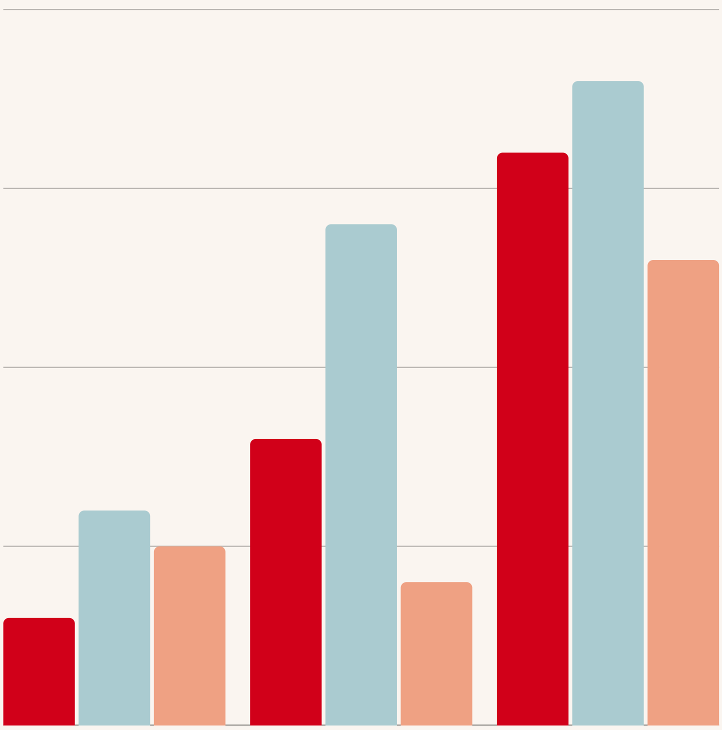
KONKURRANSE·
TILSYNET
Norwegian Competition Authority

Rapport 7/2025
Vedtak: 2023-5

PATTERNS OF SUBSTITUTION AMONG NEWSPAPERS IN NORWAY

Armando Garcia Pires, Frode Skjeret, and
Øyvind Thomassen



Prosjektet har mottatt midler fra Konkurransetilsynet.

Patterns of substitution among newspapers in Norway

Armando Garcia Pires*
Frode Skjeret†
Øyvind Thomassen‡

October 13, 2025

Contents

*armando.pires@snf.no, SNF
†frode.skjeret@snf.no, SNF
‡oyvind.thomassen@nhh.no, NHH Norwegian School of Economics.

1 Introduction

The Norwegian newspaper market is characterised by extensive "multihoming" - consumers typically read several newspapers in a given week. Unlike in markets where consumers typically choose only one product in a given decision period, such as the market for new cars, a demand model must therefore take into account the role of consuming multiple products simultaneously.

In this paper we estimate a structural demand model for newspapers and groups of newspapers. In order keep the choice set of a manageable size, we group newspapers, apart from the largest few, into categories such as 'other national', 'large regional', and 'local'. In our model, consumer can then choose any combination of these products.

Products interact in consumer utility through pairwise synergy parameters, that can be either positive, in case of complementarity in utility, or negative, in case of substitution in utility.

Heterogeneity in consumer tastes is modelled purely based on observable consumer traits. This reflects the fact that our data contain extensive information about consumers' stated preferences for different types of news, such as 'international', 'domestic', 'local', 'sport', 'entertainment', etc., as well as political preferences and usual demographics like age, sex, income, etc. This information, together with a frequency of reported reading (rather than yes/no) and a large sample, provides us with extensive information about the drivers of multihoming.

However, the absence of observed product characteristics, panel data, or second-choice data makes it challenging to identify unobserved heterogeneity in the taste for products, i.e. random coefficients. Still, we think direct information about preferences, both for types of news and politics, more than compensate for this challenge, and allows us to estimate a rich and realistic model of newspaper demand. We are able to make use of the BLP (Berry, Levinsohn, and Pakes, 1995) nested fixed point approach to estimating time-product specific utility constants, but adapt this to a maximum-likelihood (rather than GMM) estimator.

We are able to precisely estimate rich drivers of taste heterogeneity, as well as the pairwise synergy parameters. As expected, we find that newspaper pairs are overwhelmingly substitutes in utility, while consumer tastes may be correlated. Distinguishing between these too drivers of simultaneous consumption of pairs of newspapers is crucial to correctly predicting price responses to changes in market primitives.

Another challenge is that we have very limited information about the price (if any) consumers have paid to read the newspapers they say they have read. We therefore rely on external data on average prices and marginal cost to infer the price sensitivity that makes the observed prices the profit-maximizing ones.[1] This is similar to the approach used by Smith (2004) and Gentzkow (2007).

We plan to use the demand model to evaluate counterfactual experiments such as reducing the government support for newspapers, the VAT exemption for newspapers, or removing the public broadcaster's free digital written news channel NRK.no from consumers' choice set. [This still remains

[1] This part of the analysis has not been completed yet, and similarly the counterfactuals described below. However, the paper includes derivations of the relevant expressions for first-order conditions for profit maximization that will be used for this analysis.

to be done.]

Our model is most closely related to the newspaper demand model of Gentzkow (2007), but our much larger choice set creates new challenges. Our demand model is also closely related to that for supermarkets in Thomassen, Smith, Seiler, and Schiraldi (2017). Other related work on newspaper demand includes Argentesi and Filistrucchi (2007), Fan (2013), and Bhuller, Havnes, McCauley, and Mogstad (2024).

2 Data

In this section we describe the data sources used in the paper, and provide some descriptive and simple regression analysis of the data.

2.1 Readership data

Our main data source is the F&M (Consumer and media) survey from the consumer research firm Kantar, for the years 2020, 2021, and 2022. The survey is a repeated cross section or about 36 000 respondents.

For each newspaper in Norway, respondents are asked in how many of the last six (in case of paper editions) or seven (in case of digital editions) days they read the newspaper in question. They are asked separately about paper, mobile phone, tablet, and PC. For instance, when asked about the newspaper Verdens Gang (VG) an individual might respond with frequencies: VG paper 4/6, VG mobile 2/7, VG tablet 0/7, and VG PC 3/7.

