Quantifying the Benefits of New Products with Residual-Income Demand: A Comment on Petrin (2002)

Chae-Yeon Xon

Yonsei University

September 2026

Abstract

Petrin (2002) specifies automobile demand with logarithmic residual income, $\alpha_i \log(y_i - p_j)$, while the author-distributed estimation and markup programs use its first-order price term, $-\alpha_g p_j / y_i$. This Comment traces the two operators through choice probabilities, price derivatives, markups, and welfare. A finite economy gives different rankings, choice probabilities, and compensating variation under a common shock law. In a public reconstruction, the linear operator reproduces the reported mean income of minivan purchasers, 36.091 thousand dollars, with a value of 36.051. Replacing the operator by the printed affordable-set model and recontracting every product share gives 45.349; a matched seven-parameter profile gives 40.456. The demand-moment comparison changes direction relative to the CEX value. A Taylor bound measures the approximation error. Three implementations using a common utility, derivative, and support show which welfare calculations the available archive can reproduce.

JEL Codes: C35, C51, D12, L62

Keywords: discrete choice, income effects, micro moments, new products, compensating variation

Yonsei University. eunixon@yonsei.ac.kr. Early stage draft; please do not cite or circulate without author's written permission. All errors are my own.

1. Introduction

Petrin (2002) measures the welfare contribution of the minivan with product-market shares and purchaser demographics. The demand system augments the random-coefficients framework of Berry, Levinsohn, and Pakes (1995) with logarithmic utility from residual income. For consumer i and vehicle j, the price component is $\alpha_i \log(y_i - p_j)$, and the outside option contributes $\alpha_i \log y_i$. Purchaser-income moments discipline substitution across demographic groups, while the same utility system supplies demand derivatives, markups, counterfactual prices, and compensating variation.

The author-distributed GAUSS and C programs evaluate another price component. They set the group price coefficient equal to α_g / y_i and use $-\alpha_g p_j / y_i$ in utility and in the markup derivative. The April 2001 working paper already states the logarithmic model, and the distributed source files are dated July 2001. The available paper and code therefore define two estimands along the same source lineage.

This Comment follows that mapping through three exercises. First, it writes both models on a common normalized utility scale and bounds their difference. Second, it gives a two-product economy in which the operators reverse the deterministic ranking and change choice probabilities and welfare. Third, it reconstructs a reported purchaser-income moment using the public product and demographic files. The linear branch gives 36.051 thousand dollars, close to the reported model value 36.091. The logarithmic affordable-set branch gives 45.349 after both branches are recontracted to the same observed market shares. Re-estimating seven demographic coefficients under one public objective moves the logarithmic value to 40.456, while the comparison with the CEX mean 39.476 keeps the reversed direction.

The archive provides the estimation and markup paths and omits the final importance-sampling state and welfare routine. The calculation therefore reaches the demand moment directly. Tables reporting demand coefficients, marginal costs, equilibrium prices, profits, and compensating variation require a joint re-estimation and one declared welfare law. The estimator architecture, demographic micro moments, and Bertrand first-order conditions remain available after the utility and derivative operators are aligned.

The distinction also appears in later implementations. The implementations in Conlon and Gortmaker (2025) and the accompanying PyBLP reconstruction use $-p/y$ and reproduce the minivan exercise as a linear specification. Griffith, Nesheim, and O'Connell (2018) classify Petrin's application as an example of nonlinear income effects. Together, these sources yield two readings: the article states the logarithmic model, while the available computational lineage estimates its first-order branch.

The comparison begins with the two demand operators and a finite economy in which they select different products. The released programs show where each operator enters. The purchaser-income calculation and matched profiles then indicate the empirical size of the difference, after which a common-operator implementation carries the same demand law into welfare. The appendices derive the price effects and support conditions and record the numerical scope.

2. Two demand operators

Let $y_i > 0$ be consumer income, $p_j > 0$ price, and q_{ij} the nonprice part of utility conditional on observed quality, random-coefficient draws, and the market-product term. The Type-I extreme-value shock is added separately. Normalizing by the outside option, the model printed in Petrin (2002) has deterministic index

$$v_{ij}^E = q_{ij} + \alpha_i \log\left(1 - \frac{p_j}{y_i}\right), \qquad j \in \mathcal{J}_i(p) = \{j : p_j < y_i\}. \tag{1}$$

The released program uses

$$v_{ij}^L = q_{ij} - \alpha_g \frac{p_j}{y_i}. \tag{2}$$

For the comparison below, set $\alpha_i = \alpha_g = \alpha > 0$ and apply both indices on the affordable set. Write $x = p/y$ and

$$R(x) = -\log(1 - x) - x = \sum_{k=2}^{\infty} \frac{x^k}{k}. \tag{3}$$

PROPOSITION 1 (Operator gap). *For $0 \le x \le r < 1$,*

$$\frac{x^2}{2} \le R(x) \le \frac{x^2}{2(1-r)}. \tag{4}$$

The linear index exceeds the logarithmic index by $\alpha R(p_j/y_i)$. The two deterministic rankings agree whenever every pairwise logarithmic-utility margin exceeds the corresponding difference in Taylor remainders.

