HETEROGENEOUS DEMAND FOR FOOD DIVERSITY: A QUANTILE REGRESSION ANALYSIS

HETEROGENE NACHFRAGE NACH LEBENSMITTELVIELFALT: EINE QUANTIL-REGRESSIONSANALYSE

Larissa S. Drescher[1] and Ellen W. Goddard[2]

LDrescher@tum.de

[1]Marketing and Consumer Research, Technische Universität München, Alte Akademie 16, D-85354 Freising, Germany

[2]Department of Resource Economics and Environmental Sociology, University of Alberta, 515 General Services Building, T6G 2H1 Edmonton, Canada

2011

Vortrag anlässlich der 51. Jahrestagung der GEWISOLA
„Unternehmerische Landwirtschaft zwischen Marktanforderungen und gesellschaftlichen Erwartungen"
Halle, 28. bis 30. September 2011

HETEROGENEOUS DEMAND FOR FOOD DIVERSITY: A QUANTILE REGRESSION ANALYSIS

HETEROGENE NACHFRAGE NACH LEBENSMITTELVIELFALT: EINE QUANTIL-REGRESSIONSANALYSE

Abstract

Poverty and inequality studies frequently use the quantile regression approach to provide results on the impact of determinants at different points of the distribution of a dependent variable. To innovate diversity research this paper uses quantile regressions to identify determinants at different points of the food diversity distribution. Regional and household level differences in demand for food diversity are analysed based on a pooled sample of Canadian data of the Family Food Expenditure Surveys of Statistics Canada. Simple OLS regressions show that the determinants of Canadian diversity demand are similar to those of other developed countries. However, with quantile regressions significantly different effects of independent variables on diversity across quantiles are observed. In most cases, low diversified households, especially those in the lower 10% quantile of the distribution, are much more affected by key determinants such as household size, (real) income and age than higher-diversified households. Results further reveal that the demand for food diversity is not stable over time but is lower in 1996 and 2001 than in 1984. The diversity decline over time is higher for households in the middle quantiles compared to the moderate decline for households located at the ends of the diversity distribution.

Keywords: Household demand, Food diversity, Quantile regressions, Pooled samples, Canada

Zusammenfassung

In Armuts- und Ungleichheitsstudien sind Quantil-Regressionen eine häufig angewandte Methode um den Einfluss von Determinanten an verschiedenen Punkten der Verteilung einer abhängigen Variablen zu identifizieren. Um die bestehende Lebensmittelvielfaltsliteratur zu erweitern, werden in diesem Beitrag Quantil-Regressionen dafür genutzt, Nachfragedeterminanten an verschiedenen Stellen der Lebensmittelvielfaltsverteilung zu bestimmen. Dabei werden sowohl regionale als auch weitere haushaltsspezifische Unterschiede in der Nachfrage basierend auf einem gepoolten kanadischen Datensatz identifiziert. Ergebnisse einfacher Regressionen bestätigen, dass die Nachfrage nach Lebensmittelvielfalt in Kanada von ähnlichen Determinanten bestimmt ist wie in anderen Ländern. Die Quantil-Regressionen zeigen allerdings, dass je nach Quantil signifikant unterschiedliche Einflüsse auf die Vielfaltsnachfrage vorliegen. In den meisten Fällen sind die Haushalte mit der geringsten Vielfaltsnachfrage, insbesondere solche im unteren 10%-Quantil, am stärksten beeinflusst durch Schlüsselgrößen wie Haushaltsgröße, (Real)einkommen und Alter als Haushalte mit einer höheren Nachfrage nach Vielfalt. Die Ergebnisse zeigen auch, dass die Vielfaltnachfrage über die Zeit nicht konstant ist, sondern 1996 und 2001 niedriger ist als 1984. Der Rückgang in der Vielfaltsnachfrage über die Zeit ist größer bei Haushalten in den mittleren Quantilen im Vergleich Haushalten an den Enden der Vielfaltsverteilung.

