2016) take a moment-matching approach and compute the first four moments (mean, variance, skewness, and kurtosis), simulating draws from the resulting four-parameter Pearson distribution (hereafter, the *Pearson* method).[6]

Relative to the *Exact* method, each of these alternatives must balance trade-offs. The *Laplace* and *Fourier* methods are convergent approximations given tight tolerances, but they must repeat their expensive numerical calculations. The *Laplace* method pays this cost for each simulated draw, causing it to scale poorly in the total number of draws; the *Fourier* method pays this cost upfront by pre-computing the CDF for a given λ, allowing it to scale well in draws, but then scales poorly in the number of λ parameters. Conversely, the *Pearson* method scales well in both draws and parameters, but this comes at the cost of accuracy. Failing to match higher moments may affect outcomes dependent on the nested logit distribution's tails.

Table 1 presents the accuracy results. Each method is used to simulate 2,000 replications of 100,000 draws of the Figure 1, Panel (a) tree for $\lambda = \{0.1, 0.5, 0.9\}$. The first three rows

[6]Refer to the cited papers for details. The Laplace method is implemented using Ridout (2009)'s provided R function `rlaptrans`. One modification is made to fix sampling points across candidate λ values, which improves performance. The Fourier method is a custom implementation of Townsend (2025)'s provided R package `nev`, based on `fourierin`, amended to handle arbitrary tree structures and also to reuse sampling points. The Pearson method uses the R package `PearsonDS`. All computation was done on the University of Chicago compute cluster using Intel Xeon E5-2690 v4 @ 2.60GHz CPUs.

report the frequency with which Anderson-Darling goodness-of-fit tests of each marginal distribution against the analytic Gumbel$(0, 1)$ null reject at the 10%, 5%, and 1% levels. For a method that exactly generates the Gumbel, the rejection frequency should equal the nominal level. Systematic discrepancies would indicate that the simulated distribution differs from the Gumbel. The next four rows show simulated vs. analytic quantile errors at the median, 90th, 99th, and 99.9th percentiles. As expected, the *Exact*, *Laplace*, and *Fourier* methods reject at close to their nominal rates and display minimal quantile errors. In contrast, the *Pearson* method rejects the Gumbel null in a large fraction of tests for $\lambda = \{0.1, 0.5\}$ and shows greater quantile errors in the tails. At $\lambda = 0.9$ the *Pearson* rejection frequency falls to 0.096 at the 5% level, an order of magnitude closer than at $\lambda = \{0.1, 0.5\}$, though still nearly twice the nominal rate. The quantile errors are similarly more modest. This suggests that it may provide a reasonable approximation at lower levels of correlation between alternatives. Finally, the bottom row shows the mean signed error across all pairwise correlations, which is negligible for all four methods.

Table 2 presents the timing and scaling results. Panel (a) shows the time to simulate $S = 100, 1{,}000, ..., 10{,}000{,}000$ draws of a single nested logit random variable. The *Exact* and *Pearson* methods are generally fastest, while the *Laplace* method is generally uncompetitive. The *Fourier* method starts slowest due to the upfront cost of pre-computing the CDF, but as this cost is amortized over more draws, it actually becomes the fastest across methods. Hence, the *Fourier* method is feasible when many draws are needed at a known λ.

In practice, however, the λ parameter will not be known, and estimation routines will need to estimate λ. Errors must be recomputed for each candidate value. Panel (b) repeats $S = 100$ draws of a single random variable for $L = 1, 10, 100, 1{,}000$ distinct λ values. Here, the *Exact* method is universally fastest, nearly an order of magnitude faster than *Pearson* and over 100x faster than *Laplace*. The *Fourier* method is slowest, a consequence of the fact that different λ values require separate expensive upfront computations. Finally, Panel (c) repeats the same experiment for the full 25-alternative, 15-nest tree in Figure 2, which stands in for a more realistic choice set in practice. The results are the same, with even extremely large routines of the *Exact* method taking less than one second.

These results demonstrate that the Kanter method is fast, exact, and scales well. It is simple to implement in code and no more difficult to implement for arbitrarily deep trees than for a single nest. Its only error is ordinary Monte Carlo noise. Reusing draws makes repeated evaluations in an estimation routine cheap and free of resampling noise. It should be the preferred method for researchers looking to simulate nested logit draws.

References

Bunch, David S., and David M. Rocke. 2016. "Variance-Component-Based Nested Logit Specifications: Improved Formulation, and Practical Microsimulation of Random Disturbance Terms." *Journal of Choice Modelling*, 21: 30–35.

Cardell, N. Scott. 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity." *Econometric Theory*, 13(2): 185–213.

Chambers, John M., Colin L. Mallows, and B. W. Stuck. 1976. "A Method for Simulating Stable Random Variables." *Journal of the American Statistical Association*, 71(354): 340–344.

