CONSUMER AND MARKET DEMAND

AGRICULTURAL POLICY RESEARCH NETWORK

Latent consideration sets in egg demand

Principal Investigator: Vic Adamowicz
Graduate Research Assistant: Tatiana Gubanova
Department of Rural Economy
University of Alberta

Research Project Number CMD 08-08

PROJECT REPORT

July 2008

Department of Rural Economy
Faculty of Agriculture & Forestry,
and Home Economics
University of Alberta
Edmonton, Canada

Latent consideration sets in egg demand

.

Principal Investigator: Vic Adamowicz
Graduate Research Assistant: Tatiana Gubanova
Department of Rural Economy
University of Alberta

Acknowledgements: Agriculture and Agri-Food Canada Consumer and Market Demand Agriculture Policy Research Network.

Introduction

This project was to examine the use of Kuhn-Tucker demand systems (to assess price, attribute and consideration set factors in consumer demand for eggs. Data from Ontario and Alberta were used for separate analysis of consumption patterns. Typically models of this type are intractable for data sets with more than 5 or 6 alternatives. We attempted to employ a relatively new algorithm by van Haefen that has been successful for large numbers of alternatives and has been used in the examination of choice set structure.

The project also involved a descriptive analysis of the eggs data for seasonality and an assessment of the mix of egg types consumed by households. The latter analysis was carried out by using entropy as a metric of the mix of egg types consumed by households as a function of prices, incomes, and demographic characteristics.

Finally, the analysis also included an assessment of the extent to which demand inertia exists in egg consumption in Canada. This assessment used an empirical model that examine egg consumption when price data are difficult to obtain or are not reliable for the application at hand. A separate working paper on the document is provided along with this report.

Results:

Initial data analysis indicated that the expenditure / quantity data were based on "trips" and thus several assumptions had to be made regarding the construction of price series. Furthermore, there were several missing data entries or cases in which expenditure data were provided and no quantity recorded, or quantity data with no expenditures. Appendix 1 below outlines the assumptions used to develop price series for the remainder of the analysis.

Descriptive analysis of the data provided the following insights:

- Smaller and higher income families with no children seem to be more likely to consume multiple varieties of eggs.
- There is statistically significant positive association between choice entropy and

price dispersion. A larger effect of price differences on choice implies greater substitutability.

- There are two peaks in the consumption of eggs—a smaller peak occurs in the end of April (Easter) and a larger one occurs in the end of December—the beginning of January (Christmas and New Year Holidays). A gap in purchases of eggs in late July—early August may represent people leaving the corresponding provinces for vacation.
- There is considerable demand inertia in egg consumption by brand (UPC). Interestingly, this inertia is more pronounced in Alberta consumer data than in Ontario consumer data.
- Information about the panel data set:
  - Families with more than 2 household members have a higher tendency to leave the panel.
  - Families with kids under 18 years old have a higher tendency to leave the panel.
  - Families with the head under 45 years old have a higher tendency to leave the panel.
  - Families with higher incomes are less likely to stay in the panel throughout the analysis period.

Details of the descriptive analysis is presented below along with the description of the price calculation algorithm and the explanation of the Kuhn-Tucker model estimation difficulties.

Keywords: egg consumption, Kuhn Tucker demand system, consumer choice

JEL codes: D1, C1

Appendix 1: Assumptions regarding price determination

An Algorithm of Retrieving per-unit Prices

Data issues.

There are errors and omissions in data. For example, there are observations where several units were purchased but zero spending was recorded. There are also unreasonable dollar values, like 19 cents paid for one unit.

Assumptions.

(1) The dollar figures ($/trip) in the data set are dollars spent on a UPC per trip per period (from the layout description).

(2) We assume that shopping trips with no eggs purchased were not counted to obtain the $/trip values.

(3) We assume 12 eggs in a pack unless noted otherwise in the UPC description provided by Ellen.

An algorithm of retrieving per-pack prices.

(1) For each UPC/ID (i.e. for each row), first find observations where only one unit was purchased. Leave the dollar figures for such observations as is—there are per unit prices already. If there is at least one such observation in the row, use it as a "reference" price with observations where more than one unit was purchased and the price needs to be guessed.

    N units can be purchased in 1 to a maximum of N shopping trips. Go from trip=1 to N. For each trip value, calculate the total expenditure (total = $/trip*trip) and obtain the average price (average = total/units). Find the trip value with the minimum difference between the "reference" price, if available, and its average price. Use this average price value as the per-unit price for the observation in question.

Step 1 cannot retrieve prices for rows where it is impossible to get the starting "reference" price. There rows are processed in Step 2.

(2) For each unresolved (in Step 1) row, select resolved rows from other IDs with the same UPC. Average prices in those, use the average as the "reference" price and proceed as in Step 1.

Step 2 leaves some 10-15 observations per province where there reference price still could not be obtained. We then did those observations manually, by looking at other items from the same manufacturer and/or of the same type.

(3) To deal with errors in data, replace price values below 5th and above 95th percentiles with the mean for each UPC, if there are at least 10 recovered prices available. This introduces minor corruption to the data but removes many outlier values.

An algorithm of retrieving per-unit prices.

(1) Take the data set with the per-pack prices.

(2) Calculate an average per-pack price across individuals for a chosen UPC code. Divide this average by the number of eggs in a pack for a chosen UPC code.

Notation for the Final Dataset file:

p = per-egg price, in cents;

q = quantities, number of eggs;

e = total expenditures = p*q, in cents

Appendix 2: Descriptive Analysis

Part 1. Alberta

The original data set has been separated into two data sets—one represents an incomplete panel and the other one represents a complete panel.

The incomplete panel has 1,036 observations, and complete panel has 1,608 observations. The total number of people participating in the project is 2,644.

Descriptive statistics for both incomplete and complete panels is examined for five variables—household size, household income, age of the household head, the presence of children under 18 years old, and language. This analysis provides information on the factors that affect continued membership in the panels.

The language variable has been dropped from the analysis as the Mann-Whitney test for two independent samples failed to reject the null hypothesis that both incomplete and complete samples come from the same population. Test statistic is 832890 and p-value is 0.9701.

HHLD Size

Value	Description
1	Single Member
2	Two Members
3	Three Members
4	Four Members
5	Five Members
6	Six Members
7	Seven Members
8	Eight Members
9	Nine or More Members

HHLD Income

Value	Description
08	Under $10,000
10	$10000-$14999
11	$15000-$19999
13	$20000-$24999
15	$25000-$29999
16	$30000-$34999
17	$35000-$39999
18	$40000-$44999
19	$45000-$49999
21	$50000-$54999
23	$55000-$69999
27	$70000-$84999
28	$85000-$99999
29	$100000-$124999
30	$125000 +

HHLD Head Age

Value	Description
1	Under 35
2	35-44
3	45-54
4	55-64
5	65 and over

Presence of Children

Value	Description
1	Children Under 18
2	No Children Under 18

Household size analysis. Alberta

Figure 1. The distribution of households by the household size for both incomplete and complete panels