Schlagwörter: Haushaltsnachfrage, Lebensmittelvielfalt, Quantil-Regressionen, gepoolte Daten, Kanada

1 Introduction

The link between higher GDP per capita and a nation's high demand for diversity of products is well documented (CLEMENT, WU and ZHANG, 2006; THEIL and FINKE, 1983). Among various product groups, food diversity has been frequently analysed on a microeconomic level generating a number of household or individual demand studies for a varied diet in various countries (e.g. MOON ET AL., 2002; STEWART and HARRIS, 2005). Jackson's hierarchical model of consumer demand (JACKSON, 1984), with higher income as a driving force behind increasing demand for diversity, has been used as the theoretical background for many of these studies.

There has been little change in the empirical methodology of analysing food diversity: usually cross-sectional data sets are analysed with ordinary least squares (OLS) using a diversity measure as the dependent variable explained by a set of socio-economic variables. Previous studies show that besides income typically variables such as age and household size increase the demand for a varied diet (DRESCHER ET AL., 2009; THIELE and WEISS, 2003). However, as these results are based on OLS regressions, they are restricted to only providing evidence of the impact of independent variables at the mean of the respective diversity measures. Often overlooked when studying diversity is whether the variables have significantly different effects at other points of the distribution of the dependent variable than at the mean. In other contexts such as returns to schooling or school composition effects, quantile regressions turned out as an effective tool to explain the relationship between dependent and independent variables (ARIAS, HALLOCK and SOSA-ESCUDERO, 2001; SCHINDLER RANGVID, 2001). Quantile regressions appear relevant for food diversity, as well, as they would identify population subgroups at the tails of the diversity distribution. This may provide relevant information for economic and nutrition policy, particularly in public health where specific information about vulnerable sub-groups could generate targeted interventions.

As a methodological improvement to economic diversity studies the use of panel data for analysing individual diversity demand has been suggested (THIELE and WEISS, 2003). This would allow the elimination of unobservable time-invariant individual effects and would additionally provide insights into trends in food consumption and inertia in consumption behaviour (ibid., p. 109). In fact, to the best of the authors' knowledge, there is still no empirical confirmation of Jackson's theory, empirical studies which would follow a country's diversity demand over various years. MOON ET AL. (2002) observe the impact of different time spans in their diversity study for Bulgaria. Results demonstrate that weekly and monthly food demand does not exhibit different consumption patterns, but that daily variety is different from other consumption time frames. Their study does not consider longer than monthly time spans. This paper aims at considering demand for food diversity at more than one point in time.

Since most of the research on food diversity has been conducted in the United States (JACKSON, 1984; STEWART and HARRIS, 2005), there is a lack of empirical evidence on the demand for food diversity in Canada. Even if the consumption structures between the US and overall Canada are similar, it is possible that differences between the Canadian provinces might exist, due to differences in resource endowments and socio-demographic characteristics.

As introduced above, studying food diversity is important as it gives a picture of the economic well-being of a country. This paper has three objectives considering the empirical and methodological research gaps noted above. Since evidence on the demand for food diversity in Canada is missing, a first objective of this paper is to present a Canadian diversity study, testing extensively for regional and household level differences in the demand for food diversity based on Jackson's hierarchical model of consumer demand. The second and third objectives addressed in this paper are methodological. With the second one, this paper provides evidence on changes in Canadian demand for food diversity over time using a pooled sample. Although analysing the demand for food diversity with panel data would be desirable, panel data for individual consumption behaviour remain quite rare and expensive. Thus, the next step taken in this paper is to compare changes in the cross-sectional demand over time. The third objective is to observe whether the impact of independent variables is different at different points of the distribution of the dependent diversity measure. To the best of the authors' knowledge, this is the first study to use quantile regressions to analyse food diversity demand in this way.

The remainder of this paper is as follows. Section 2 explains the diversity measurement and the data used. The empirical strategy is explained in section 3. Results are presented in section 4. The paper ends with conclusions in section 5.