Dingel, Jonathan I., and Felix Tintelnot. 2026. "Spatial Economics for Granular Settings." *Econometrica*, 94(2): 407–464.

Galichon, Alfred. 2022. "On the Representation of the Nested Logit Model." *Econometric Theory*, 38(2): 370–380.

Grigolon, Laura, and Frank Verboven. 2014. "Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation." *Review of Economics and Statistics*, 96(5): 916–935.

Herriges, Joseph A., and Catherine L. Kling. 1999. "Nonlinear Income Effects in Random Utility Models." *Review of Economics and Statistics*, 81(1): 62–72.

Kanter, Marek. 1975. "Stable Densities Under Change of Scale and Total Variation Inequalities." *Annals of Probability*, 3(4): 697–707.

Kaye, Aaron P. 2026. "The Personalization Paradox: Welfare Effects of Personalized Recommendations in Two-Sided Digital Markets." Working paper, Boston University and MIT.

McFadden, Daniel. 1978. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models.* , ed. Anders Karlqvist, Lars Lundqvist, Folke Snickars and Jorgen W. Weibull, 75–96. Amsterdam:North-Holland.

McFadden, Daniel. 1999. "Computing Willingness-to-Pay in Random Utility Models." In *Trade, Theory and Econometrics: Essays in Honour of John S. Chipman.* , ed. James R. Melvin, James C. Moore and Raymond Riezman, 253–274. London:Routledge.

Miravete, Eugenio J., Katja Seim, and Jeff Thurk. 2023. "Pass-Through and Tax Incidence in Differentiated Product Markets." *International Journal of Industrial Organization*, 90: 102985.

Ramea, Kalai, David S. Bunch, Christopher Yang, Sonia Yeh, and Joan M. Ogden. 2018. "Integration of Behavioral Effects from Vehicle Choice Models into Long-Term Energy Systems Optimization Models." *Energy Economics*, 74: 663–676.

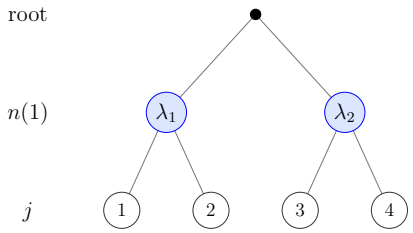
Ridout, M. S. 2009. "Generating Random Numbers from a Distribution Specified by Its Laplace Transform." *Statistics and Computing*, 19(4): 439–450.

Stephenson, Alec. 2003. "Simulating Multivariate Extreme Value Distributions of Logistic Type." *Extremes*, 6(1): 49–59.

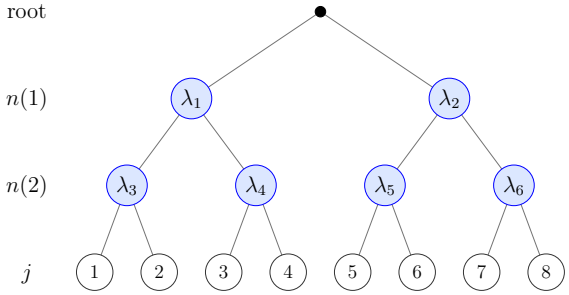
Townsend, Wilbur. 2025. "A Novel Algorithm for Drawing Nested Extreme Value Random Variables." *Journal of Choice Modelling*, 57: 100575.