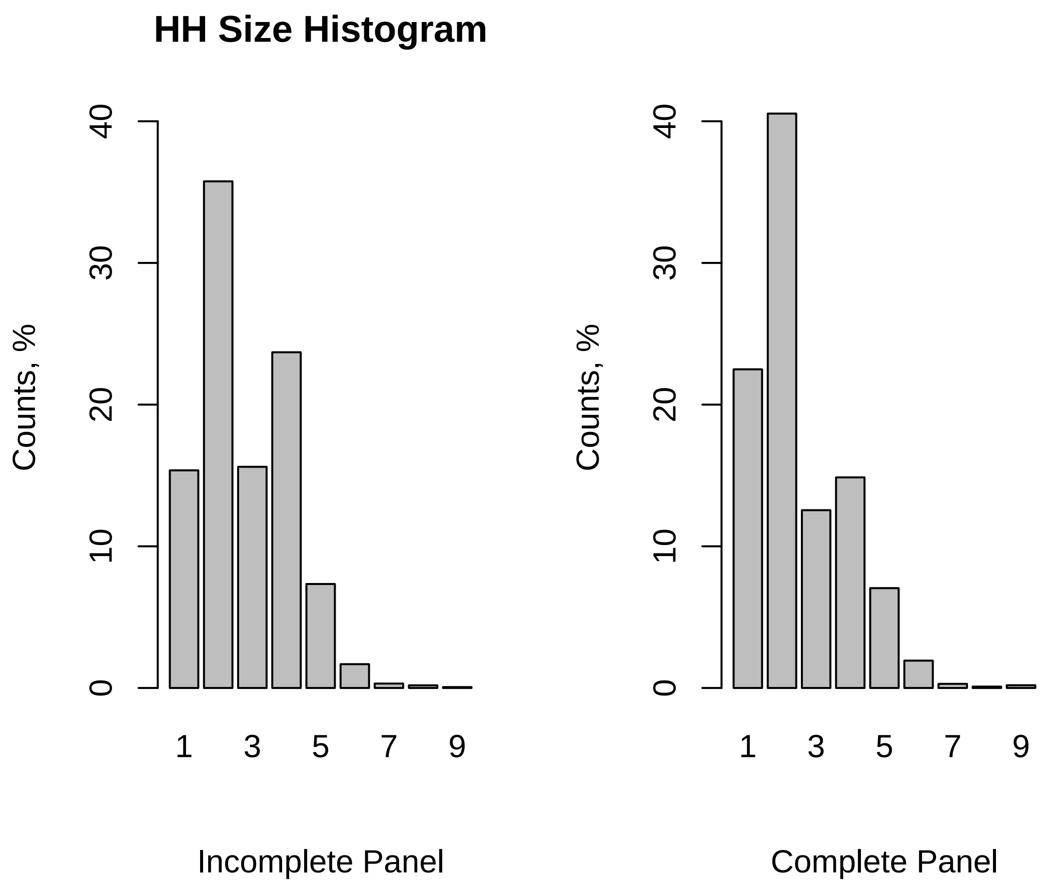


Table 1. Frequency table for the incomplete panel

Size group	1	2	3	4	5	6	7	8	9	Total
Frequency	247	575	251	381	118	27	5	3	1	1608
Percentage of total	15.4	35.8	15.6	23.7	7.3	1.7	0.3	0.2	0.1	100.0

Table 2. Frequency table for the complete panel

Size group	1	2	3	4	5	6	7	8	9	Total
Frequency	233	420	130	154	73	20	3	1	2	1036
Percentage of total	22.5	40.5	12.5	14.9	7.0	1.9	0.3	0.1	0.2	100.0

Families with three, four, and five members have a tendency to leave the panel before the project ends. Families with one and two members prefer to stay in the panel till the project ends (see "percentage of total", Tables 1 and 2). Families with the number of members from six to nine have about equal chances of either staying in the panel or leaving it.

Quantile tables show the difference in the sampling distribution of households between incomplete and complete panel samples.

Table 3. Quantile table for incomplete panel by household size

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Size group	1	1	1	2	2	4	4	5	6

Table 4. Quantile table for complete panel by household size

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Size group	1	1	1	2	2	3	4	5	6

A Two-sample Kolmogorov-Smirnov test rejected the null hypothesis that both incomplete and complete samples come from the same population. The test statistic is 0.1191 and p-value is approximately zero.

Household income analysis. Alberta

Figure 2. The distribution of households by household income for both incomplete and complete panels

Table 5. Frequency table for the incomplete panel

Income group	8	10	11	13	15	16	17	18	19	21	23	27	28	29	30	Total
Frequency	23	34	27	52	65	104	92	102	108	131	256	236	152	122	104	1608
Percentage of total	1.4	2.1	1.7	3.2	4.0	6.5	5.7	6.3	6.7	8.1	15.9	14.7	9.5	7.6	6.5	100.0

Table 6. Frequency table for the complete panel

Income group	8	10	11	13	15	16	17	18	19	21	23	27	28	29	30	Total
Frequency	14	36	37	76	52	70	53	60	47	60	166	126	90	83	66	1036
Percentage of total	1.4	3.5	3.6	7.3	5.0	6.8	5.1	5.8	4.5	5.8	16.0	12.2	8.7	8.0	6.4	100.0

Families with the income of less than $35,000 per year (group 17) have high chances of staying in the panel till the project ends. Families with the annual income of more than $35,000 per year are likely to leave the panel before the project ends.

Table 7. Quantile table for incomplete panel by the household income

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Income group	8	11	15	18	23	27	29	30	30

Table 8. Quantile table for incomplete panel by the household income

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Income group	8	11	13	16	23	27	29	30	30

A Two-sample Kolmogorov-Smirnov test rejected the null hypothesis that both incomplete and complete samples come from the same population. The test statistic is 0.0854 and p-value is approximately zero.

Household head age analysis. Alberta

Figure 3. The distribution of households by household head age for both incomplete and complete panels

Table 9. Frequency table for the incomplete panel

Size group	1	2	3	4	5	Total
Frequency	146	478	441	305	238	1608
Percentage of total	9.1	29.7	27.4	19	14.8	100

Table 10. Frequency table for the complete panel

Size group	1	2	3	4	5	Total
Frequency	72	243	311	208	202	1036
Percentage of total	6.9	23.5	30	20.1	19.5	100

Families with the head under 45 years old have a tendency of leaving the panel before the project ends. Families with the head above 45 years old have a tendency of staying in the panel till the end of the project.

Table 11. Quantile table for incomplete panel by the household head age

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Head age group	1	1	2	2	3	4	5	5	5

Table 12. Quantile table for incomplete panel by the household head age

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Head age group	1	1	2	2	3	4	5	5	5

A two-sample Kolmogorov-Smirnov test rejected the null hypothesis that both incomplete and complete samples come from the same population. Test statistic is 0.084 and p-value is approximately zero.

Household children analysis. Alberta

Figure 4. The distribution of households by the presence of children under 18 years old for both incomplete and complete panels