2 Diversity measurement and data description

To analyse the Canadian household demand for diversity, food diversity is measured with the Berry-Index (BERRY, 1971). While there are many different measures to analyse food diversity (an overview of measures together with an exploration of advantages and disadvantages can be found in DRESCHER, 2007), especially the Berry-Index has been applied frequently in economic studies (THIELE and WEISS, 2003). To obtain results that are comparative to previous economic diversity studies, the Berry-Index is applied here based on the following equation:

$$(1) \qquad BI = 1 - \sum_{i=1}^{n} s_i^2$$

Where in this paper, s_i is the consumption share of food i in grams with reference to the total consumption bundle. Due to the definition of the Berry-Index, diversity is higher the more foods are eaten in equal (quantitative or expenditure) proportions of all foods eaten. Since the Berry-Index lies in the boundaries of 0 and 1-1/n (with n = number of food items distinguished in the data set), the Berry-Index can only be compared over different data sets provided that the number of foods distinguished (not eaten) is the same. It has to be noted that the values of the Berry-Index are difficult to interpret in absolute terms. Higher values indicate a higher level of diversity. Given that the number of foods distinguished in data sets is identical, the Berry-Index values can then be compared.

The data used to analyse Canadian demand for food diversity is the Family Food Expenditure Survey provided by Statistics Canada (STATISTICS CANADA, 2008). Observing changes in demand over time is possible by using three different cross-sectional data sets from 1984, 1996 and 2001. In order to make the Berry-Index comparable over all years, the foods distinguished in each year need to be the same. Therefore, certain foods that were only listed in one or two but not all of the years had to be excluded from each of the datasets. 9.2% of foods have been deleted from the 1984 data set, 5.38% for 1996 and 1.68% for 2001, respectively. However, the share of each of the deleted foods in total food consumption is relatively small.[1] Using the daily quantities consumed of the same 176 foods, the two diversity indices are calculated for 5,360 households in 1984, 10,459 households in 1996, and 4,885 households in 2001. Apparently, sample sizes vary considerably over the years. As noted above, the Berry-Index lies in the range of 0 and almost 1. In order to assure the assumption of normality, the logit transformation is used which applies a 0-1 bound variable into a variable bounded between minus and plus infinity (GREENE, 1997). After the logistic transformation, it is possible to run an OLS regression. The transformation of the Berry-Index is achieved by

[1] Mean quantitative share of deleted foods from 1984, 1996, and 2001: 0.024, 0.011, and 0.002, respectively.

$$(2) \qquad TBI = \ln\left(\left(BI / \left(1 - BI\right)\right)\right).$$

Each sample of the Canadian Family Food Expenditure Survey contains various socioeconomic characteristics of the household. Table 1 gives the descriptive statistics of the variables as used in this paper.

Table 1: Descriptive statistics of variables used in OLS and quantile regressions (pooled sample, 1984, 1996 and 2001)