In addition to almost 200 newspapers, the survey includes responses on digital-only channels such as NRK.no, TV2.no, ABC nyheter, Nettavisen, and major social media channels.

Importantly, the survey also contains rich information on demographic variables. These include age, sex, income, education, home municipality centrality and population, political views, political party vote, interest in different types of news.

In addition to party voted for in the last parliamentary election, the respondents give a Likert-scale agreement/disagreement answer to each of 22 different statements about politics, such as "A person's salary should be based on their individual effort", "We should reduce the use of social welfare", or "It is important to respect traditions".

Respondents also give a Likert-scale response about their interest for diffent types of news, such as international news, domestic news, local news, news debate and analysis, business news, sports news, culture news, celebrity news, interior/housing, and health/nutrition/relationship material.

2.2 Other data sources

From the Norwegian Media Authority (Medietilsynet), we have obtained data on newspaper revenues and cost, by year (2018-2022) and newspaper (ca. 230 newspapers). The data contain daily readership, advertising revenue, digital ad revenue, subscription revenue, other sales revenue, salary costs, cost of inputs purchased, other production and administrative cost, distribution costs (only 2022), and paper/print costs (only 2022)

We also use publicly available information about company ownership to construct a data set of final owners of all newspaper.

2.3 Descriptive statistics

Table 1 shows the distribution among the Kantar survey respondents of the number of newspapers read at least once. The most common outcome is three different newspapers, but two, four and five are also common. Only 2.4 percent of respondents have not read any newspapers during the last week. These numbers do not include digital-only alternatives, such as NRK.no or social media - only traditional newspapers.

	Freq.	Percent	Cum. perc.
0	857	2.4	2.4
1	3203	8.9	11.3
2	5800	16.1	27.4
3	7335	20.3	47.8
4	7239	20.1	67.9
5	6203	17.2	85.1
6	3809	10.5	95.7
7	1325	3.6	99.4
8	225	0.6	100
Total	35996	100	

Table 1: Number of newspapers read at least once last week.

In Table 2 we look at the distribution of the number of newspapers read at least six days during the last week. About 70 percent of respondents read 1, 2 or 3 newspapers every day, with 1 being the modal response, and 2 a close second.

	Freq.	Percent	Cum. perc.
0	6374	17.7	17.7
1	9728	27.0	44.7
2	8907	24.7	69.5
3	5991	16.6	86.1
4	3102	8.6	94.7
5	1403	3.9	98.6
6	400	1.1	99.8
7	84	0.2	99.98
8	7	0.02	100
Total	35996	100	

Table 2: Number of newspapers read at least 6 issues last week.

Table 3 shows the same type of distribution as in Table 1, but now paper and digital editions of the same newspaper are treated as distinct alternatives. Reading 5 or 6 different newspapers at least once per week is fairly common.

	Freq.	Percent	Cum. perc.
0	1002	2.8	2.8
1	2633	7.3	10.1
2	4225	11.7	21.8
3	5475	15.2	37.0
4	5989	16.6	53.7
5	5432	15.1	68.8
6	4396	12.2	81.0
7	3091	8.6	89.6
8	1817	5.0	94.6
9	980	2.7	97.3
10	539	1.5	98.8
11	247	0.7	99.5
12	108	0.3	99.8
13	48	0.13	99.96
14	11	0.03	99.99
15	3	0.008	100
Total	35996	100	

Table 3: Number of newspapers read at least once last week; paper and digital version of the same newspaper treated as distinct choices.

In Table 4 we still consider paper and digital editions as different alternatives, but now count only newspapers read at least six days in the last week.

	Freq.	Percent	Cum. perc.
0	6449	17.92	17.92
1	8603	23.90	41.82
2	8153	22.65	64.47
3	6025	16.74	81.20
4	3633	10.09	91.30
5	1830	5.084	96.38
6	851	2.364	98.74
7	308	0.856	99.60
8	95	0.264	99.86
9	31	0.0861	99.95
10	10	0.0278	99.98
11	3	0.00833	99.99
12	4	0.0111	100.00
14	1	0.00278	100
Total	35996	100	

Table 4: Number of newspapers read all issues or all issues except one last week; paper and digital version of the same newspaper treated as distinct choices.

In Table 5 we look at the number of respondents who report to have read specific newspapers (or groups of them, such as large regional newspapers) at least once during the last week. Note that the frequencies do not sum to 100 percent, since it is common (as shown in Tables 1-4) to read multiple newspapers. The largest newspaper, VG, is read at least once during the week by 72 percent of respondents.

variable	**sum**	**share of respondents (%)**
AftPpapir	6174	17.1
VGpapir	6317	17.5
DagBpapir	4873	13.5
DNpapir	2673	7.4
KlaKpapir	3167	8.8
riksanpapir	5718	15.8
regiopapir	7822	21.7
lokalpapir	20877	58.0
AftPdigit	14232	39.5
VGdigit	26136	72.6
DagBdigit	19690	54.7
DNdigit	6536	18.2
KlaKdigit	2585	7.2
riksandigit	16739	46.5
regiodigit	8903	24.7
lokaldigit	20699	57.5

Table 5: Readers of each newspaper: at least once in a week.

variable	sum	share of respondents (%)
AftPpapir	2844	7.9
VGpapir	753	2.1
DagBpapir	343	0.9
DNpapir	578	1.6
KlaKpapir	1405	3.9
riksanpapir	2481	6.8
regiopapir	3864	10.7
lokalpapir	11095	30.8
AftPdigit	5123	14.2
VGdigit	16777	46.6
DagBdigit	9222	25.6
DNdigit	1719	4.7
KlaKdigit	384	1.0
riksandigit	5477	15.2
regiodigit	5574	15.4
lokaldigit	11996	33.3

Table 6: Readers of each newspaper: all or all except one editions in a week.

Table 6 reports the same type of information as Table 5, but now requiring that the respondent has read the newspaper every day or every day except one during the last week. We see that VG is read every day by almost half of the respondents.

3 Regression results

Two features of the data are important for modelling demand in this market: 1) there is a large number of different products, and 2) consumers typically read multiple newspapers in a week.