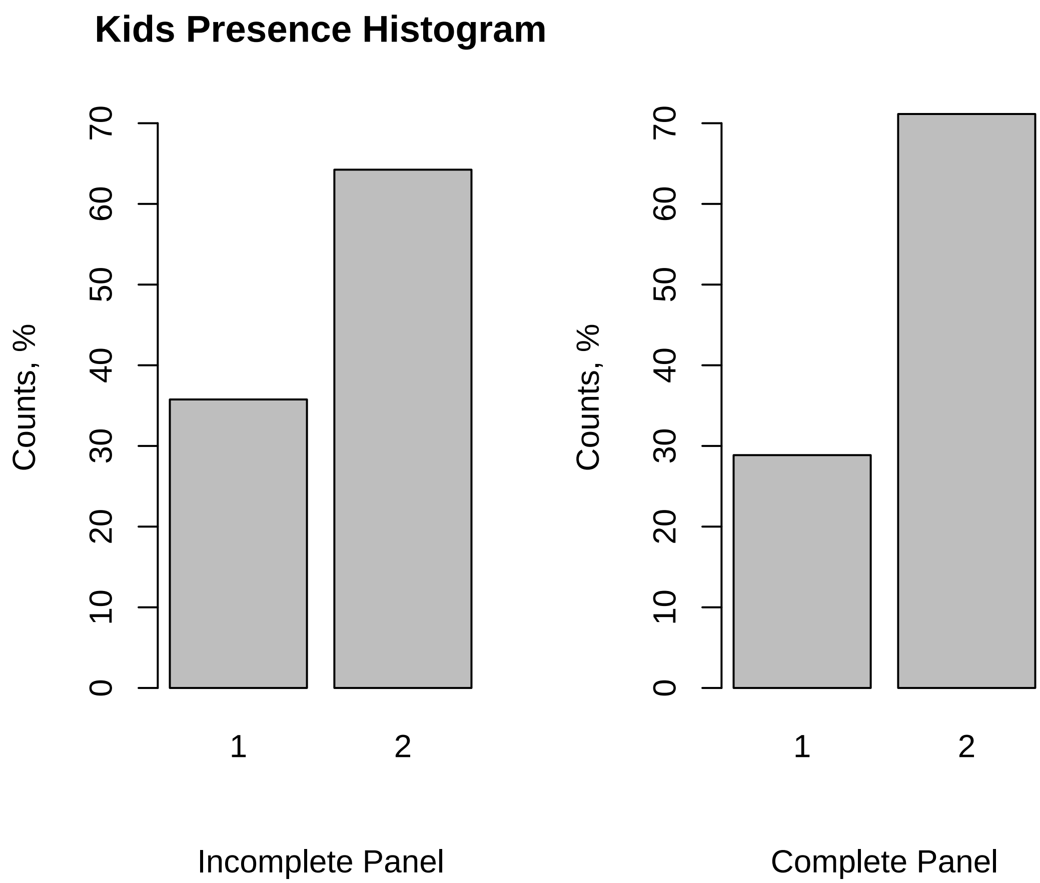


Table 13. Frequency table for the incomplete panel

Child group	1	2	Total
Frequency	575	1033	1608
Percentage of total	35.8	64.2	100

Table 14. Frequency table for the complete panel

Child group	1	2	Total
Frequency	299	737	1036
Percentage of total	28.9	71.1	100

Families with kids under 18 years old have a tendency to leave the panel before the project ends. Families without kids under 18 years old have a tendency to stay in the panel until the end of the project.

The Mann-Whitney test for two independent samples failed to accept the null hypothesis that both incomplete and complete panels come from the same population. Test statistic is 775490 and p-value is approximately zero.

Part 2. Ontario.

The original data set has been separated into two data sets—one represents an incomplete panel (those who enter the project from the beginning but left it before the project ends) and the other one represents a complete panel.

The incomplete panel has 2,843 observations, and the complete panel has 2,031 observations. The total number of people participating in the project is 4,874.

Descriptive statistics for both incomplete and complete panels is examined for five variables—household size, household income, age of the household head, the presence of children under 18 years old, and language. This analysis provides information on the factors that affect continued membership in the panels.

The language variable has been dropped down from the analysis as the Mann-Whitney test for two independent samples showed no difference in the two samples—incomplete panel and complete panel—based on this variable. Test statistic is 2896211 (under two-sided alternative hypothesis) and p-value is 0.3561.

Household size analysis. Ontario.

Figure 5. The distribution of households by household size for both incomplete and complete panels

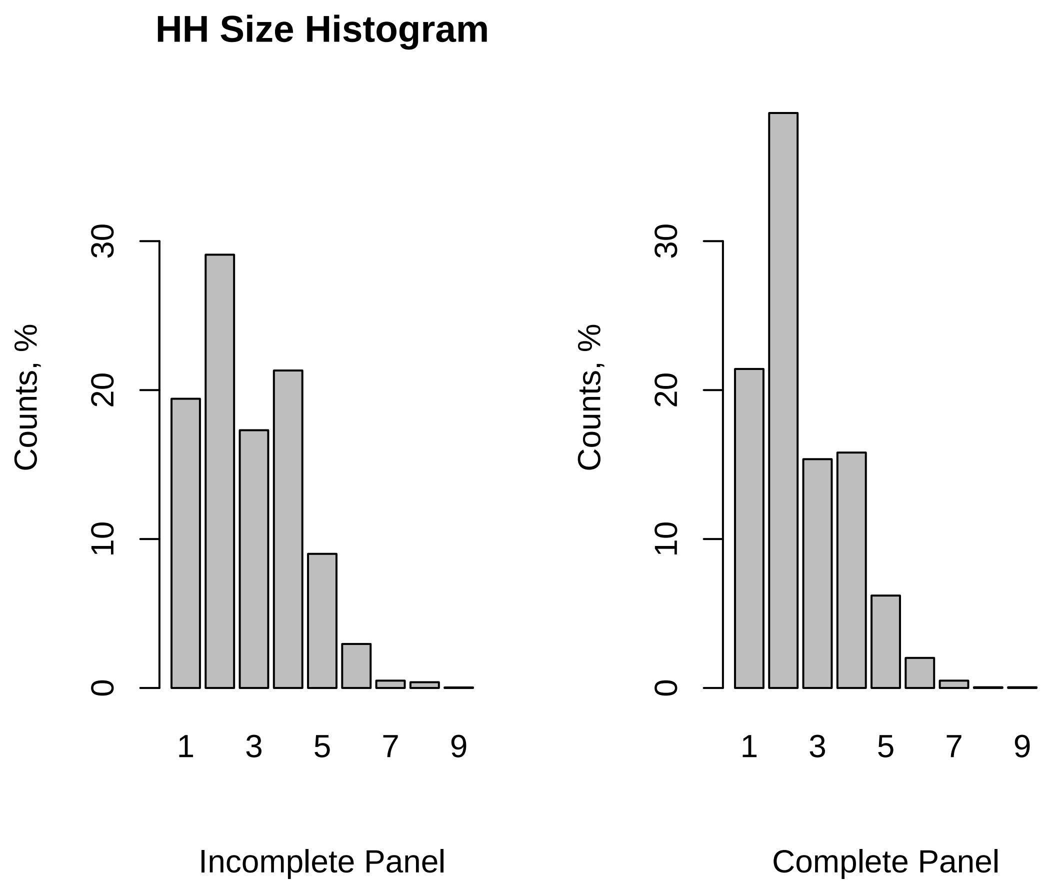


Table 16. Frequency table for the incomplete panel

Size group	1	2	3	4	5	6	7	8	9	Total
Frequency	552	827	492	606	256	84	14	11	1	2843
Percentage of total	19.4	29.1	17.3	21.3	9.0	3.0	0.5	0.4	0.0	100

Table 17. Frequency table for the complete panel

Size group	1	2	3	4	5	6	7	8	9	Total
Frequency	435	784	312	321	126	41	10	1	1	2031
Percentage of total	21.4	38.6	15.4	15.8	6.2	2.0	0.5	0.0	0.0	100

Families with two+ members have a tendency to leave the panel before the project ends. Families with one and two members prefer to stay in the panel till the project ends (see “percentage of total”, Tables 16 and 17). Probably, families with at least one child do not have time to stay in the panel constantly. And the benefits from participating in the project do not seem to be sufficient motivation for those families in Ontario.

Quantile tables show the difference in the sampling distribution of households between incomplete and complete panel samples.