Variable	Description	Mean (Std.dev)	Min (Max)
Dependent variables			
BI	Berry-Index	0.88 (0.11)	0.02 (0.98)
TBI	Transformed Berry-Index	2.22 (0.77)	-3.90 (3.83)
Independent variables			
HHSIZE	Number of people in household	2.65 (1.36)	1 (10)
lnHHSIZE	Log of HHSIZE	0.83 (0.55)	0 (2.30)
REALINC	Real income (personal income in past 12 months, income of all family members)	30,440 (21,780)	204 (577,344)
lnREALINC	Log of REALINC	10.07 (0.76)	5.32 (13.27)
AGE	Age of reference person	47.06 (16.03)	20 (80)
lnAGE	Log of AGE	3.79 (0.35)	3.00 (4.38)
MALE	Reference person is male	0.52 (0.50)	0 (1)
FEMALE	Reference person is female	0.48 (0.50)	0 (1)
FAFH	No. of meals purchased at restaurants locally and on day trips and overnight and longer (food away from home)	3.89 (4.81)	0 (68.15)
lnFAFH[a]	Log of FAFH	-0.77 (3.60)	-6.91 (4.22)
SINGLE	Single household	0.15 (0.35)	0 (1)
NON-SINGLE	Household is a married or common-low household or non-single household	0.85 (0.35)	0 (1)
ATLANTIC	Household is living in Newfoundland and Labrador, Prince Edward Island, Nova Scotia, or New Brunswick	0.19 (0.39)	0 (1)
QUÉBEC	Household is living in Québec	0.18 (0.38)	0 (1)
PRAIRIES	Household is living in Manitoba, Saskatchewan, or Alberta	0.25 (0.43)	0 (1)
ONTARIO	Household is living in Ontario	0.26 (0.44)	0 (1)
BC	Household is living in British Columbia	0.12 (0.32)	0 (1)
YEAR1984	Household surveyed in 1984	0.26 (0.44)	0 (1)
YEAR1996	Household surveyed in 1996	0.51 (0.50)	0 (1)
YEAR2001	Household surveyed in 2001	0.24 (0.43)	0 (1)

[a]A value of 0.001 has been added to FAFH to avoid that missing values occur when logarithmising FAFH=0.

Source: Own table

3 Empirical strategy

Before analysing the demand for diversity with statistical hypothesis testing, descriptive statistics will be used to provide first insights on Canadian household demand for diversity. The next step in the empirical strategy includes a multiple OLS regression to observe whether socioeconomic determinants in Canada are similar to those found in previous studies. The Berry-Index functions as a dependent diversity variable (*D*) that is explained by a set of socio-economic variables. The demand is estimated for a pooled sample over three years (1984, 1996 and 2001) and has the following form:

$$(3) \qquad D = \alpha_0 + \beta_i \mathbf{x} + \varepsilon$$

where **x** represents a vector of independent variables such as regional and household characteristics and ε is the error term. α describes the constant coefficient and the β's reflect the estimated coefficients for the independent variables.

In addition to OLS, quantile regressions are applied. Since OLS estimators reflect the impact of independent variables at one point of the distribution of the dependent variable, the conditional mean (HAO and NAIMANM 2007), any information obtained by OLS regression is limited to this point of the distribution. This can lead to imprecise (cf. KANDPAL and MCNAMARA, 2009), or at least incomplete findings. It is possible that the effects of independent variables are different at other points of the distribution and therefore quantile regressions have been proposed. Quantile regressions have their roots in the mid-18th century (KOENKER and HALLOK, 2001) and became known first as median regression. The term quantile regression was introduced by KOENKER and BASSETT (1978). It is a robust alternative to OLS with error terms not following a normal distribution. Instead of least squares estimates, median regression is based on least absolute distance regression and as such requires much higher computational power. Quantile regressions have frequently been applied in social science studies, especially poverty or inequality studies (HAO and NAIMAN, 2007; KOENKER and HALLOK, 2001). For many research questions, information about the effect of the independent variables on special decentral locations of the dependent variable instead of the mean is decisive, e.g. in tax studies to compare the effects on the poor (lower tail) versus rich (higher tail) (see also KANDPAL and MCNAMARA, 2009). The median itself describes the most central location of the distribution but any other noncentral locations, expressed as the p^{th} quantile, can be used to describe the relationship between dependent and independent variables at any point of the distribution (HAO and NAIMAN, 2007). Symmetric weights are used for the median (0.5) whereas for all other quantiles (e.g., 0.1 or 0.9) asymmetric weights are employed (ZIETZ, ZIETZ and SIRMANS, 2008).

