Replication summary of *Nevo (2001): Measuring Market Power in the Ready-to-Eat Cereal Industry*[*]

Antto Jokelainen Adam Rybarczyk

[*]Antto Jokelainen, Aalto University & FCCA. Email: antto.jokelainen@aalto.fi. Adam Rybarczyk, University of Helsinki. Email: adam.rybarczyk@helsinki.fi.

1 Replication

In this document we summarize our replication of Measuring Market Power in the Ready-to-Eat Cereal Industry by Aviv Nevo (Nevo 2001). This paper has an existing replication package made by Aviv Nevo that is available at Eric Rasmusen's website[1]. The replication package includes a dataset although it is not the same dataset as in the original paper but rather a simulated sample. Unfortunately, as the original dataset is not publicly available, it is impossible to replicate the results from the original paper, even with the simulated dataset. To circumvent this, we estimate demand using the simulated dataset with both our model and a Python package called PyBLP (Conlon and Gortmaker 2020) and compare the results across the two.

PyBLP is an easy-to-use Python package for estimating BLP-style demand models[2]. In fact, this package uses this specific paper with the simulated dataset as an example for estimating a BLP-style demand model.[3] Consequently, we are using the results from this package as a benchmark for our own model. As we are more familiar with Python than with Matlab, we have also coded our own model using Python instead of Matlab, unlike the original replication package by Nevo.

Our replication has two files: demand_estimation.py and nevo_cereal_replication.py. The former includes a Python-class that implements a BLP-style demand estimator (akin to PyBLP) and the latter initializes the dataset, sets the parameters for the estimator and saves the outputs. The workflow of our replication is similar to Nevo's paper: define the demand system, estimate the demand parameters using a two-stage GMM estimator, estimate the marginal costs with a chose supply specification, and print out the results. One thing to note note is that we haven't included any new extensions or modifications of the original analysis to our replication. We felt that coding the original analysis from scratch with a new software was a task big enough in itself.

For our model and for the model estimate with PyBLP, we set the utility function to depend linearly on prices and product fixed-effects and non-linearly on constant, prices, sugar content and mushiness. As the product fixed effects exhibit perfect multicollinearity with sugar content and mushiness, they are not included in the linear part of the utility function. Nevo estimates these parameters by taking the estimated fixed effects and regressing these with the product characteristics in the post-estimation phase. We do not do this exercise here. Other specification choices that we made are that we simulate the integration of the non-linear part with Monte Carlo and using standard normal

1. https://www.rasmusen.org/zg604/lectures/blp/frontpage.htm
2. https://pyblp.readthedocs.io/en/stable/index.html
3. https://pyblp.readthedocs.io/en/stable/_notebooks/tutorial/nevo.html

distribution with 1000 draws and that use numerical approximation of the gradients of the GMM objective function instead of using analytical gradients. The results of the demand estimations are shown in Table 1. Although the table includes estimates from Nevo's paper, these are using different data and different set of variables, so we are going to focus on the estimates from PyBLP and our model.

Table 1: Demand model estimates

	Our model	PyBLP	Nevo
prices	-30.051	-30.049	-27.198
	(1.415)	(1.056)	(5.248)
RC: Intercept	0.000	0.063	0.330
	(23.688)	(1.937)	(0.609)
RC: prices	0.000	0.000	2.453
	(43.128)	(20.612)	(2.978)
RC: sugar	0.000	0.005	1.661
	(47.611)	(0.263)	(5.866)
RC: mushy	0.000	0.000	0.244
	(53.322)	(8.666)	(0.623)

RC refers to random coefficient. Standard errors are reported in parentheses. Product fixed effects are included but not reported in table. Nevo's specification also includes other variables not reported here.

Staring from the random coefficients at Table 1, one can see that the estimation of the random coefficients doesn't seem to work for either our model or PyBLP. In the PyBLP tutorial they use the same utility specification, although with a Monte Carlo size of 50 instead of 1000, and they get non-zero results for random coefficients. After experimenting with the model, it seems that these estimates vary significantly over random draws of heterogeneity. When increasing the number of draws to 1000, the estimates are approximately zero. Therefore, this is not the optimal dataset to use for replication. Since we didn't have time to simulate our own dataset, we decided to continue with this regardless of the issues.

Although we couldn't estimate any reasonable random coefficients, the estimation of price coefficient, alpha, seems to be quite consistent for both models. The small difference between can be explained by the two non-zero random coefficients for the PyBLP model. Alternatively, differences in convergence criteria, e.g. with respect to fixed-point iteration, can cause such differences in estimates.

Last thing to notice in the table is the standard errors. As we can see, the estimates are quite different across models. This would imply that there is probably some errors or differences in the calculation of Jacobians of the moment conditions with respect to parameters. It could be that the way we approximate partial derivatives is either too inaccurate or done erroneously. Especially for the random coefficients, we are not sure how to calculate the partial derivatives when the coefficient is at its bound at zero. Whatever the case, some difference in the moment Jacobians could explain the fact that we have a difference in standard errors but not really in point estimates. Also, the standard errors are different even when using one-stage GMM, so the issue is unlikely with the weighting matrices.

Then, the next step is to estimate the markups. For these we estimate the costs assuming that 1) all products are produced by different firms, 2) products belong to the firms that are reported in data, and 3) products are all owned by a monopolist. The results are reported in Table 2. Again, we also report the estimates from Nevo's paper although these are produced with different data and demand specification.

Table 2: Estimated markups

	Our model	PyBLP	Nevo
Independent	24.2%	27.4%	35.8%
Current ownership	28.8%	31.6%	42.2%
Monopolist	47.2%	52.8%	72.0%

Table shows median markups in percentages ($(p - mc)/p$). Independent refers to independent products with no cross-ownership, current ownership refers to the ownership structure defined in the data, and monopolist refers to ownership structure with only one producing firm.

From Table 2 we can see that the markup estimates behave as expected: as the cross-ownership increases, so do the estimated markups. Unfortunately as this is simulated data, it is no use to try to compare this to the ones reported in Nevo's paper. However, although the markups seem overall reasonable, there is still a difference in the markups between our model and the PyBLP. This difference can result from the fact that there are still some minor differences in the demand estimates, as seen in Table 1. Another, a more probable thing is that the issue of Jacobians persists here. When calculating the first-order conditions, we need to calculate demand Jacobians, that is the partial derivatives of market shares with respect to prices. As the way we calculate demand Jacobians follows the similar methodology as with the moment Jacobians, the same issues probably arise here as well. This could explain the differences, especially as the difference in markups seems somewhat in the same ballpark as the difference in the standard errors of price coefficients in Table 1 (the comparison of standard errors of random coefficients is not reasonable).

References

Conlon, Christopher and Jeff Gortmaker. 2020. 'Best practices for differentiated products demand estimation with PyBLP'. *The RAND Journal of Economics* 51 (4): 1108–1161. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.12352.

Nevo, Aviv. 2001. 'Measuring Market Power in the Ready-to-Eat Cereal Industry'. *Econometrica* 69 (2): 307–342.