This means that the total number of product bundles, i.e. combinations of newspapers, that are in fact chosen by consumers is very large.

In order to reduce the dimensions of the choice set to a point where it can be modelled realistically, we aggregate over products except for the largest few, to get the following eight newspaper products:

- *NRK.no* (free, no advertising, government funded) [owner: government]
- *Other digital only* (free and freemium digital) (TV2.no, ABC nyheter, Nettavisen,) [owner: other 1]
- *VG* (paper and freemium digital) [owner: Schibsted]
- *Dagbladet* (paper and freemium digital) [owner: Allergruppen]
- *Aftenposten* (paper and premium digital) [owner: Schibsted]
- *Other national* (DN, Klassekampen, ...) [owner: other 2]
- *Major regional newspapers* (BT, Adresseavisen, Stavanger Aftenblad, Fædrelandsvennen, Nordlys)

- *All other not national*

In the remainder of the paper these eight newspapers or groups of newspapers are our 'products', of which consumers can choose any combination.

In order to motivate our full demand model, presented in Section 4, we next do a preliminary regression analysis.

The aim of the paper is to understand how consumption of one newspaper affects demand for other newspapers. The most obvious approach to look at this question, is to regress consumption of newspaper A on the consumption of its competitors.

Concretely, we construct consumption measures for each respondent in our data set by taking the maximum across the frequencies for paper, mobile phone, tablet and PC. For instance, if a respondent reports frequencies VG paper 4/6, VG mobile 2/7, VG tablet 0/7, and VG PC 3/7, our VG consumption variable will be $4/6 = 2.67$ for this respondent. We do the same for each of the eight products (newspapers or newspaper groups) listed above.

Table 7 displays results from OLS regressions of these frequencies for four different "products": Vg, OtherNational, Regional, and Other, where the explanatory variables are the remaining seven products.

The most noticeable thing in this table is that most coefficients have a positive sign. If taken at face value as causal effects, this means, that higher consumption of one newspaper in many cases increases consumption of other newspaper.

For instance, in regression (1), where Vg is the dependent variable, Dagbladet has a coefficient of 0.452, which is precisely estimated. Understood as a causal effect, this means that if consumption of Dagbladet goes up by 10 percentage points for some exogenous reason (e.g. a reduction in price), then consumption of VG will increase by 4.5 percentage points.

	Dependent variable:			
	Vg	OtherNational	Regional	Other
	(1)	(2)	(3)	(4)
Vg		−0.024***	0.010*	0.044***
		(0.005)	(0.006)	(0.007)
NRK	0.025***	0.100***	0.091***	0.010
	(0.005)	(0.005)	(0.005)	(0.006)
Aftenposten	0.025***	0.262***	−0.163***	−0.044***
	(0.006)	(0.005)	(0.006)	(0.007)
Dagbladet	0.452***	0.070***	0.056***	0.041***
	(0.005)	(0.006)	(0.006)	(0.007)
OtherNational	−0.026***		0.047***	−0.006
	(0.006)		(0.006)	(0.007)
Regional	0.009*	0.038***		−0.088***
	(0.005)	(0.005)		(0.006)
Other	0.029***	−0.004	−0.066***	
	(0.004)	(0.004)	(0.005)	
OnlyDigit	0.155***	0.043***	0.051***	0.121***
	(0.005)	(0.005)	(0.006)	(0.007)
Constant	0.326***	0.084***	0.194***	0.443***
	(0.004)	(0.005)	(0.005)	(0.005)
Observations	32,870	32,870	32,870	32,870
R^2	0.270	0.129	0.041	0.025
Adjusted R^2	0.270	0.129	0.040	0.025
Residual Std. Error (df = 32862)	0.363	0.345	0.385	0.447
F Statistic (df = 7; 32862)	1,736.603***	694.029***	199.184***	120.505***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 7: Regress "probability of reading" e.g. VG, on probabilities of reading other newspapers

However, there are strong reasons to believe that these regressions do not provide consistent estimators of the causal effects of interest. There are two distinct mechanisms that we cannot hope to disentangle with this kind of analysis.

The first is that certain drivers of Dagbladet readership will also drive VG readership. We can call this a *correlation in tastes* - i.e. the things that make certain readers like VG may also make them like Dagbladet. This mechanism will create a correlation in reading of the two newspapers, but it is not a causal effect: if Dagbladet readership increases for an exogenous reason (e.g. a price drop) it will not increase VG readership at all, as long as correlation in tastes is the only reason we get a positive coefficient in regression (1).

The second mechanism is what we can call *complementarity or substitution in utility*: it could be that reading VG is either more valuable (complementarity) or less valuable (substitution) when one also reads Dagbladet, than if one reads VG on its own, without reading Dagbladet. This is a driver of a causal substitution or complementarity effect: if readership of Dagbladet increases for an exogenous reason, it will causally impact the readership of VG, either positively (if they are complements) or negatively (if they are substitutes).

A simple approach to disentangle correlation in taste from complementarity/substitution in utility is to include as control variables in the regression observable variables that are likely to be important components of tastes. Suppose we could control for all factors that determine tastes for VG consumption. Then the coefficient on Dagbladet would give us the partial correlation of VG and Dagbladet consumption *when holding tastes for VG fixed*, and could then be interpreted as an estimate of the substition or complementarity in utility.

In Table 8 we report results from regressions like those in Table 7 but with the difference that we have no included an extensive set of demographic control variables. These control variables fall into three main categories: preferences for different kinds of news, standard demographics, and political preferences.

Preferences for news include nine variables based on the Likert score for how strong the respondents' expressed interest in international news, domestic news, etc. is. To make the interpretation of coefficients easier, we have transformed the Likert-scores into z-scores, by subtracting the sample mean and dividing by the sample standard deviation. This means that the unit of the variable is one standard deviation.

The standard demographics are age, age squared, a dummy for female, household income, a dummy for whether the person in the household with the highest education has a university degree, a score from Statistics Norway that combines various measures of a municipality's centrality (from 0 to 1000 where Oslo scores 1000), and municipality population size.