A couple of other recent papers have used quantile regressions in other contexts than diversity (ARIAS, HALLOCK and SOSA-ESCUERDO, 2001; SCHINDLER RANGVID, 2007). For food diversity, the p^{th} quantile regressions can be expressed accordingly as (BUCHINSKY, 1998; SCHINDLER RANGVID, 2007):

$$(4) \qquad D = x'_i \beta_p + u_{pi} \qquad Quant_p(D|x_i) = x'_i \beta_p \ \ p \in (0,1)$$

D describes the diversity indicator and $Quant_p(D|x_i)$ is the quantile of D. It becomes clear that the error term distribution is not specified but it is assumed that the quantile restriction $Quant_p(D|x_i) = 0$ is fulfilled for the error term. The special feature of the quantile regression approach is that the coefficients of the independent variables (β_p) can differ across quantiles (see SCHINDLER RANGVID, 2007).

Results of estimating (3) and (4) are presented in section 4.2. (4) is estimated for the selected percentiles 0.1q, 0.25q, 0.5q, 0.75q and 0.9q using Stata's QREG command. In addition, simultaneous-quantile regressions are estimated. These differ from quantile regressions only in that the equations are estimated simultaneously instead of separately. Using Stata's SQREG procedure, the entire variance-covariance matrix is obtained with 500 bootstrap repetitions. It is then possible to test the equivalence of coefficients for an independent variable across quantiles.

4 Empirical results

4.1 Descriptive statistics for Canadian household demand for food diversity

The first results of household demand for food diversity in Canada are introduced using descriptive statistics. Regional differences are described first. Canada has ten provinces and three territories. However, in the Canadian Family Food Expenditure Survey, no data is available for the territories. Also, in the 2001 survey the 10 provinces were aggregated to five regions: Atlantic Provinces (including Newfoundland and Labrador, Prince Edward Island, Nova Scotia, and New Brunswick), Québec, Ontario, Prairies (including Manitoba, Saskatchewan, Alberta), and British Columbia. This restricts the analysis to those five regions for all years. Figure 1 shows the extent of diversity by province in the pooled sample.

Figure 1: Mean Berry-Index by Canadian province over years (1984, 1996 and 2001).