Table 18. Quantile table for incomplete panel by the household size

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Size group	1	1	1	2	3	4	5	5	6

Table 19. Quantile table for complete panel by the household size

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Size group	1	1	1	2	2	3	4	5	6

A two-sample Kolmogorov-Smirnov test failed to accept the null hypothesis that both incomplete and complete panels come from the same population. Test statistic is 0.1151 and p-value is approximately zero.

Household income analysis. Ontario.

Figure 6. The distribution of households by household income for both incomplete and complete panels

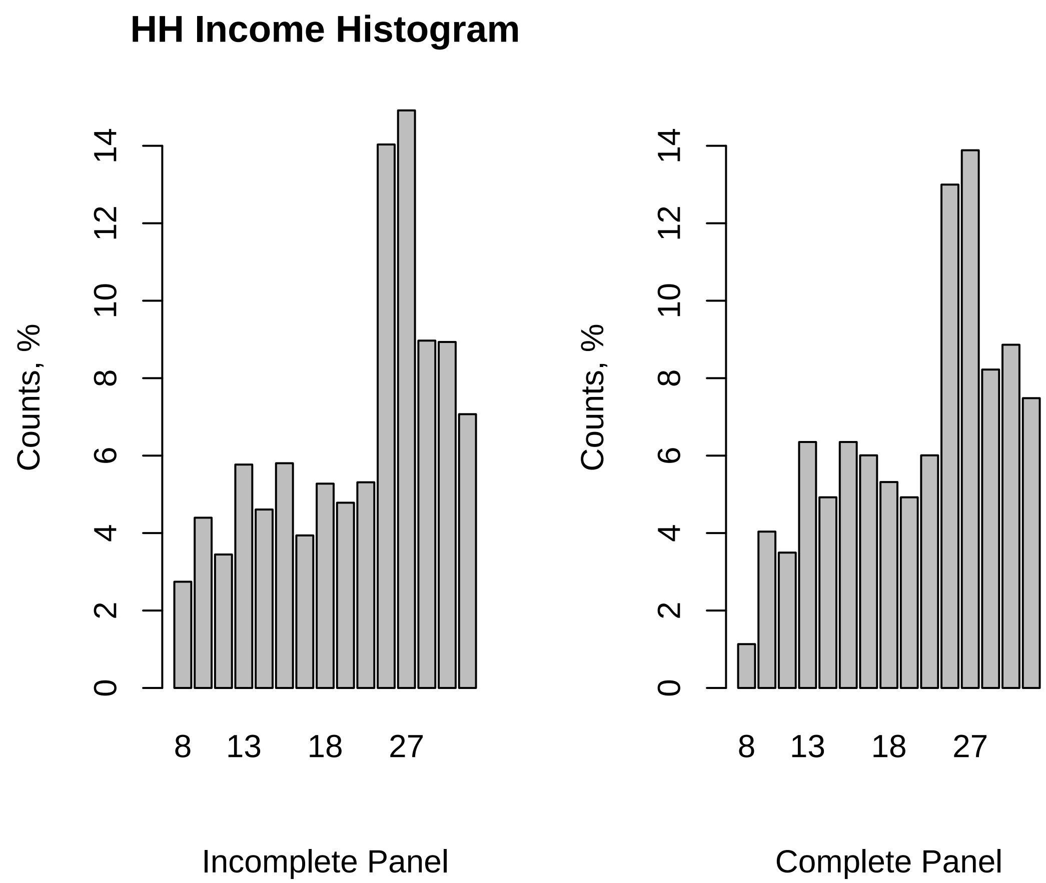


Table 20. Frequency table for the incomplete panel

Income group	8	10	11	13	15	16	17	18	19	21	23	27	28	29	30	Total
Frequency	78	125	98	164	131	165	112	150	136	151	399	424	255	254	201	2843
Percentage of total	2.7	4.4	3.4	5.8	4.6	5.8	3.9	5.3	4.8	5.3	14.0	14.9	9.0	8.9	7.1	100

Table 21. Frequency table for the complete panel

Income group	8	10	11	13	15	16	17	18	19	21	23	27	28	29	30	Total
Frequency	23	82	71	129	100	129	122	108	100	122	264	282	167	180	152	2031
Percentage of total	1.1	4.0	3.5	6.4	4.9	6.4	6.0	5.3	4.9	6.0	13.0	13.9	8.2	8.9	7.5	100

Families with an income of under $10,000 per year (group 8) and of over $55,000 per year (group 23+) have high chances of leaving the panel before the project ends.

Table 22. Quantile table for incomplete panel by the household income

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Income group	8	10	11	16	23	27	29	30	30

Table 23. Quantile table for incomplete panel by the household income

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Income group	8	10	13	16	23	27	29	30	30

A two-sample Kolmogorov-Smirnov test failed to reject the null hypothesis that both incomplete and complete panels come from the same population. Test statistic is 0.0247 and p-value is 0.4654.

Household head age analysis. Ontario.

Figure 7. The distribution of households by household head age for both incomplete and complete panels

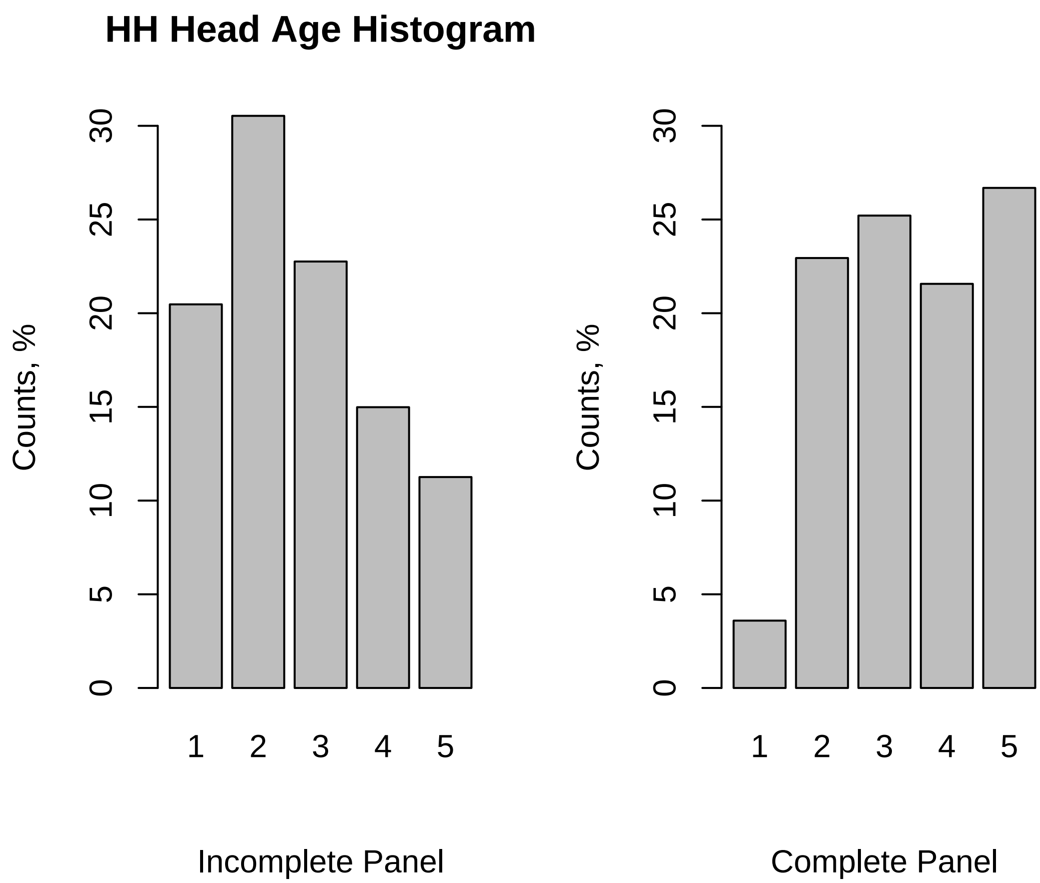


Table 24. Frequency table for the incomplete panel

Size group	1	2	3	4	5	Total
Frequency	582	868	647	426	320	2843
Percentage of total	20.5	30.5	22.8	15.0	11.3	100.0

Table 25. Frequency table for the complete panel

Size group	1	2	3	4	5	Total
Frequency	73	466	512	438	542	2031
Percentage of total	3.6	22.9	25.2	21.6	26.7	100.0

Families with the head under 44 years old (groups 1 and 2) have a tendency to leave the panel before the project ends. Families with the head over 45 years old (groups

3, 4 and 5) have a tendency to stay in the panel until the end of the project.