The political preferences are the two variables zLrecon and zGaltan. These are also z-scores, so that the unit of measurement is a sample standard deviation. Here we have aggregated the Likert-score responses for political preferences into two dimensions, based on methods from the political science literature, like in Cincotta and Thomassen (2025). A high score for zLrecon indicates political preferences on the right of the classical political left-right spectrum, in the sense of a preference for less state involvement in the economy, less redistribution, etc. A high score for zGaltan indicates political preference on the right in terms of socially conservative, traditional, and nationalist preferences, while a low value is associated with support for same-sex marriage, immigration, etc. We aggregate political preferences in this way because it is a parsimonious way of capturing what the political science literature considers key dimensions of political preferences. For instance the party Venstre would be associated with a right-of-centre zLrecon score and a left-of-centre zGaltan score. On the other hand, the party Senterpartiet would be associated with a left-of-centre zLrecon score and a right-of-centre zGaltan score. Sosialistisk Venstreparty would be left-of-centre on both dimensions, while Fremskrittspartiet is right-of-centre on both.

Turning to the estimates in Table 8, the overall conclusion is that results to not change in qualitative way from the regressions without controls in Table 7. In particular, many coefficients on newspaper consumption are still positive. Some limitations of this model is that we need to run separate regressions for each newspaper, so that not all data is brought to bear at the same time; and that the functional form is somewhat restrictive, in that there is no interaction between demographics and newspapers, e.g. Dagbladet in regression (1).

	Dependent variable:			
	Vg	OtherNational	Regional	Other
	(1)	(2)	(3)	(4)
Vg		-0.012^{**}	0.017^{***}	0.070^{***}
		(0.005)	(0.006)	(0.007)
NRK	0.039^{***}	0.055^{***}	0.079^{***}	0.032^{***}
	(0.005)	(0.005)	(0.006)	(0.006)
Aftenposten	0.042^{***}	0.161^{***}	-0.140^{***}	0.035^{***}
	(0.006)	(0.005)	(0.006)	(0.007)
Dagbladet	0.426^{***}	0.058^{***}	0.047^{***}	0.027^{***}
	(0.005)	(0.005)	(0.006)	(0.007)
OtherNational	-0.014^{**}		0.041^{***}	0.020^{***}
	(0.006)		(0.006)	(0.007)
Regional	0.015^{***}	0.031^{***}		-0.111^{***}
	(0.005)	(0.005)		(0.006)
Other	0.049^{***}	0.012^{***}	-0.090^{***}	
	(0.005)	(0.004)	(0.005)	
OnlyDigit	0.118^{***}	0.056^{***}	0.043^{***}	0.056^{***}
	(0.005)	(0.005)	(0.006)	(0.006)
zNewsInternat	0.009^{***}	0.012^{***}	-0.003	-0.032^{***}
	(0.003)	(0.003)	(0.003)	(0.004)
zNewsDomestic	0.024^{***}	-0.007^{**}	-0.011^{***}	-0.005
	(0.003)	(0.003)	(0.004)	(0.004)
zNewsAnalysis	-0.015^{***}	0.032^{***}	0.003	-0.004
	(0.003)	(0.002)	(0.003)	(0.003)
zNewsBusiness	0.015^{***}	0.061^{***}	0.008^{***}	-0.001
	(0.003)	(0.002)	(0.003)	(0.003)
zNewsCulture	-0.029^{***}	0.012^{***}	-0.007^{***}	-0.003
	(0.002)	(0.002)	(0.003)	(0.003)
zNewsLocal	-0.004	-0.018^{***}	0.048^{***}	0.076^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
zNewsSport	0.055^{***}	-0.018^{***}	0.006^{***}	0.012^{***}
	(0.002)	(0.002)	(0.002)	(0.003)
zNewsCeleb	0.027^{***}	-0.017^{***}	0.00000	0.008^{***}
	(0.002)	(0.002)	(0.002)	(0.003)
zNewsEntertain	-0.001	-0.014^{***}	-0.005^{**}	-0.006^{**}
	(0.002)	(0.002)	(0.002)	(0.003)
age	0.004^{***}	-0.002^{***}	-0.00001	0.007^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
ageSq	-0.077^{***}	0.032^{***}	0.009	-0.028^{***}
	(0.008)	(0.007)	(0.008)	(0.009)
female	-0.019^{***}	-0.074^{***}	0.011^{**}	0.009^{*}
	(0.005)	(0.004)	(0.005)	(0.006)
hhIncomeMill	0.023^{***}	0.027^{***}	0.014^{***}	0.019^{***}
	(0.003)	(0.003)	(0.004)	(0.004)
hhEducUniv	-0.008^{*}	0.043^{***}	0.026^{***}	-0.012^{**}
	(0.004)	(0.004)	(0.005)	(0.005)
municipCentr	0.014^{***}	0.022^{***}	-0.089^{***}	-0.042^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
municipSize	-0.001	-0.005^{**}	0.077^{***}	-0.036^{***}
	(0.003)	(0.002)	(0.003)	(0.003)
zLrecon	0.017^{***}	-0.008^{***}	0.004	-0.003
	(0.002)	(0.002)	(0.002)	(0.003)
zGaltan	0.009^{***}	-0.018^{***}	-0.015^{***}	0.003
	(0.002)	(0.002)	(0.002)	(0.003)
Constant	0.286^{***}	0.064^{***}	0.126^{***}	0.466^{***}
	(0.022)	(0.021)	(0.024)	(0.026)
Observations	32,870	32,870	32,870	32,870
R^2	0.330	0.221	0.083	0.145
Adjusted R^2	0.329	0.221	0.082	0.144
Residual Std. Error (df = 32844)	0.348	0.327	0.377	0.418
F Statistic (df = 25; 32844)	645.833^{***}	373.390^{***}	118.232^{***}	222.839^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8: Regress "probability of reading" e.g. VG, on probabilities of reading other newspapers, plus controls

4 Structural demand model and estimation

We now turn to the structural model that is designed to make use of all the information in the data by incorporating the consumer's full choice of product bundles. We start by setting out a small toy version of the model, before turning to the full model.