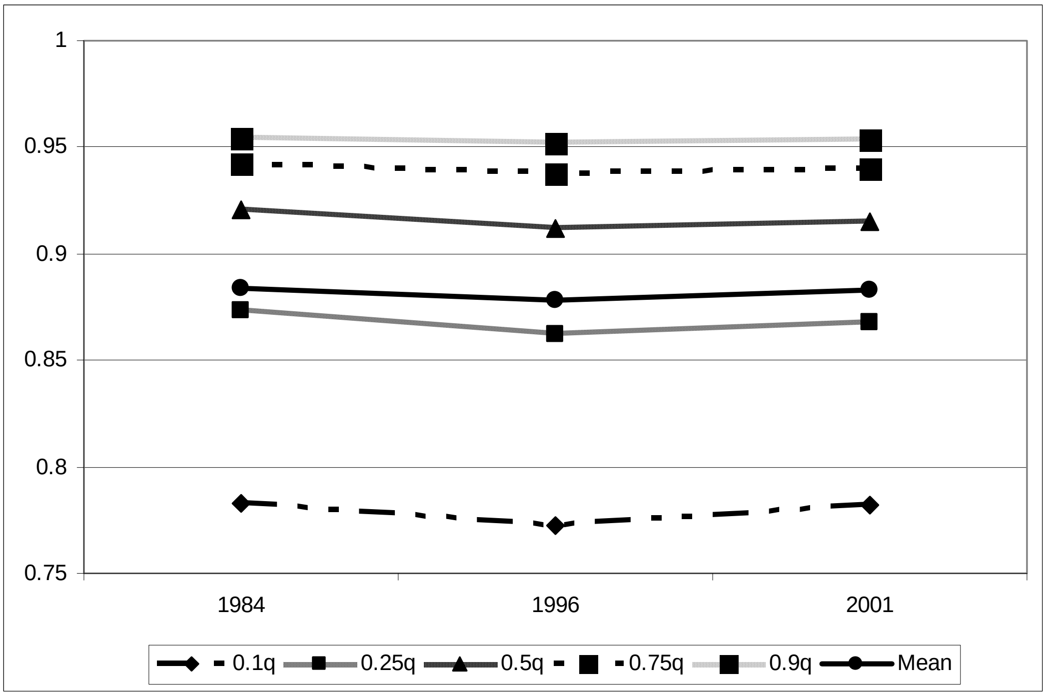
Source: own figure

The extent of diversity is given in four categories. The categories are calculated applying the method of natural breaks using the so called Jenks-Caspall Algorithm. With this algorithm, differences within categories are minimized but they are maximized between classes. The algorithm establishes classes so that the standard deviation within a class is as low as possible. Diversity scores are highest in the cross shaded provinces. Diagonal lined provinces indicate a medium demand for diversity. The dotted category reflects provinces with the lowest demand for diversity. The blank category displays regions for which no data is available (i.e. Yukon Territories, Northwest Territories, and Nunavut).

Figure 1 shows that across the years 1984, 1996 and 2001, the Berry-Index is lowest in the Prairies. Ontario and the Atlantic Provinces have medium Berry-Index values. The Berry-Index is highest in British Columbia and Québec.

To demonstrate the importance of looking at other points of the diversity distribution than at the mean, the following figures illustrate selected quantiles as well as the mean for both diversity indicators separately for all three years.

Figure 2: Berry-Index at different points of the distribution in 1984, 1996, and 2001



Source: own figure

Figure 2 shows that the diversity indicators have in fact extremely different values at decentral locations of the dependent variable compared to the mean which suggests that there could be different influential factors at those points. There is an increase in diversity from the lowest to the highest quantiles. Diversity was higher in 1984 than in 1996 or 2001. Exactly, the deepest points are reached in 1996. It recovered somewhat in 2001 albeit not reaching the original diversity level from 1984. Noticeable is further that the graphs for the higher quantiles (0.5q, 0.75q, 0.9q) are almost flat. This leads to following hypothesis: the decrease in diversity is highest for households with the lowest demand for diversity.

4.2 Determinants of Canadian household demand for food diversity

OLS and quantile regressions are estimated and shown in Table 2. Directions and significances are similar between OLS and quantile regressions which underscore the robustness of the results. The best model fit was obtained by taking logarithms of all metric variables in the equation. Thus, the estimated coefficients are to be interpreted as percentages. Table 2 reports coefficient estimates for the transformed Berry-Index based on the pooled sample (1984, 1996 and 2001).[2] For comparison reasons, OLS estimates are presented together with quantile regression estimates.

[2] Single year regressions for the transformed Berry-Index vary only slightly from the pooled regression models. They are available from the authors upon request.

Table 2: (Quantile) regression coefficients: determinants of the transformed Berry-Index (TBI), pooled sample (1984, 1996, 2001) (n = 19,540)[3]