Table 26. Quantile table for incomplete panel by the household head age

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Head age group	1	1	1	2	2	4	5	5	5

Table 27. Quantile table for incomplete panel by the household head age

Quantile	0.01	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.99
Head age group	1	2	2	2	3	5	5	5	5

A two-sample Kolmogorov-Smirnov test rejected the null hypothesis that both incomplete and complete panels come from the same population. Test statistic is 0.2446 and p-value is approximately zero.

Household children analysis. Ontario.

Figure 8. The distribution of households by the presence of children under 18 years old for both incomplete and complete panels

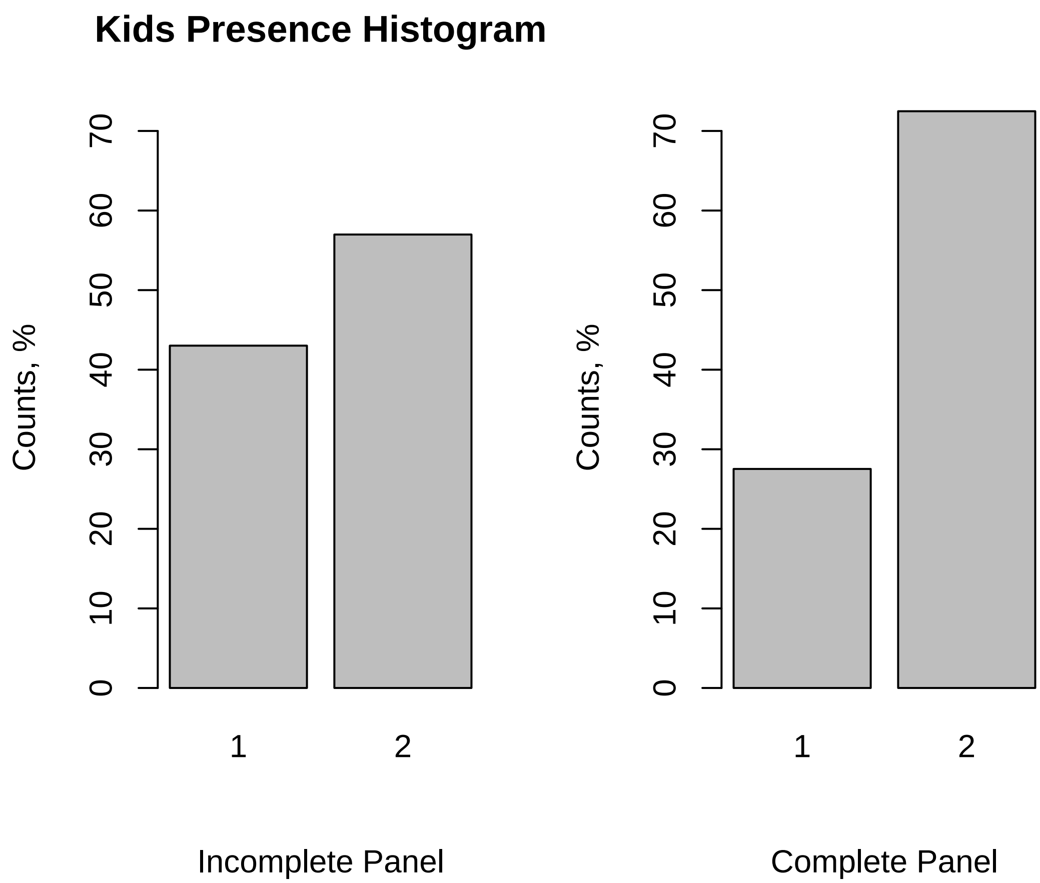


Table 28. Frequency table for the incomplete panel

Child group	1	2	Total
Frequency	1223	1620	2843
Percentage of total	43.0	57.0	100

Table 29. Frequency table for the complete panel

Child group	1	2	Total
Frequency	559	1472	2031
Percentage of total	27.5	72.5	100

Families with kids under 18 years old have a tendency to leave the panel before the project ends. Families without kids under 18 years tend to stay in the panel till the end of the project.

The Mann-Whitney test for two independent samples rejected the null hypothesis that both incomplete and complete panels come from the same population. Test statistic is 2439729 and p-value is approximately zero.

Analysis of Variation in Expenditures on Eggs

The number of unique brands of eggs purchased in Alberta is 187. The number of unique brands of eggs purchased in Ontario is 236.

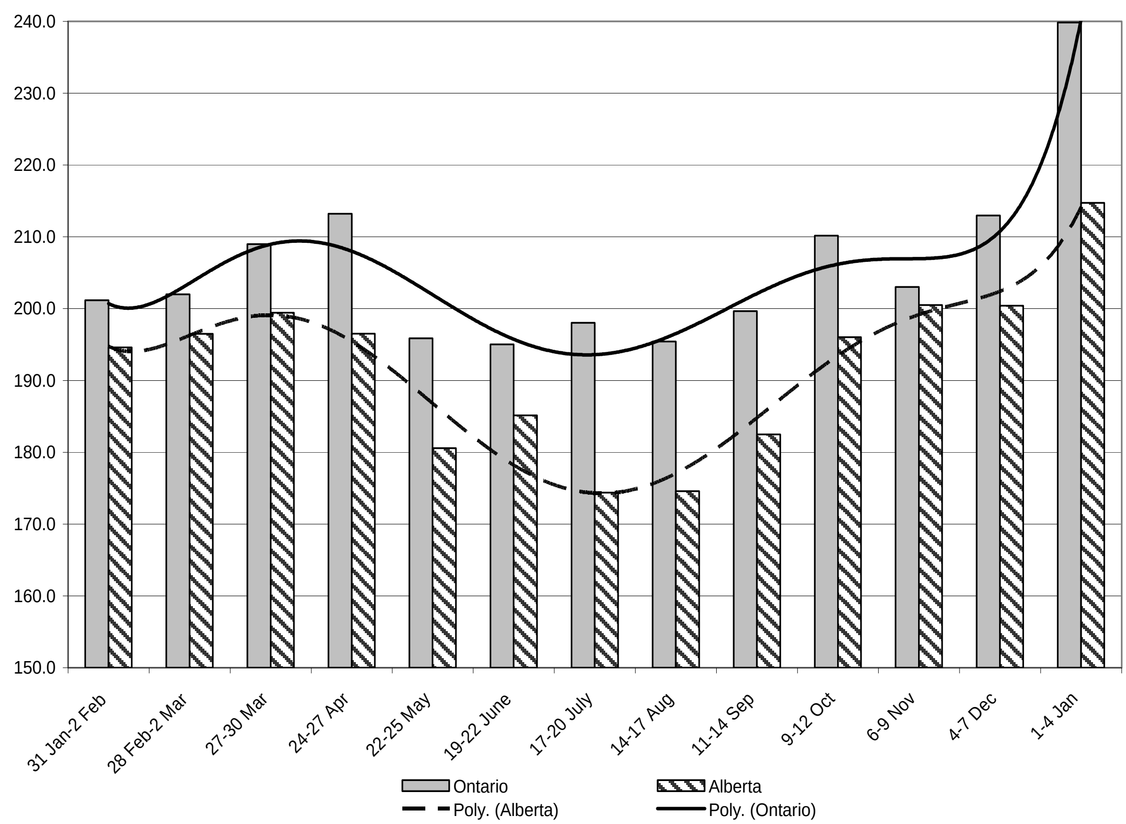
I. Seasonal Variation

Seasonal variation in eggs expenditures was checked by a nonparametric test—the Friedman rank sum test. The Friedman test is applied to the complete panel, i.e. those people who bought eggs during at least one four-week period within each year.