4.1 Toy model

To build intuition, we start with a simple model with two newspaper products $k = 1, 2$, where consumers make a discrete choice between the bundles $\emptyset$, $\{1\}$, $\{2\}$, and $\{1, 2\}$, denoted respectively by $j = 0, 1, 2, 3$. A price p_k and a gross benefit δ_{ik} is associated with each product, so that the consumption of k in itself gives the consumer a net payoff

$$v_{ik} = \delta_{ik} - \alpha p_k$$

where α is a price sensitivity parameter. Indirect utility from the consumption of each bundle is

$$\begin{aligned} u_{i0} &= \varepsilon_{i0} \\ u_{i1} &= v_{i1} + \varepsilon_{i1} \\ u_{i2} &= v_{i2} + \varepsilon_{i2} \\ u_{i3} &= v_{i1} + v_{i2} + \Gamma + \varepsilon_{i3}, \end{aligned}$$

where ε_{ij} are standard Gumbel (type I extreme value) distributed independent random shocks, and Γ is a synergy parameter which is negative if the two products are substitutes and positive if they are complements.

The probabilities of i choosing bundles $j = 1$ and bundle $j = 3$ are, respectively,

$$\begin{aligned} P_{i1} &= \frac{\exp[v_{i1}]}{1 + \exp[v_{i1}] + \exp[v_{i2}] + \exp[v_{i1} + v_{i2} + \Gamma]} \\ P_{i3} &= \frac{\exp[v_{i1} + v_{i2} + \Gamma]}{1 + \exp[v_{i1}] + \exp[v_{i2}] + \exp[v_{i1} + v_{i2} + \Gamma]} \\ &= P_{i1} \exp(v_{i2}) \exp(\Gamma). \end{aligned}$$

4.2 Full demand model

Each consumer i is observed in one of T years t, written as $t(i)$. The set of consumers observed in t is denoted by I_t, the number of elements (consumers) in the set I_t is n_t, and the total number of consumers is $n = \sum_{t=1}^{T} n_t$. Let newspaper products be indexed by k, where the available products are $1, \ldots, K$. After aggregating over smaller products, we get $K = 8$. Consumers choose a weekly bundle of newspaper products that can contain from zero to all eight products. This gives a total number of $J = 2^K$ possible bundles, including the one that contains no newspaper products. We drop some combinations which are chosen by very few consumers, and end up with $J = 201$ alternatives in the choice set. If newspaper k belongs to combination j, we write $k \in j$.[2]

Conditional on choosing bundle j, consumer i gets indirect utility

$$u_{ij} = \sum_{k \in j} v_{ik} + \Gamma_j + \xi_{t(i)j} + \varepsilon_{ij} \tag{1}$$

where ξ_{tj} is a bundle-specific unobserved effect and ε_{ij} is i.i.d. standard Gumbel (type I extreme value). The utility component v_{ik} associated with each product j is given by

$$v_{ik} = \bar{\xi}_{t(i)k} - \alpha_k p_{t(i)k} + z_i \beta_k, \tag{2}$$

where $\bar{\xi}_{tk}$ is a time- and product-specific constant term, p_{tj} is the average price of reading k, α_k is a scalar price sensitivity parameter, β_k is a vector of k-specific parameters, z_i a vector of consumer attributes, which importantly includes i's stated interest in different types of news, and political preferences, as well as more standard demographics like age, sex, income, education, and centrality of home municipality. It will be convenient to define

$$\bar{\delta}_{tj} = \bar{\xi}_{tk} - \alpha_k p_{tk}, \tag{3}$$

so that

$$v_{ik} = \bar{\delta}_{t(i)k} + z_i \beta_k. \tag{4}$$

The time-invariant bundle-specific constant Γ_j is given by

$$\Gamma_j = \sum_{k \in j} \sum_{k' \in j, k' > k} \gamma_{kk'} + (\#j)\gamma_1^c + (\#j)^2 \gamma_2^c.$$

Here the parameter $\gamma_{kk'}$ captures the complementarity (if positive) or substitutability (if negative) in use between products k and k', while the terms involving the cardinality of the set j, $\#j$, determine the scale effects of reading multiple newspapers (e.g. diminishing marginal utility or inconvenience of reading a large number of papers). For the outside option $j = 0$ of reading no newspapers, we normalize utility to zero plus the standard Gumbel error term: $u_{it0} = \varepsilon_{it0}$.

For each (j, t) we gather the utility terms that do not vary across consumers i as

$$\delta_{tj} = \sum_{k \in j} \bar{\delta}_{tk} + \sum_{k \in j} \sum_{k' \in j, k' > k} \gamma_{kk'} + (\#j)\gamma_1^c + (\#j)^2 \gamma_2^c + \xi_{tj}. \tag{5}$$

[2]If consumers can choose up to $l \leq K$ different products, the number of possible bundles is $\prod_{b=0}^{l} \binom{K}{b}$.

We can then rewrite (2) as

$$u_{ij} = \sum_{k\in j} z_i\beta_k + \delta_{t(i)j} + \varepsilon_{ij}. \tag{6}$$

The probability of i choosing bundle j is now

$$P_{ij}(\beta,\delta_{t(i)}) = \frac{\exp[\sum_{k\in j} z_i\beta_k + \delta_{t(i)j}]}{1+\sum_{j'=1}^{J}\exp[\sum_{k\in j'} z_i\beta_k + \delta_{t(i)j'}]}$$

where β stacks β^k for all k and δ_t stacks δ_{tj} for all j. We use the fact Berry (1994) that for each β, there is a unique $J\times 1$ vector δ_t such that

$$s_{tj} = P_j(\beta,\delta_t), \quad \text{for } j=1,\dots,J \tag{7}$$

where

$$s_{jt} = \frac{1}{n_t}\sum_{i\in I_t} s_{ij},$$

is the observed market share of bundle j, with

$$s_{ij} = \text{Observed probability that } i \text{ choose } j \text{ on a given day},$$

and

$$P_j(\beta,\delta_t) = \frac{1}{n_t}\sum_{i\in I_t} P_{ij}(\beta,\delta_t) \tag{8}$$

is the corresponding predicted market share. For a given $J\times 1$ vector s_t (stacking the s_{tj}) the system of equations (7) defines δ_t as an implicit function of β:

$$\delta_t(\beta) = \text{the value of } \delta_t \text{ that solves (7)}. \tag{9}$$