	OLS	**Quantile**				
	Mean	0.10	0.25	0.50	0.75	0.90
Constant	-0.327** (-2.98)	-3.384** (-11.09)	-1.302** (-6.31)	-0.014* (-0.09)	1.067** (9.47)	1.624** (15.12)
lnHHSIZE	0.264** (21.33)	0.363** (9.74)	0.252** (10.48)	0.205** (12.02)	0.191** (15.34)	0.174** (14.92)
lnREALINC	0.127** (15.17)	0.246** (9.99)	0.181** (11.35)	0.136** (11.72)	0.091** (10.75)	0.083** (10.49)
lnAGE	0.324** (18.28)	0.556** (12.20)	0.358** (11.03)	0.263** (10.80)	0.183** (10.07)	0.127** (7.21)
MALE	-0.067** (-6.05)	-0.107* (-3.47)	-0.089** (-4.32)	-0.067** (-4.44)	-0.063** (-5.61)	-0.057** (-5.33)
lnFAFH	0.014** (8.60)	0.015** (3.47)	0.013** (4.30)	0.007** (3.35)	0.005* (2.83)	0.004* (2.87)
SINGLE	-0.058** (-3.20)	-0.016 (-0.33)	-0.108* (-3.24)	-0.055* (-2.25)	-0.050* (-2.51)	-0.054** (-3.12)
ATLANTIC	-0.042* (-0.03)	-0.084 (-1.55)	-0.052* (-1.48)	-0.022 (-0.85)	-0.017 (-0.88)	-0.001 (-0.05)
QUÉBEC	0.077** (3.90)	0.187** (3.94)	0.123* (3.87)	0.084* (3.58)	0.078** (4.50)	0.058* (3.59)
PRAIRIES	-0.138** (-7.43)	-0.221** (-4.79)	-0.185** (-5.91)	-0.099** (-4.34)	-0.086** (-5.12)	-0.068** (-4.31)
ONTARIO	-0.101** (-5.46)	-0.152** (-3.27)	-0.111** (-3.60)	-0.091** (-4.03)	-0.073** (-4.45)	-0.061** (-4.01)
YEAR1996	-0.113** (-8.73)	-0.079* (-2.42)	-0.109** (-4.87)	-0.125** (-7.57)	-0.085** (-6.95)	-0.078** (-6.73)
YEAR2001	-0.077** (-4.98)	-0.090* (-2.37)	-0.093** (-3.61)	-0.097** (-5.18)	-0.083** (-5.93)	-0.062** (-4.72)
(Pseudo) R^2	0.100	0.072	0.062	0.052	0.049	0.048

Note: t-values are displayed in parentheses. ** = p<0.01, * = p<0.05. Analyses adjusted for weights. OLS t-values are based on heteroscedasticity consistent estimates of the covariance matrix using White's method. The standard errors, and, therefore, the t-values for quantile regression coefficients are obtained using KOENKER and BASSETT's (1982) method. The coefficient of determination (R^2) for the quantile regressions are pseudo R^2.

The F-value of the OLS regression is 177.06 and highly significant at 1% level. The coefficient of determination reveals that the independent variables in the OLS regressions explain 10% of the variance of the transformed Berry-Index. Quantile regressions Psuedo-R^2's lie between 5% and 7% and are also lower than OLS regression R^2. Other quantile regression studies report comparable differences between OLS regression R^2 and quantile regression Pseudo-R^2 (e.g. VARIYAM, BLAYLOCK and SMALLWOOD, 2002).

The constant in the regression tables indicates the reference household with a female reference person who is 47 years old. The household is a non-single, three person household (mean household size is 2.65) with a mean real income of 30,439.54 Can$ per year. The household purchases, on average, 4 meals away from home (mean of 3.89). The household lives in British Columbia and was surveyed in 1984.

The first notable effect of quantile regressions shows in the sign change of the constant coefficient. The constant is negative and significant at the mean (p<0.05) (OLS). The sign change occurs at the 0.75q, where the coefficient becomes positive (p<0.05). At the 0.9q, the constant is also positive (p<0.01). This reflects that there is a break in the effect of the base

[3] For the quantile regressions, standard errors (not shown) have to be regarded as inflated due to sampling issues. However, the magnitude of the coefficients is not inflated.

level of diversity: only if the reference household is located at the upper tails of the distribution is the base level demand of diversity positive.

An increase in household size of 1% increases the Berry-Index about 26% based on OLS. This relationship is in line with previous studies reporting that a higher number of people in the household[4] increases the demand for diversity (LEE 1987; MOON ET AL. 2002). Looking at the quantile regressions, it becomes clear that this positive effect is higher at the lower tails of the distribution of the dependent variable. At 0.1q, the effect amounts to 37% but steadily declines with higher quantiles demonstrating that the positive effect of household size is getting smaller in magnitude for more diversified households.