The complete panel size for the Friedman test is 1,036 for Alberta. The Friedman rank sum test rejected the null hypothesis that there was no seasonal variation in the data for Alberta. The test statistic is 137.9 and the p-value is approximately zero.

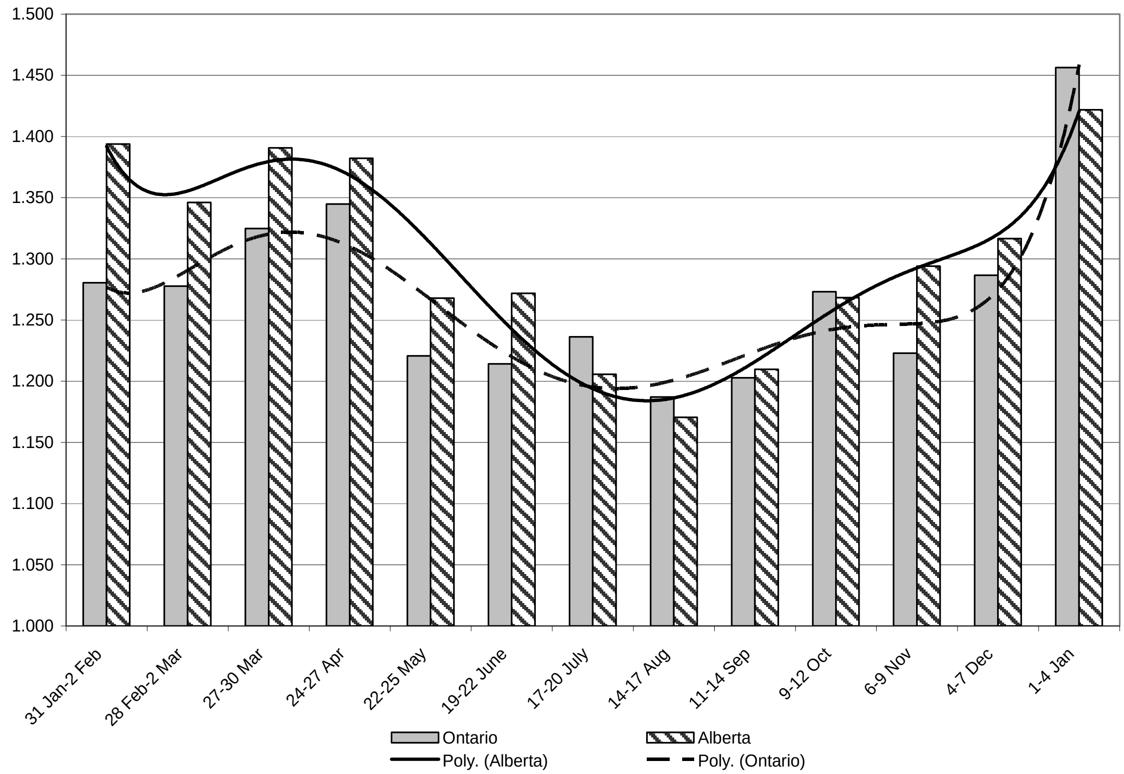
The complete panel size for the Friedman test is 2,031 for Ontario. The Friedman rank sum test rejected the null hypothesis that there was no seasonal variation in the data for Ontario. The test statistic is 376.3 and the p-value is approximately zero.

Figure 9. Average 4-week egg expenditures per person per trip across three years (2002-2004) for Alberta and Ontario: Complete Panel, in cents



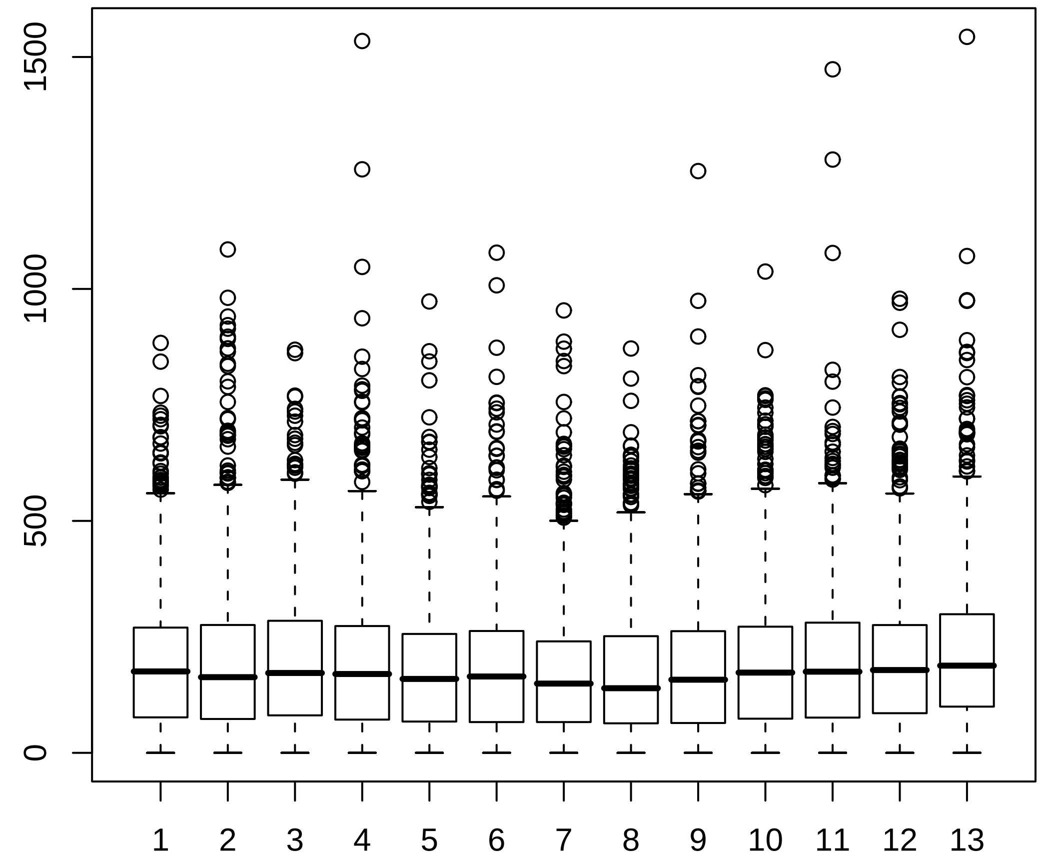
There are two peaks in the consumption of eggs—a smaller peak occurs in the end of April (Easter) and a larger one falls in the end of December—the beginning of January (Christmas and New Year Holidays). A gap in purchases of eggs in late July—early August may represent people leaving the corresponding provinces for vacation.

