Economic analysis of supermarket de-mergers – A code review

Date: 23 October 2025

Authors: Dr Stuart Donovan and Dr Leon Stirk-Wang

Summary

Motu Economic and Public Policy Research ("Motu Research") was engaged by Sense Partners to review the code that was used to undertake their economic analysis of supermarket de-mergers. This document outlines the results of our review. Based on our review, we find:

- The implementation is a faithful representation on the spatial competition model developed in the paper by Ellickson et al (2020)[1]; and
- There are no apparent substantive issues with implementation in terms of data inputs/outputs, parameter estimation, or modelling assumptions.

In the following sections, we summarise 1) our understanding of the model, 2) our review of its implementation, and 3) potential opportunities for further development of the model.

Our understanding of the model

Although this review focuses on the code used to analyse supermarket de-mergers, we first briefly summarise the underlying model to provide a basis for subsequent discussions.

The model developed in Ellickson et al (2020) begins by assuming utilities are defined at a store and geographic level via a combination of store attributes and tract-level consumer demographic characteristics. Formally, the contribution of a unit of expenditure i at a store s to the utility of a consumer in tract t is given by the following expression (cf. Equation 1):

$$u_{st} = \tau_0 d_{st} + \tau_1 d_{st} z_t + \gamma_0 x_s + \gamma_1 (x_s \otimes z_t) + \varepsilon_{sti}$$

Where d_{st} denotes distance, z_t denotes demographics, x_s denote store characteristics, and ε_{sti} denotes an idiosyncratic preference shock.

Heterogeneity enters this model via the interaction terms between tract-level demographics, z_t, with distance ($d_{st} z_t$) and store characteristics ($x_s \otimes z_t$). By allowing the effects of distance and store attributes to vary with tract-level demographic characteristics, such as income and family size, this model can accommodate relatively heterogeneous substitution patterns.

The inclusion of store characteristics, x_s, also enable the model to be used to analyse the competition effects of mergers and de-mergers. By assuming stores are nested within chains, idiosyncratic preferences, ε_{sti}, can be correlated within chains more so than between chains.

[1] Ellickson, P. B., Grieco, P. L., & Khvastunov, O. (2020). Measuring competition in spatial retail. The RAND Journal of Economics, 51(1), 189-232.

All other factors equal, we expect greater substitution *within chains* than *between chains* due to, for example, similarities in product range, pricing levels, and store environments.

By assuming preference shocks ε_{sti} are distributed GEV, Ellickson et al (2020) derive closed-form analytical expressions that describe parameters in the model, specifically:

- First, *the consumer expenditure level* defines the unconditional share, $p_{st}(\theta)$, of overall consumer expenditure at store, s, in tract, t, using a nested logit model (cf. Equation 3).
- Second, *the store revenues level* defines the total revenue per store, $R_s(\theta, \alpha)$, as the sum of revenue per store per tract, $\hat{R}_{st}(\theta, \alpha)$ (cf. Equation 4).[2]

Ellickson et al (2020) estimate the parameters in these two levels of the model using non-linear least squares, where they define a loss function that minimises the variance between predicted, $\hat{R}_s(\theta, \alpha)$, and observed, R_s, store revenues (cf. Equation 5).

To finish, Ellickson et al 2020 use the estimated parameter values to show how the model can be used to analyse the retail landscape, for example, in terms of *demographic effects*, such as income effects; *substitution patterns*, such as store-level diversion ratios; *market concentration*, both tract-level and aggregate; and *mergers or de-mergers*. Naturally, the latter application of the model is the most relevant to the following sections of this review.

In our view, the model of retail competition developed in Ellickson et al (2020) provides a relatively elegant, tractable, and flexible framework for analysing spatial competition. The model draws on and extends seminal papers in the competition literature, such as Berry et al (1994).[3] Other aspects of the model specification – specifically, the logit model of consumer expenditure, the inclusion of interactions between store-level and tract-level characteristics, and preferences with correlated structures – aligns with a large body of spatial economic literature that uses similar methods (sometimes called "sorting models") to analyse, for example, home and work location choice.[4] For these reasons, we consider the model in Ellickson et al (2020) to be an appropriate quantitative framework for the analysing the economic effects of mergers and de-mergers of supermarkets in New Zealand.

Reviewing the model's implementation

The supermarket competition analysis is implemented at the **SA2 level**, using census and geospatial data to capture local demographic and accessibility characteristics.
The modelling framework consists of three linked components:

- **Spatial competition model:** a model of consumer store choice implemented at the SA2 level, which relates the share of household grocery expenditure allocated across

[2] Revenue by store-tract, $\hat{R}_{st}(\theta, \alpha)$, is itself a function of the consumer expenditure share $p_{st}(\theta)$, the tract population, n_t; average income in a tract, inc_t; and share of groceries in consumer expenditure, α.
[3] Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841-890.
[4] See, for example, Mulalic, I., & Rouwendal, J. (2020). Does improving public transport decrease car ownership? Evidence from a residential sorting model for the Copenhagen metropolitan area. Regional Science and Urban Economics, 83, 103543.

competing stores based on relative accessibility and local demographic characteristics. The model captures systematic variation in store preferences between SA2s and quantifies the probability that consumers substitute between supermarket brands when either relative accessibility or store attributes change.