PROOF. The series in (3) gives the lower bound. Since $1/k \le 1/2$ for $k \ge 2$,

$$R(x) \le \frac{x^2}{2} \sum_{m=0}^{\infty} x^m \le \frac{x^2}{2(1-r)}. \tag{5}$$

For products j and k, subtracting their two pairwise margins shows that the change equals $\alpha\{R(p_j/y_i) - R(p_k/y_i)\}$. The stated margin condition preserves its sign. □

The remainder varies across products and consumers. A market-product fixed effect cannot absorb it once purchaser demographics and product prices vary. The support also changes with prices because (1) assigns a finite index only on $p_j < y_i$.

3. A finite economy

Consider one consumer with income $y = 2$, coefficient $\alpha = 1$, and two products. The old product has $(p, q) = (0.2, 0)$ and the new product has $(p, q) = (1, 0.5)$. Both products lie in the affordable set. Under (1), their normalized deterministic indices are −0.1053605 and −0.1931472, so the old product ranks first. Under (2), the indices are −0.1 and 0, so the new product ranks first.

Apply one common Type-I extreme-value shock law to the outside option and both products. The resulting choice probabilities are

$$(P^E_0, P^E_{\text{old}}, P^E_{\text{new}}) = (0.3670586, 0.3303527, 0.3025887) \tag{6}$$

and

$$(P^L_0, P^L_{\text{old}}, P^L_{\text{new}}) = (0.3442533, 0.3114933, 0.3442533). \tag{7}$$

The expected-maximum-utility gain from adding the new product is 0.3603799 under (1) and 0.4219808 under (2). The corresponding exact ex ante compensating variation is 0.8243606, while the fixed-coefficient log-sum money metric is 0.8439615. The same income, prices, qualities, products, and shock law generate each comparison; the price operator produces the difference.

4. The published and released computation

The published choice probability evaluates $\log(y_i - p_l)$ for every available vehicle l. Its price derivative is

$$\frac{\partial u^E_{ij}}{\partial p_j} = -\frac{\alpha_i}{y_i - p_j}. \tag{8}$$

The released call path runs from `minivan.prg` through `del_markc` and `minivan.arc` to `contractmark` in `libminivan.c`. It sets `alphai=parm1[group]/income`, evaluates the exponential of quality minus `alphai*price`, and uses the same coefficient in the markup derivative. Its derivative is

$$\frac{\partial u^L_{ij}}{\partial p_j} = -\frac{\alpha_g}{y_i}. \tag{9}$$

Let $D_{rj} = \partial s_r / \partial p_j$ and let H_{rj} indicate common ownership. Both branches can use the multiproduct Bertrand condition

$$s(p) + \{D(p) \odot H\}'(p - c) = 0, \tag{10}$$

provided D comes from the same utility system as s. The operator choice therefore propagates from demand through marginal-cost recovery and counterfactual prices.

The released archive has programs for demand estimation and markups. Its alternative importance-sampling branches load or save a final draw object that is absent from the distributed files. The archive also lacks the welfare routine used for Tables 8 and 13. These omissions leave the exact integration and realized-shock compensation branch unspecified.

TABLE 1. Model components in the article and distributed programs

Object	Evidence	Implication
$\alpha_i \log(y_i - p_j)$	Article and April 2001 working paper	Affordable-set demand and derivative (8)
$-\alpha_g p_j / y_i$	Released estimation and markup programs	Linear demand and derivative (9)
Purchaser-income moment	Public product and demographic proxy	Common-share comparison in Table 2
Joint GMM and welfare state	Final draws and welfare routine absent from the archive	Coefficient, markup, equilibrium-price, and compensation tables require the original integration state

5. Purchaser income and matched profiles

The public reconstruction uses the product records and demographic integration proxy distributed with the current Petrin tutorial. For each demand branch, it fixes the nonlinear coefficients reported in Tables 4 and 5, recontracts market mean utilities to the observed shares, recovers marginal costs from (10), removes minivans, and solves the remaining-product price system. Both branches reproduce every baseline product share within 3.2×10^{-16}; central-difference and analytic derivatives agree within 2.2×10^{-12}; the largest accepted price-system residual is 1.6×10^{-10}.