Figure 10. Average 4-week units of eggs per person per trip across three years (2002-2004) for Alberta and Ontario: Complete Panel, in packs.



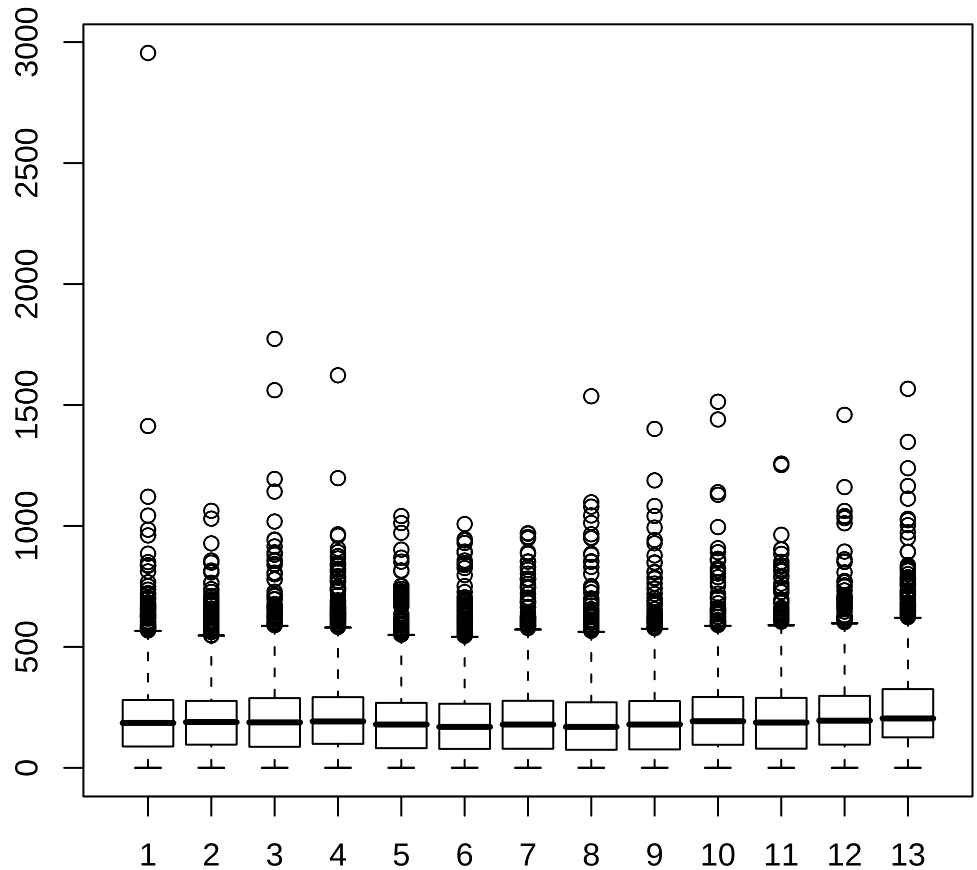
II. Variation Across Households

Figure 11. Average 4-week Variation of Expenditures on Eggs across Households for Alberta, Complete Panel, in cents



	31 Jan-2 Feb	28 Feb-2 Mar	27-30 Mar	24-27 Apr	22-25 May	19-22 June	17-20 July	14-17 Aug	11-14 Sep	9-12 Oct	6-9 Nov	4-7 Dec	1-4 Jan
Time	1	2	3	4	5	6	7	8	9	10	11	12	13
Min	0	0	0	0	0	0	0	0	0	0	0	0	0
0.25	77	73	81	72	68	66	66	64	64	74	76	86	100
0.50	176	163	172	170	159	165	149	139	158	173	175	179	188
Mean	195	197	199	197	181	185	174	175	183	196	201	200	215
0.75	270	276	285	273	256	263	240	251	262	272	281	275	299
Max	884	1085	869	1534	973	1078	954	872	1254	1037	473	979	1543

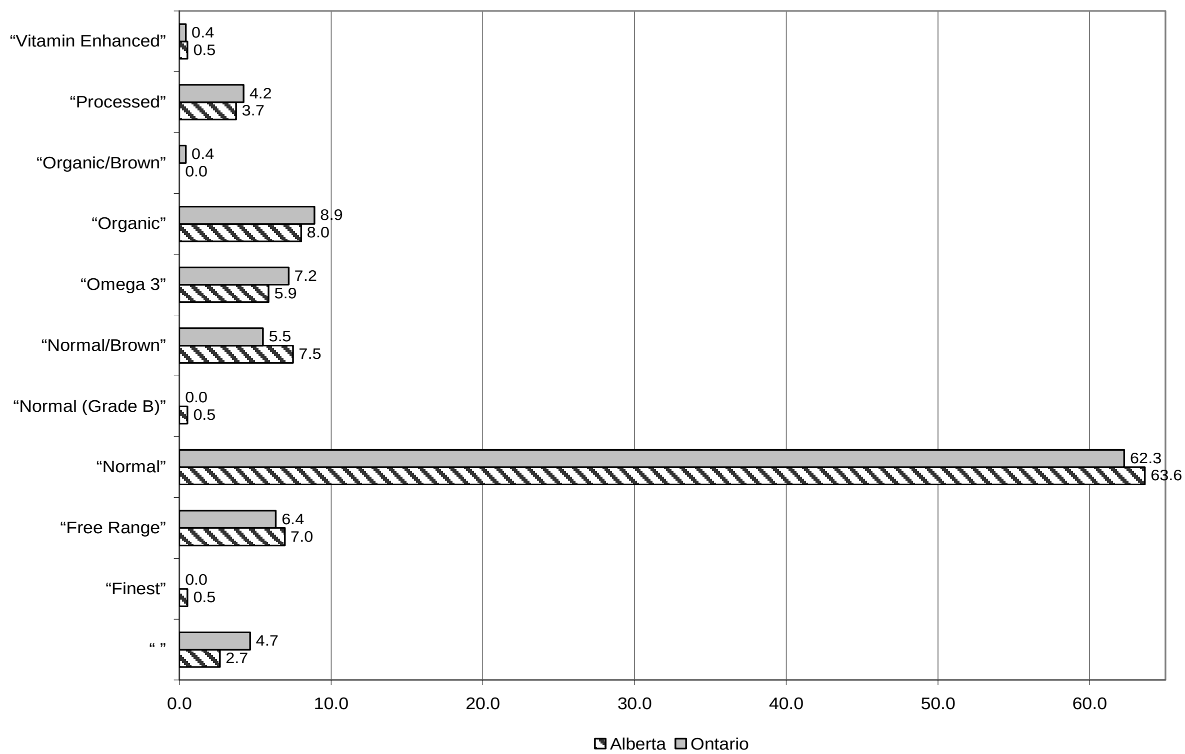
Figure 12. Average 4-week Variation of Expenditures on Eggs across Households for Ontario, Complete Panel, in cents



	31 Jan-2 Feb	28 Feb-2 Mar	27-30 Mar	24-27 Apr	22-25 May	19-22 June	17-20 July	14-17 Aug	11-14 Sep	9-12 Oct	6-9 Nov	4-7 Dec	1-4 Jan
Time	1	2	3	4	5	6	7	8	9	10	11	12	13
Min	0	0	0	0	0	0	0	0	0	0	0	0	0
0.25	89	96	87	99	81	78	79	75	76	96	80	96	126
0.50	186	189	188	192	179	169	179	169	179	193	188	195	204
Mean	201	202	209	213	196	195	198	195	200	210	203	213	240
0.75	280	277	288	292	269	265	277	271	276	293	289	298	325
Max	955	63	773	623	1041	8	970	536	401	514	257	459	1567

Analysis of Types of Eggs Purchased

Figure 13. **Distribution of Types of Eggs for Unique Brands of Eggs Purchased by Consumers in Alberta and Ontario: 2002--2004**



In Alberta, consumers buy more "Normal" types of eggs (both white and brown), more "Free Run", less "Processed", less "Omega 3" and less "Organic" (both white and brown).