Consumer i's likelihood of making his or her observed choice is then

$$L_i(\beta) = \prod_{j=1}^{J}\left\{P_{ij}[\beta,\delta_{t(i)}(\beta)]\right\}^{s_{ij}}, \tag{10}$$

and the log likelihood

$$\log L_i(\beta) = \sum_{j=1}^{J} s_{ij}\log P_{ij}[\beta,\delta_{t(i)}(\beta)].$$

Our maximum likelihood estimator of β is then

$$\hat{\beta} \quad = \quad \arg\min_{\beta} \left\{ -\sum_{i=1}^{n} \log L_i(\beta) \right\}.$$

For each trial value of β during the numerical minimization of the objective function, we compute $\delta_t(\beta)$ by iterating until convergence on the following contraction Berry, Levinsohn, and Pakes (1995):

$$\delta_t^{r+1} \quad = \quad \delta_t^r + \log s_t - \log \{P(\beta, \delta_t^r)\}$$

where $P(\beta, \delta_t)$ is the $J \times 1$ vector of choice probabilities that stacks (8).

Let $\hat{\delta}_t = \delta_t(\hat{\beta})$. In a second step, we estimate the parameters in (5) by regressing $\hat{\delta}_{tj}$ on dummies for year-product and product pair, as well as $(\#j)$ and $(\#j)^2$.

4.3 Newspaper profit maximization

Advertisers at newspaper k in year t are willing to pay a_{tk} kroner per daily reader. We assume that a_{tk} is exogenously given by features of the advertising demand side; newspapers are price takers in the advertising market. The average price per daily reader is p_{tk}, and the marginal cost (printing, paper, and distribution costs) per reader is c_{tk}.

From (3) and (5) we see that δ_t is a function of the $K \times 1$ vector of prices p_t for each product k. Accordingly we write

$$\delta_{tj}(p_t) \quad = \quad \sum_{k \in j} [\hat{\bar{\xi}}_{tk} - \alpha_k p_{tk}] + \sum_{k \in j} \sum_{k' \in j, k' > k} \hat{\gamma}_{kk'} + (\#j)\hat{\gamma}_1^c + (\#j)^2 \hat{\gamma}_2^c + \hat{\xi}_{tj},$$

where the estimates are from the second-stage regression of $\hat{\delta}_{tj}$ on dummies for year-product and product pair, as well as $(\#j)$ and $(\#j)^2$.

The predicted number of readers of bundle j is

$$Q_{tj}(p_t) \quad = \quad \sum_{i \in I_t} P_{ij}[\hat{\beta}, \delta_t(p_t)].$$

This number is based on the set of respondents I_t in the data set. To generate nationwide reader predictions, the number can be scaled up by the inverse of the share of the population that was sampled in the survey. For a newspaper k the predicted number of readers is:

$$Q_{tk}(p_t) \quad = \quad \sum_{j : k \in j} Q_{tj}(p_t).$$

The profit in year t of a corporation f that owns a subset K_f of the K products is

$$\sum_{k \in K_f} (a_{tk} + p_{tk} - c_{tk}) Q_{tk}(p_t).$$

The first-order conditions for profit maximization with respect to the p_{tk} are given by the system of K equations,

$$0 \;=\; Q_{tk}(p_t) + \sum_{k' \in K_{f(k)}} (a_{tk'} + p_{tk'} - c_{tk'}) \frac{\partial Q_{tk'}(p_t)}{\partial p_{tk}}, \quad \text{for } k = 1, \ldots, K, \tag{11}$$

where $f(k)$ denotes the firm that owns product k.

Writing $P_{ij} = P_{ij}[\hat{\beta}, \delta_t(p_t)]$, we have

$$\begin{aligned}
\frac{\partial Q_{tk'}(p_t)}{\partial p_{tk}} &= \sum_{j:k' \in j} \sum_{i \in I_t} \frac{\partial}{\partial p_{tk}} P_{ij} \\
&= \sum_{j:k' \in j} \sum_{i \in I_t} \frac{\partial}{\partial p_{tk}} \frac{\exp[\sum_{k'' \in j} z_i \hat{\beta}_{k''} + \delta_{tj}(p_t)]}{1 + \sum_{j'=1}^{J} \exp[\sum_{k'' \in j'} z_i \hat{\beta}_{k''} + \delta_{tj'}(p_t)]} \\
&= \sum_{j:k' \in j} \sum_{i \in I_t} \left[P_{ij}(-\alpha_k) 1[k \in j] - P_{ij} \sum_{j':k \in j'} P_{ij'}(-\alpha_k) \right] \\
&= -\alpha_k \sum_{j:k' \in j} \sum_{i \in I_t} P_{ij} \left[1[k \in j] - \sum_{j':k \in j'} P_{ij'} \right]
\end{aligned}$$

Using this expression for $Q_{tk}(p_t)$ in (11) and rearranging, we then have:

$$\alpha_k \;=\; Q_{tk}(p_t) \Big/ \sum_{k' \in K_{f(k)}} (a_{tk'} + p_{tk'} - c_{tk'}) \sum_{j:k' \in j} \sum_{i \in I_t} P_{ij} \big[1[k \in j] - \sum_{j':k \in j'} P_{ij'} \big], \quad \text{for } k = 1, \ldots, K,$$