Figure 1: Illustrative nested-logit trees

(a) Two-level nested logit



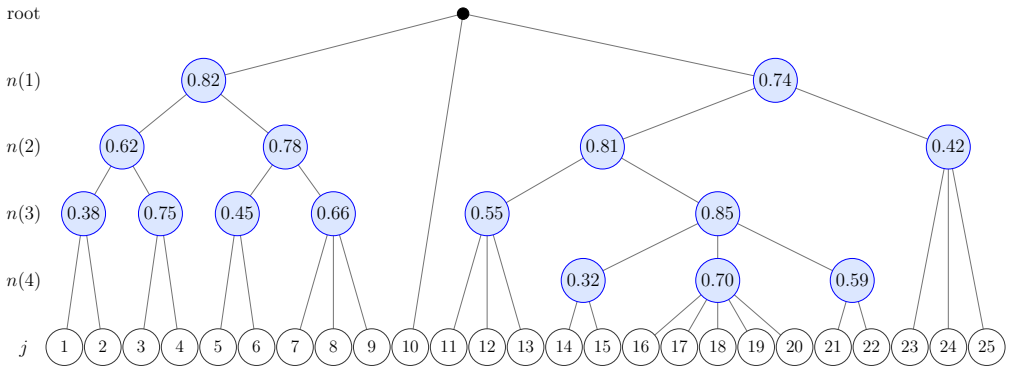
(b) Three-level nested logit



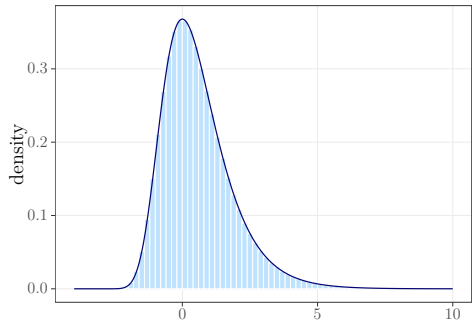
Notes: Panel (a)—two-level tree with two nests of two choices each. Panel (b)—three-level tree with two nests, two subnests within each nest, and two choices within each subnest. Blue nodes are nests with parameter λ_n. White nodes are alternatives indexed by j. The black dot indicates the root node, which has $\lambda_0 = 1$ by convention. The left margin labels each row, from the *root* at top through depth $n(d)$ for each nest level to j for the alternatives.

Figure 2: Validation of the Kanter method

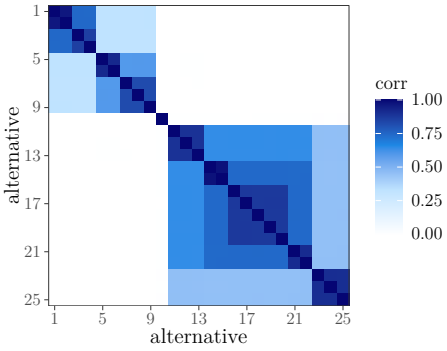
(a) Five-level nested logit



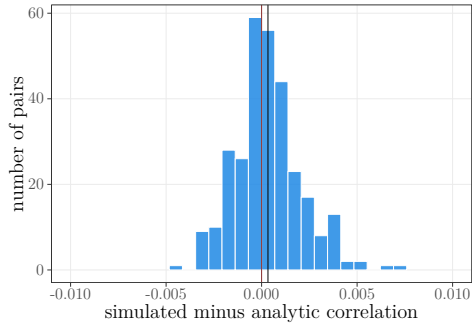
(b) Histogram of draws vs. analytic pdf



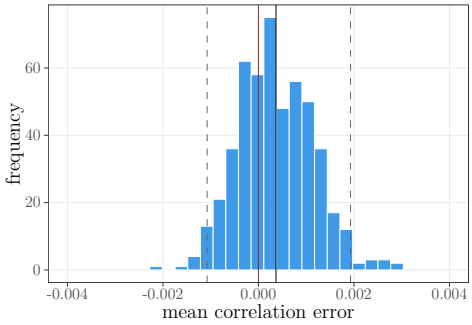
(c) Simulated correlation matrix



(d) Pairwise correlation errors



(e) Bootstrapped mean correlation error



Notes: Panel (a)—tree structure, with each blue node containing nest parameter λ_n. Panel (b)—histogram of all simulated nested logit errors from 200,000 draws of 25 errors. Analytic Gumbel$(0, 1)$ density overlain. Panel (c)—simulated pairwise correlation matrix between alternatives across draws. Shading reflects strength of correlation. Panel (d)—histogram of simulated minus analytic correlation errors for all 300 unique off-diagonal pairwise correlations. Solid black line is the mean error (0.0003). Panel (e)—distribution of the mean error over unique off-diagonal pairwise correlations across 500 bootstrap resamples with replacement of the same 200,000 draws of error vectors. Solid black line is the bootstrap mean (0.0004) and dashed lines are 95% confidence interval bounds ($[-0.0011, +0.0019]$).