Table 31. Distribution of Types of Eggs for Unique Brands of Eggs Purchased by Consumers in Alberta and Ontario: 2002--2004

Type of Eggs	Alberta		Ontario	
	Number	Percentage of total	Number	Percentage of total
“ ”	5	2.7	11	4.7
“Finest”	1	0.5	0	0.0
“Free Range” or “Free Run”	13	7.0	15	6.4
“Normal”	119	63.6	147	62.3
“Normal (Grade B)”	1	0.5	0	0.0
“Normal/Brown”	14	7.5	13	5.5
“Omega 3”	11	5.9	17	7.2
“Organic”	15	8.0	21	8.9
“Organic/Brown”	0	0.0	1	0.4
“Processed”	7	3.7	10	4.2
“Vitamin Enhanced”	1	0.5	1	0.4
Total	187	100.0	236	100.0

Analysis of Egg Type Purchase Entropy

To compute egg consumption variety purchase entropy was calculated for each ID. All unique UPCs were first categorized into 6 types:

- Normal (also includes "normal grade B" and "finest")
- Normal/brown
- Free Range/Run
- Omega-3 (also includes "Vitamin Enhanced")
- Organic
- Processed

ID-specific purchase entropy for the outcomes in the 6 groups above was calculated as $e = -\sum_{i=1}^{6} \hat{p}_i \log_2 \hat{p}_i$

Choice probabilities $\hat{p}_i$ for $i = 1...6$ egg types were obtained empirically, dividing the number of periods in which the respective type of eggs was purchased by the individual by the total number of periods in which eggs were purchased (max = 13 x 3 years = 39). Note that numbers of units purchased per period were disregarded; if any purchases were made during a period, the latter would be marked as 1, and 0 otherwise.

The entropy was contrasted with demographics and price dispersion by computing Kendall's concordance and conducting Kendall's tau-test for association. The price dispersion was measured by finding the variance of average within-type prices. The latter were calculated by averaging all non-zero prices across both time and UPCs. If a person were buying eggs of only one type, then the entropy and price variance would both be zero for that person. Only non-zero entropy-price pairs were used to test for association between the choice entropy and price dispersion.

Alberta

22% of IDs (585 out of 2644) purchased more than one type of eggs and, correspondingly, had non-zero entropy and variance.

Household Size

```
z = -3.3601, p-value = 0.0007791
alternative hypothesis: true tau is not equal to 0
sample estimates:
        tau
-0.05588785
```

Household Income

```
z = 6.0999, p-value = 1.061e-09
alternative hypothesis: true tau is not equal to 0
sample estimates:
       tau
0.09552532
```

Household Head Age

```
z = 0.5204, p-value = 0.6028
alternative hypothesis: true tau is not equal to 0
sample estimates:
        tau
0.008643588
```

Presence of Children in Household

(levels: 1- present, 2 – not present)

```
z = 2.9616, p-value = 0.00306
alternative hypothesis: true tau is not equal to 0
sample estimates:
       tau
0.05449989
```

Statistically significant positive association was found with household income and the absence of children; negative association was found with the household size. **Smaller and more well-off families with no children seem to be less likely to stick to just one type of eggs.** Note that the tests were run pair-wise, so no partial association effects are available. For reference, concordance between the demographic variables appears below.

	[size,1]	[inc,2]	[age,3]	[child,4]
[1,]	1.0000000	0.2520823	-0.2821121	-0.6634806
[2,]	0.2520823	1.0000000	-0.1269996	-0.1288240
[3,]	-0.2821121	-0.1269996	1.0000000	0.4378673
[4,]	-0.6634806	-0.1288240	0.4378673	1.0000000

Price Dispersion

```
z = 3.7082, p-value = 0.0001044
alternative hypothesis: true tau is greater than 0
sample estimates:
      tau
0.1029127
```

There is statistically significant positive association between choice entropy and price dispersion. The estimate has the proper sign (a greater substitutability implies a larger effect of price differences on choice). **When considered composite goods, different types of eggs are substitutes to some degree.**

Appendix 3: Reports of Kuhn-Tucker Demand Estimation on Alberta and Ontario Eggs Databases

Alberta

Variables in ψ: Constant, declared as random

Variables in repack (attributes):

Pack (6, 12, 18) --- number of eggs in a retail package

Brown, binary indicator of whether the egg type is brown

Omega3, binary indicator of whether the egg type is Omega3

Free_run, binary indicator of whether the egg type is free-run

Organic, binary indicator of whether the egg type is organic

Model estimated with 50 grouped UPC (as commodities) and 39 time periods as "consumers"; budget available was assumed to be the total egg expenditure during the period. UPC grouping was done on the basis of the 3-year-long price level

Halton sequence-based quasi-random integration was enabled. All UPCs were assumed to be available for consumption for each time period.

```
Iteration  Func-count       f(x)        Step-size       optimality
    20         231          7264.69             1             404
    21         242          7263.73             1             421
    22         253          7262.77             1             432
    23         264          7261.51             1             435
Line search cannot find an acceptable point along the current
 search direction.
```

Estimation halted at iteration 23 due to the impossibility of carrying out a line search. Errors like this may occur when the likelihood is ever-increasing or has a discontinuity.

Alternative formulations examined:

Fixed constant for all time periods as opposed to the random parameter model above. Algorithm exited at around iteration 25 for the same reason.
Fixed parameters for all time periods, including seasonal consumption trend and dummies for the second and third years. Algorithm failed at around iteration 25 for the same reason.
Number of grouped UPCs reduced to 42. Algorithm failed at around iteration 25 for the same reason.

Ontario

Variables in ψ: Constant, declared as random
Variables in repack (attributes):
Pack (6, 12, 18) --- number of eggs in a retail package
Brown, binary indicator of whether the egg type is brown
Omega3, binary indicator of whether the egg type is Omega3
Free_run, binary indicator of whether the egg type is free-run
Organic, binary indicator of whether the egg type is organic

Model estimated with 50 grouped UPC (as commodities) and 39 time periods as "consumers"; budget available was assumed to be the total egg expenditure during the period. UPC grouping was done on the basis of the 3-year-long price level

Halton sequence-based quasi-random integration was enabled. All UPCs were assumed to be available for consumption for each time period.

Estimation could not start due to the singularity of the Hessian.

<u>Alternative formulations tried:</u>

Fixed constant for all time periods as opposed to the random parameter model above. Algorithm could not start for the same reason.
Fixed parameters for all time periods, including seasonal consumption trend and dummies for the second and third years. Algorithm could not start.
Number of grouped UPCs reduced to 42, and then further reduced to 30. Algorithm could not start for the same reason.

Conclusion

The KT demand model as formulated here is not suitable for these egg consumption datasets. The latter have severe attrition problems, income is unavailable, and regular eggs (12 pack, white) dominate consumption with a share of 85% to 95% in a period. Furthermore there may not be sufficient variation in price over the time periods and ranges of products to identify price parameters.