Here σ_{jt}, a_{jt}, and c_{jt} are observed, and

$$\begin{aligned}\frac{\partial Q_{j't}}{\partial p_{jt}} &= \sum_{r:j'\in r} \frac{\partial Q_{rt}}{\partial p_{jt}} \\ &= \frac{1}{x_t} \sum_{r:j'\in r} \sum_{i\in I_t} \frac{\partial}{\partial p_{jt}} P_{itr} \\ &= \frac{1}{x_t} \sum_{r:j'\in r} \sum_{i\in I_t} (-\alpha_k) \frac{\partial}{\partial \bar{\delta}_{jt}} P_{itr},\end{aligned}$$

where

$$\begin{aligned}\frac{\partial}{\partial \bar{\delta}_{jt}} P_{itr} &= \frac{\partial}{\partial \bar{\delta}_{jt}} \frac{\exp[\sum_{j'\in r} v_{itj'} + \Gamma_r]}{\exp[z_i\beta_0] + \sum_{r'=1}^{R} \exp[\sum_{j'\in r'} v_{itj'} + \Gamma_{r'}]} \\ &= \frac{\exp[\sum_{j'\in r} v_{itj'} + \Gamma_r] \cdot \mathbb{1}[j\in r]}{\exp[z_i\beta_0] + \sum_{r'=1}^{R} \exp[\sum_{j'\in r'} v_{itj'} + \Gamma_{r'}]} \\ &\quad - \frac{\exp[\sum_{j'\in r} v_{itj'} + \Gamma_r]}{\left\{\exp[z_i\beta_0] + \sum_{r'=1}^{R} \exp[\sum_{j'\in r'} v_{itj'} + \Gamma_{r'}]\right\}^2} \sum_{r'=1}^{R} \exp[\textstyle\sum_{j'\in r'} v_{itj'} + \Gamma_{r'}] \cdot \mathbb{1}[j\in r'] \\ &= P_{itr} \cdot \left[\mathbb{1}[j\in r] - \sum_{r'=1}^{R} \left(P_{itr'} \cdot \mathbb{1}[j\in r']\right)\right]\end{aligned}$$

5 Estimates

In this section we present estimates from the full structural model. Table 9 reports estimates of the parameters β_k in (4), that is, the parameters that determine how observable demographics, political preferences, and preferences for news influence the taste for each newspaper product k. These are the parameters that determine *correlation in tastes* i.e. the similarity of tastes between different newspaper products. Table 10 reports estimates from (5), i.e. the year-newspaper effects δ_{tk}, the newspaper-pair effects $\gamma_{kk'}$, and the bundle size effects γ_1^c and γ_2^c. Here the newspaper-pair effects determine the *complementarity or substition in utility*, which are the drivers of substitution patterns.

Considering first the estimates of the parameters that govern taste heterogeneity, in Table 9, we can for instance look at the role of political preferences. The variable zLrecon, where a high value means being against redistribution and state involvement in the economy. This variable has positive, and precisely estimated, coefficients on both Aftenposten and VG, which are traditionally rightwing newspapers, while it is negative for Dagbladet, which is traditionally a leftwing newspaper. NRK.no has an even larger negative coefficient, presumably reflecting the social-democrat profile of the public broadcaster.

Turning to more standard demographics, we see that a university degree (hhEducUniv) reduces the probability of reading local newspapers (the 'Other' category) and particularly the digital-only written news outlets, while it strongly increases the taste for other national newspapers and NRK.no.

As for news preferences, the estimates overall look sensible. A higher preference for local news, zNewsLocal, strongly increases the taste for both the large regional (Regional) and local (Other) categories. A taste for news analysis (zNewsAnalysis) has a strong positive coefficient for OtherNational, which includes the outlets (e.g. Klassekampen, Dagens Næringsliv) characterized by more long-form journalism.

Table 9: Estimates from logit model: interactions between demographics and products.

	Aftenposten	Dagbladet	Vg	OtherNational	Regional	Other	NRK	OnlyDigit	InsideGoods
age	0.01	0	-0.027	0.007	0.003	0.017	-0.018	0.01	0.024
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)
female	0.177	-0.114	-0.126	-0.504	0.076	0.06	0.024	-0.126	0.384
	(0.034)	(0.032)	(0.034)	(0.034)	(0.034)	(0.029)	(0.031)	(0.03)	(0.235)
hhIncomeMill	0.062	0.053	0.161	0.082	0.087	0.127	0.026	-0.042	0.104
	(0.028)	(0.026)	(0.036)	(0.027)	(0.026)	(0.027)	(0.027)	(0.026)	(0.439)
indIncomeMill	0.031	0.037	0.069	0.095	-0.023	-0.021	0.06	-0.053	0.81
	(0.038)	(0.036)	(0.052)	(0.036)	(0.034)	(0.034)	(0.04)	(0.035)	(0.792)
hhEducUniv	0.48	0.042	-0.037	0.299	0.126	-0.079	0.262	-0.227	-0.004
	(0.032)	(0.03)	(0.032)	(0.033)	(0.032)	(0.028)	(0.028)	(0.028)	(0.215)
zMunicipSize	-0.187	0.063	-0.002	-0.085	0.745	-0.244	0.043	-0.017	-0.091
	(0.029)	(0.027)	(0.028)	(0.028)	(0.029)	(0.025)	(0.025)	(0.025)	(0.202)
zMunicipCentr	0.688	-0.106	0.06	0.233	-0.799	-0.313	-0.06	-0.077	0.364
	(0.029)	(0.027)	(0.028)	(0.029)	(0.027)	(0.025)	(0.026)	(0.025)	(0.199)
zLrecon	0.138	-0.118	0.104	-0.051	0.023	0.001	-0.212	0.072	-0.104
	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.014)	(0.015)	(0.014)	(0.117)
zGaltan	-0.114	-0.033	0.088	-0.135	-0.086	0.038	-0.163	0.251	-0.081
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.126)
zNewsInternat	0.151	0.059	0.035	0.116	-0.029	-0.18	0.084	-0.01	0.014
	(0.024)	(0.021)	(0.022)	(0.025)	(0.021)	(0.02)	(0.02)	(0.019)	(0.137)
zNewsDomestic	0.102	0.004	0.131	-0.048	-0.095	-0.034	0.11	0.016	0.157
	(0.026)	(0.023)	(0.025)	(0.027)	(0.025)	(0.022)	(0.022)	(0.021)	(0.146)
zNewsAnalysis	0.197	0.047	-0.081	0.246	0.028	-0.021	0.198	-0.042	-0.053
	(0.02)	(0.018)	(0.019)	(0.021)	(0.019)	(0.017)	(0.017)	(0.017)	(0.141)
zNewsBusiness	0.11	-0.037	0.1	0.464	0.055	-0.012	-0.012	0.059	-0.081
	(0.019)	(0.017)	(0.018)	(0.02)	(0.018)	(0.016)	(0.017)	(0.016)	(0.12)
zNewsCulture	0.195	0.011	-0.17	0.092	-0.048	-0.015	0.108	-0.128	0.001
	(0.018)	(0.017)	(0.018)	(0.018)	(0.018)	(0.016)	(0.016)	(0.016)	(0.125)
zNewsLocal	-0.244	0.026	-0.03	-0.134	0.355	0.392	-0.032	0.121	-0.115
	(0.02)	(0.019)	(0.02)	(0.02)	(0.022)	(0.018)	(0.019)	(0.018)	(0.137)
zNewsSport	-0.109	-0.063	0.337	-0.103	0.041	0.057	-0.052	0.145	-0.02
	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.014)	(0.014)	(0.014)	(0.12)
zNewsCeleb	-0.082	0.011	0.169	-0.1	0.006	0.049	-0.101	0.141	0.001
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.125)
zNewsEntertain	-0.044	0.012	0.009	-0.096	-0.051	-0.034	-0.064	0.022	-0.01
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.112)