Table 1: Accuracy of methods in simulating nested logit draws across λ

Panel A: $\lambda = 0.1$				
Statistic	Exact (Kanter)	Laplace (Ridout)	Fourier (Townsend)	Pearson (Bunch-Rocke)
AD reject rate - 10%	0.098 (0.004)	0.109 (0.005)	0.105 (0.005)	0.996 (0.001)
AD reject rate - 5%	0.049 (0.003)	0.050 (0.003)	0.052 (0.003)	0.965 (0.003)
AD reject rate - 1%	0.011 (0.002)	0.008 (0.001)	0.012 (0.002)	0.674 (0.007)
p_{50} error	+0.0001	+0.0002	−0.0006	−0.0015
p_{90} error	−0.0000	−0.0002	−0.0001	+0.0073
p_{99} error	−0.0007	+0.0003	+0.0008	−0.0297
$p_{99.9}$ error	−0.0025	−0.0008	+0.0000	−0.0064
Corr. error	+0.0000	−0.0001	−0.0000	+0.0001
Panel B: $\lambda = 0.5$				
Statistic	Exact (Kanter)	Laplace (Ridout)	Fourier (Townsend)	Pearson (Bunch-Rocke)
AD reject rate - 10%	0.111 (0.004)	0.106 (0.004)	0.103 (0.004)	0.986 (0.001)
AD reject rate - 5%	0.057 (0.003)	0.057 (0.003)	0.053 (0.003)	0.911 (0.003)
AD reject rate - 1%	0.010 (0.001)	0.014 (0.001)	0.012 (0.001)	0.510 (0.006)
p_{50} error	−0.0000	+0.0000	−0.0003	−0.0024
p_{90} error	−0.0002	−0.0003	−0.0002	+0.0077
p_{99} error	−0.0003	+0.0000	−0.0001	−0.0283
$p_{99.9}$ error	+0.0005	−0.0001	+0.0002	−0.0052
Corr. error	−0.0000	−0.0001	+0.0001	−0.0000
Panel C: $\lambda = 0.9$				
Statistic	Exact (Kanter)	Laplace (Ridout)	Fourier (Townsend)	Pearson (Bunch-Rocke)
AD reject rate - 10%	0.095 (0.003)	0.100 (0.003)	0.104 (0.003)	0.186 (0.004)
AD reject rate - 5%	0.048 (0.002)	0.052 (0.003)	0.051 (0.002)	0.096 (0.003)
AD reject rate - 1%	0.011 (0.001)	0.008 (0.001)	0.011 (0.001)	0.019 (0.002)
p_{50} error	+0.0001	−0.0001	−0.0002	−0.0020
p_{90} error	+0.0001	+0.0000	−0.0004	+0.0042
p_{99} error	+0.0002	+0.0003	−0.0017	−0.0121
$p_{99.9}$ error	−0.0009	−0.0000	−0.0103	−0.0155
Corr. error	−0.0000	−0.0001	−0.0001	−0.0001

Notes: Each panel repeats 2,000 simulations of 100,000 draws of a two-level nested logit with two nests of two choices each as in Figure 1, Panel (a), with the nest parameter λ fixed at the panel's value. Each simulation runs one Anderson-Darling goodness-of-fit test per alternative against an analytic Gumbel$(0, 1)$ null hypothesis, giving $2{,}000 \times 4 = 8{,}000$ tests of marginal distributions per method. AD rejection rate is the fraction of those tests with p-value below the nominal level; parentheses report Monte Carlo standard errors clustered by simulation-nest pair, since the two alternatives within a nest share a common nest component. Quantile errors are differences between simulated and analytic Gumbel$(0, 1)$ quantiles at the median, 90th, 99th, and 99.9th percentiles. Correlation error is the mean of all pairwise simulated-minus-analytic correlation errors. Quantile and correlation errors report means across simulations.

Table 2: Computation time and scaling across methods

Panel A: seconds, S draws of single random variable				
S	Exact (Kanter)	Laplace (Ridout)	Fourier (Townsend)	Pearson (Bunch-Rocke)
100	< 0.0001	0.0064	0.0799	0.0003
1,000	0.0004	0.0420	0.0755	0.0007
10,000	0.0041	0.2597	0.0761	0.0042
100,000	0.0407	2.1520	0.0996	0.0384
1,000,000	0.4245	19.3700	0.3550	0.3828
10,000,000	4.6300	112.7280	3.2290	4.0080
Panel B: seconds, $S = 100$ draws of single random variable for L distinct λ values				
L	Exact (Kanter)	Laplace (Ridout)	Fourier (Townsend)	Pearson (Bunch-Rocke)
1	< 0.0001	0.0068	0.0798	0.0004
10	0.0003	0.0808	0.7892	0.0033
100	0.0029	0.7220	7.9020	0.0342
1000	0.0284	7.7450	78.9580	0.3502
Panel C: seconds, $S = 100$ draws of Figure 2 tree random vector for L distinct λ vectors				
L	Exact (Kanter)	Laplace (Ridout)	Fourier (Townsend)	Pearson (Bunch-Rocke)
1	0.0015	0.0959	1.1160	0.0056
10	0.0068	0.9655	11.3170	0.0523
100	0.0650	9.5910	112.3090	0.5315
1000	0.6315	102.6060	1132.9480	5.3130

Notes: Each cell is wall-clock time for an independent experiment. Within an experiment, draws and inputs are reused across parameter evaluations where possible, but they are not reused across experiments. Panel (a)—S draws of a single nested logit random variable with $\lambda = 0.5$. Panel (b)—repeats $S = 100$ draws of a single random variable (as in the first row of Panel A) for L distinct λ values. Panel (c)—repeats $S = 100$ draws of the 25-alternative, 15-nest tree in Figure 2 for L distinct λ vectors. To reduce noise in short-run measurements, each cell is repeated until total time is at least one second, and the average over repetitions is reported.