- **Demerger model:** a model that extends the spatial supermarket choice framework to analyse how changes in ownership structure affect retail prices, market shares, and welfare outcomes. Using the estimates of store level revenue elasticities derived from the spatial competition model, it then uses modified ownership matrices that represent the separation of supermarket groups into independent entities. The model uses the Almost Ideal Demand System (AIDS) solver[5] to estimate how changes in the ownership configurations would alter equilibrium shares, margins and revenues under the assumption of Nash–Bertrand competition.
- **Household welfare model:** a model that links the price and welfare effects simulated in the demerger model to household-level expenditure data. It applies the modelled supermarket price changes to household spending across household types and income quintiles to estimate changes in consumer welfare (equivalent or compensating variation). The model does not re-solve the market equilibrium but allocates the aggregate welfare impacts from the AIDS framework across households to show how gains and losses are distributed by income and household composition.

Spatial competition model

The model is implemented in R using SA2 data from the census and store-level attributes. Utility is specified as a linearly-additive function of distance, demographic interactions, and store characteristics. Due to correlations in idiosyncratic preferences within chains, choice probabilities are computed via a nested-logit function with the resulting tract-store spending shares aggregated to match observed revenue totals. The baseline calibration reproduces observed data for the North and South Islands in New Zealand, which subsequently serves as the reference case for all subsequent simulations.

Implementation

- Utility specification follows Ellickson's form, with correct treatment of distance, income interactions, and store characteristics.
- Choice probabilities computed via a nested-logit GEV structure; the "outside" good (non-grocery expenditure) is handled consistently.
- Aggregation and revenue matching routines reproduce observed sales totals.

Audit findings

- **Urban cutoff:** Code appears to apply a 40 min travel-time threshold although the comments cite a 30 min travel-time cutoff. Although the behaviour in the model is consistent, we suggest correcting the documentation.
- **Variable demeaning:** Demeaning occurs before filtering unavailable tracts, which gives rise to non-zero means for long-tailed variables (notably Pop_density). Although we do not expect this problem to affect the results, re-demeaning post-alignment would

[5] From the *antitrust* R package.

ensure variables are zero at their means, which may aid in the interpretation of interaction effects.

- **Distance matrix alignment:** Store IDs and SA2 codes match exactly; no missing entries after trimming. Code is implemented robustly with numerous safety checks for matrix alignment.
- **Choice probabilities:** Column sums for the N modelled stores generate broadly expected distributions of choice probabilities where the median share for unmodelled stores (>=N+1) is estimated to be approximately 9%. The latter number implies that unmodelled substitutes for supermarkets, such as local specialised grocery stores and cafes, account for the balance of expenditure. This seems appropriate.
- **Parameter replication:** Optimisation results are reproducible; mean relative difference in predicted revenues are non-zero but small and there is some minor variation in the estimated lambda coefficients. This may be due to the change in the urban cutoff parameter above or small changes to the tolerance parameters in the optimised nlminb. For example, a replication run required ~56 iterations to converge while the equivalent number reported is over 200.
- **Nesting and store classification:** mu values and store-type groupings behave as expected; substitution patterns appear plausible.
- **Convergence:** As above, the model converges rapidly with current tolerances; results numerically stable.
- **Reproducibility:** The use of saved environments and timestamped model outputs mean that the exact results in later scripts were reproducible despite the small variations in the estimated parameters that we derived here. We did not investigate the full effect of any flow on effects of these variations through the entire script pipeline. Instead we treated each script as separate and verified replication of the results given the inputs provided and checked that each script functioned as intended by verifying the results of our runs were identical to the saved (expected) outputs. In this way we established that later scripts did not introduce their own deviations that might compound.

Overall assessment: The spatial competition model is a faithful implementation of the Ellickson framework with reproducible outputs and internally consistent data handling. Issues identified are minor and relate primarily to centring order, documentation, and probability normalisation rather than substantive modelling errors.

Demerger model

The model is implemented in R using the *antitrust* package. It reads calibrated elasticities and revenue data from the spatial model, reconstructs ownership matrices to represent the demerger scenarios, and introduces proportional changes in marginal costs through cost-of-goods-sold (COGS) parameters for each owner. Each scenario is benchmarked against the same baseline, and dedicated sensitivity scripts vary the assumed COGS changes by ±3 percent to test robustness. All runs produce deterministic, reproducible outputs saved as .Rdata files.

Audit findings

- **Ownership configuration:** The ownership matrices in the unilateral Foodstuffs and Woolworths demerger scenarios are correctly implemented. In each case the ownership matrices are adjusted to reflect the assumptions laid out in Table 5 (Sense

Partners Report). Namely, Pak'nSave is demerged from the other Foodstuffs brands; and half of the Woolworths group stores that operate under the Woolworths banner are spun off into a NewCo banner. These demergers are benchmarked consistently against the same baseline.