Notes:

Table 10 reports estimates from (5), i.e. the year-newspaper effects δ_{tk}, the newspaper-pair effects $\gamma_{kk'}$, and the bundle size effects γ_1^c and γ_2^c.

As expected, the interaction parameters (in the right-hand side of the table) are negative, with only two exceptions. This means that, everything else equal, the utility of reading one newspaper falls in the frequencey with which one reads the other newspaper. For instance, local newspaper (Other) has a strong negative interaction, -1.186, and precisely estimated, with large regional (Regional) - presumably because they both cover local news, so that consumers choose one or the other to cover this need.

The coefficient on the number of elements in the bundle ($\#j$) is negative, but with a positive coefficient on its square. This means that everything else equal (including the utility one gets from each product in the bundle) consumers experience a cost associated with a larger bundle size, presumably in terms of time in particular.

Table 10: Estimates from logit model: linear parameters

	y2020	y2021	y2022	Dagbladet	Vg	OtherNational	Regional	Other	NRK	OnlyDigit	numProd
Aftenposten	-1.649	-1.632	-1.76	-0.154	-0.347	-0.083	-1.494	-0.523	-0.187	-0.604	
	(0.106)	(0.11)	(0.109)	(0.083)	(0.092)	(0.089)	(0.087)	(0.087)	(0.082)	(0.086)	
Dagbladet	-1.278	-1.379	-1.448		1.473	-0.477	-0.572	-0.65	-0.443	0.051	
	(0.104)	(0.117)	(0.119)		(0.092)	(0.084)	(0.086)	(0.083)	(0.079)	(0.076)	
Vg	1.999	2.169	2.106			-0.616	-0.435	-0.489	-0.685	-0.008	
	(0.104)	(0.112)	(0.114)			(0.095)	(0.098)	(0.092)	(0.095)	(0.078)	
OtherNational	-1.175	-1.121	-1.137				-0.574	-0.553	-0.386	-0.46	
	(0.122)	(0.12)	(0.127)				(0.09)	(0.089)	(0.083)	(0.085)	
Regional	-0.364	-0.263	-0.379					-1.186	-0.374	-0.553	
	(0.114)	(0.106)	(0.108)					(0.086)	(0.087)	(0.085)	
Other	0.222	0.179	0.155						-0.547	-0.511	
	(0.106)	(0.106)	(0.112)						(0.086)	(0.078)	
NRK	1.489	2.009	1.925								
	(0.119)	(0.119)	(0.122)								
OnlyDigit	-0.622	-0.567	-0.391								
	(0.1)	(0.098)	(0.103)								
$\#j$											-2.758
											(0.031)
$(\#j)^2$											0.471
											(0.029)

Notes: regression of δ_{tj} on product-year and product-pair dummies, as well as number of products and its square.

6 Conclusion

We estimate a rich model of the demand for newspaper, that fully takes into account the consumption of multiple newspapers, and the complex drivers of this observed behaviour, in terms of taste heterogeneity and pairwise interactions of newspaper products in utility. We exploit the benefits of rich information on news preferences and political preferences in our data, to estimate a rich, but yet tractable, demand model. We are able to precisely estimate both parameters governing heterogeneity in tastes for different products, and those governing the interactions between products in utility. Overall, this gives a fairly complete view of the drivers of substitution between newspapers in Norway. More work remains within the scope of the paper to analyse and quantify substitution in more detail, as well as for applying the framework to study counterfactual experiments of interest.

The framework can clearly be used for different kinds of aggregation of products into categories, all depending on the question one wishes to analyse. For instance, in a competition investigation where a merger between two newspaper is in question, one could separate out these two as categories on their own, and run the same analysis as we have done, to exactly quantify substitution between the relevant products or product groups.

References

Argentesi, E., and L. Filistrucchi (2007): "Estimating market power in a two-sided market: The case of newspapers," *Journal of Applied Econometrics*, 22(7), 1247–1266.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Berry, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 25(2), 242–262.

Bhuller, M., T. Havnes, J. McCauley, and M. Mogstad (2024): "How the Internet Changed the Market for Print Media," *American Economic Journal: Applied Economics*, 16(2), 318–58.

Cincotta, C., and Ø. Thomassen (2025): "Electric vehicle ownership and political preferences in Norway," *Transportation Research Part D: Transport and Environment*, 139, 104518.

Fan, Y. (2013): "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *American Economic Review*, 103(5), 1598–1628.

Gentzkow, M. (2007): "Valuing New Goods in a Model with Complementarity: Online Newspapers," *American Economic Review*, 97(3), 713–744.

Smith, H. (2004): "Supermarket Choice and Supermarket Competition in Market Equilibrium," *The Review of Economic Studies*, 71(1), 235–263.

Thomassen, Ø., H. Smith, S. Seiler, and P. Schiraldi (2017): "Multi-category Competition and Market Power: A Model of Supermarket Pricing," *American Economic Review*, 107(8), 2308–51.