The demand-side comparison uses the income distribution of predicted 1984 minivan purchasers. Table 2 reports the values. The linear reconstruction differs from the published model value by 0.040 thousand dollars. With published coefficients, the logarithmic branch lies 5.873 thousand above the CEX mean. Re-estimation absorbs part of the operator change and leaves a gap of 0.980 thousand above the CEX mean. The public objective pools thirteen

TABLE 2. Mean income of minivan purchasers, 1984

Object	Thousand 1982–84 dollars
Published model value	36.091
Published CEX value	39.476
Public linear reconstruction	36.051
Printed operator, published coefficients	45.349
Printed operator, matched seven-parameter profile	40.456

The public reconstructions use the current tutorial product and agent proxy. Every branch recontracts aggregate market shares. The matched profile estimates three income-group coefficients and four family-vehicle loadings against the same ten public CEX targets and weighting matrix.

markets and holds six random-coefficient scales fixed, so the matched profile supplies an operator comparison within that objective. A replacement for the article's joint demand-supply GMM estimate requires the original moment normalization, weighting updates, and final simulation state.

With the published coefficients, predicted low-, middle-, and high-income purchaser shares move from (21.66, 7.90, 70.44) percent under the linear branch to (0.52, 5.13, 94.35) percent under the logarithmic branch. This movement explains the purchaser-income result: aggregate vehicle shares remain fixed, while their allocation across the income distribution changes.

6. Demand and welfare after alignment

Three implementations correspond to three economic objects. The logarithmic model builds the affordable set $\mathcal{J}_i(p)$, evaluates (1), uses (8), and solves compensating income from the same logarithmic utility. A Taylor model builds the same affordable set, evaluates (2), and re-evaluates both outside utility and p/y when compensation changes income. A quasi-linear reconstruction treats α_g/y_i as a fixed demographic price coefficient and uses the associated fixed-coefficient log-sum money metric.

The first implementation matches the published utility statement. The second gives a finite first-order model whose approximation error is bounded by Proposition 1. The third matches the released demand operator and the PyBLP replication. Each implementation can use micro moments and (10); welfare inherits the selected utility law.

For Tables 4 and 6–13, an aligned rerun proceeds in four steps. It fixes one integration design and support rule, estimates demand under one operator, recovers costs and prices with that operator's derivative, and evaluates compensation under the same utility and common shock coupling. The public files reproduce the Table 2 demand moment and the matched profile. The published welfare integration and joint estimation state remain unavailable.

7. Literature scope

Conlon and Gortmaker (2025) estimate the Petrin micro-BLP specification with price interacted with inverse income and report consumer-surplus replications of 429.89 and 425.91 million beside the published 367.29 million. Their implementation and the PyBLP tutorial belong to the linear branch. Griffith, Nesheim, and O'Connell (2018) cite Petrin as an application of $\alpha \log(y - p)$ and develop a separate flexible-expenditure estimator. Their results use their own operator. Together, the article and released programs classify the implemented Petrin application as the linear branch. The wider literature using demographic micro moments can retain that method with a declared utility, derivative, support, and welfare map.

8. Conclusion

The article and the distributed programs implement two income-effect models. Their gap is $\alpha R(p/y)$, varies with the consumer-product pair, and enters demand, derivatives, markups, and welfare. A finite economy changes rankings and welfare, while the public minivan reconstruction changes the reported purchaser-income comparison after aggregate shares are held fixed. The available archive supports this demand-side result and the matched profile. Reproducing the coefficient, markup, price, profit, and compensating-variation tables requires the original integration state and a common operator throughout the calculation.

References

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–890.

Conlon, Christopher and Jeff Gortmaker. 2025. "Incorporating Micro Data into Differentiated Products Demand Estimation with PyBLP." *Journal of Econometrics* 249: 105926.

Griffith, Rachel, Lars Nesheim, and Martin O'Connell. 2018. "Income Effects and the Welfare Consequences of Tax in Differentiated Product Oligopoly." *Quantitative Economics* 9 (1): 305–341.

Petrin, Amil. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy* 110 (4): 705–729.

Appendix A. Probability derivatives and support

For conditional logit probabilities P_{ij} and a price derivative u'_{ij}, differentiation gives

$$\frac{\partial P_{ir}}{\partial p_j} = P_{ir}\{1(r = j) - P_{ij}\}u'_{ij}. \tag{A1}$$

Equation (A1) applies with $u'_{ij} = -\alpha_i/(y_i - p_j)$ in the logarithmic model and $u'_{ij} = -\alpha_g/y_i$ in the linear model. In the public proxy, 833,364 of 2,407,000 product-agent pairs have $y_i \leq p_j$. The affordable-set implementation assigns zero probability to these pairs. The count diagnoses support in the proxy integration file and carries no claim about the proprietary final draws.

Appendix B. Relation to the published welfare calculations

The fixed-coefficient public calculation produces five-year ex ante expected utility under both demand laws. Their difference totals −444.08 million in the proxy economy. These values use declared ex ante criteria. Table 8 conditions on realized minivan purchasers and includes realized logit errors, while Table 13's integration routine is absent from the archive. The proxy totals therefore compare the two operators within the released inputs. Table 2 uses choice probabilities directly and is independent of this welfare-branch difference.