- **Cost assumptions:** The only restructuring cost represented in the code is the proportional increase in COGS**.** A +1 percent increase is applied in the central case to reflect reduced efficiency and buyer power following demerger, with ±3 percent sensitivity tests implemented in the run_sensitivity_analysis() routines. Other restructuring (cost) assumptions discussed in the Sense Partners report (e.g. capital expenditure on new distribution centres, duplication of head-office or IT functions, etc) do not appear in the model code. They are instead applied in the subsequent spreadsheet summaries of the benefit-cost analysis that take the model's welfare results as input. Consequently, the quantitative outputs of the R model reflect only the effects of ownership separation and variable-cost changes on equilibrium prices and welfare.
- **Potential enhancements:** The Woolworths divestment logic assigns every second store (by ID) to NewCo, ensuring reproducibility but not reflecting potential geographic clustering in ownership. A robustness test that randomises or spatially balances the allocation of stores might better capture plausible divestment configurations and regional concentration effects.

Overall assessment: The demerger model is a faithful and technically consistent application of an AIDS-based equilibrium framework. Within the scope of the implemented assumptions the model performs as intended and produces reproducible, internally coherent results. The main limitations relate to the exclusion of fixed or restructuring costs and the simplified representation of ownership separation, rather than to any coding or conceptual errors.

Household welfare model

The model is implemented through the final sections of the demerger scripts, and in the summary scripts. It imports scenario results, merges them with household expenditure data (income, region, and spending weights), and computes equivalent or compensating variation (EV/CV) for representative household groups. Welfare changes are reported both in monetary terms and as a percentage of household income or expenditure. The code is deterministic and produces population-weighted national and regional summaries.

Audit findings

- **Scope and alignment:** The welfare calculations are correctly based on modelled price changes from the demerger simulations. The process maintains consistency between aggregate welfare results and household-level allocations.
- **Price vs. non-price effects:** The code measures welfare impacts arising **solely from price changes**. Although the Sense Partners report refers to "non-price offerings" or product-range effects, these are not implemented in the model. Store characteristics are fixed across scenarios, consistent with the report's later statement that no change in range is assumed.
- **Structure and robustness:** Calculations are purely arithmetic and transparent. Scripts execute without errors and results are reproducible. The logic for aggregating households by income quintile and region is sound and correctly weighted.

- **Potential enhancements:** The welfare scripts could be generalised to reduce reliance on globally defined variables and parameterised for portability. Consolidating repeated code blocks for merging and summarisation would improve maintainability.

Overall assessment: The household welfare model is a straightforward and reliable implementation that accurately translates simulated price effects into household-level welfare impacts. The scope is intentionally limited to price-driven effects, with no modelling of variety or non-price changes. Within this scope, the implementation is methodologically consistent, numerically stable, and well-aligned with the broader modelling framework.

Opportunities for further development

Parameter estimation

The non-linear, multi-level nature of the model may be suited to full Bayesian estimation. The latter presents an opportunity to relax some of the underlying assumptions in the models, such as the nature of measurement error. Adopting a Bayesian multi-level model for parameter estimation would also provide the opportunity to estimate a statistical model of the response variable (specifically, observed revenues at the store level), including its associated distributional parameters, rather than a simple loss function.

Style, robustness, and maintainability

In our view, the implementation of the models is generally excellent, and the codebase represents a significant quantity of work even before considering the compressed timeframe in which it appears to have been produced. In the interests of supporting style, robustness, and maintainability, however, we note some minor improvements that might be made.

- **Function generalisation and scope control:** Several scripts, particularly the household welfare and sensitivity routines, rely on globally defined variables (e.g. Scenario_name, Island, or file paths). Generalising these to accept explicit arguments would improve modularity and make reruns less environment-dependent.
- **Reduction of code replication:** Similar to the above the sensitivity analysis scripts replicated the functions for sensitivity analysis. Generalising these functions and consolidating them into a sourced utilities file (similar to the scenario helper functions) would reduce brittleness.
- **Idempotent variable assignments:** Ensuring that transformation blocks can be rerun safely without side-effects would simplify debugging and improve reproducibility when re-executing partial workflows.
- **Parameterisation over hard-coding:** A few scripts reference matrices or file paths directly (e.g. distance matrices in Add_Costco.R). Parameterising these references or sourcing them from a central configuration file would prevent breakage if file structures change.
- **Deprecation and syntax hygiene:** One instance of data_table should be replaced with either data.table or tibble. Path references could be standardised for portability.

Overall, the codebase is clean and maintainable and has no structural issues affecting results. The recommendations above are intended to enhance robustness, and support longer-term re-use of the modelling framework.

Limitations of this peer review

This review was conducted under tight time constraints, i.e. within 1-2 weeks. These time constraints limited our ability to investigate the effects of alternative implementations of the code, for example, estimating parameters using a multi-level Bayesian model.