There is a positive effect of (real) income on diversity throughout the distribution of the dependent variable. A 1% increase in (real) income increases the demand for food diversity in Canada about 13%. This impact is significant at 1% level in all regressions presented.

Again, quantile regressions show that the effect of income is higher at the lower ends of the distribution. Households with the 10% lowest demand for diversity (0.1q) show a 22% increase in diversity generated by a 1% increase in real income. Contrary, households with the 10% highest demand for diversity (0.9q) increase their demand for diversity by only about 7% in response to 1% more real income. Thus, those households already demanding the highest diversity are less affected by increasing income. A similar pattern is observed for the variable AGE. As people age, their Berry-Index increases considerably, and increases even more for low-diversified households (58% increase at 0.1q versus 11% increase at 0.9q).

The Berry-Index of male reference persons in the household is 7% lower than that of females (reference household) at the mean of the distribution of the Berry-Index (OLS). This effect remains highly significant at all considered points of the distribution of the Berry-Index. The negative effect has a higher magnitude for the low-diversified households. Also LEE (1987) shows that females have a higher demand for food diversity than males in the US.

According to Statistics Canada, Canadian households spent 27% more on FAFH from 1997 to 2003 (STATISTICS CANADA, 2006). The Canadian Food Expenditure Survey includes a count of the total number of meals that are purchased away from home. The food items from the FAFH meals are not included in the diversity measures but using this variable as a proxy for FAFH shows that a 1% increase in number of meals consumed away from home increases the diversity of foods at home also about 1% (OLS). The positive effect of FAFH is 3% and 2% at the 0.1q and 0.25q, respectively and evens out at 1% at the remaining higher quantiles. The coefficients are always significant ($p<0.01$) with the exception of the 0.9q, where the impact of FAFH is less precise ($p<0.05$). Thus, eating away from home more frequently tends to increase the diversity consumed at home especially as one moves down the distribution of diversity.

Previous diversity studies have also shown that being single results in a lower demand for diversity as compared to non-single households (reference household) (DRESCHER ET AL., 2009) which is confirmed here by OLS and quantile regression estimates. A single hosuehold's Berry-Index is 6% lower than that of non-single households with OLS regression. At 0.1q, the effect appears insignificant at conventional levels. The negative effect of SINGLE declines from 9% at 0.25q to 5% at the 0.75q ($p<0.05$). At 0.9q, SINGLE reduces the Berry-Index about 6%. Thus, being single has a higher negative effect on diversity at the ends of the diversity distribution. One might question the introduction of both SINGLE and HOUSEHOLD SIZE in the models. They are included here, likewise to previous diversity

[4] To avoid multicollinearity among independent variables, the variable "number of children in the household" has been deleted from the model in favour of HHSIZE. Pearson's correlation coefficient between household size and number of children is 0.8. There is no harmful correlation (Person's correlation coefficient = 0.33) between household size and being married or common-law (NONSINGLE) or SINGLE respectively.

studies, as they both deliver important insights into the demand for diversity. SINGLE allows a better understanding of consumption patterns of a special group of households that has been steadily increasing in numbers in developed countries over the last decades. Household size allows capturing household economies of scale and household composition effects. By definition, single households and household size must be correlated, but both variables have been included only after controlling that multicollinearity does not impose a problem to the model (see footnote 4).

There is a difference in demand for diversity between households in the Atlantic region and households in BC (reference household) with OLS regression. In the Atlantic regions, it is lower. Also the 0.25q shows a a significant negative effect of ATLANTIC ($p<0.05$). Québécois households Berry-Index is 8% higher than that of households in BC. This positive effect is captured across all quantiles but is most precise at the 0.1q of the distribution. The magnitude of this effect declines from the lower to the upper end of the distribution (15% at 0.1q to 4% at 0.9q). At the mean of the distribution, the demand for diversity in the PRAIRIES is 14% lower than in BC. The size of this effect amounts to 24% and 19% at the 0.1q and 0.25q, respectively. The impact of living in the Prairies has on diversity declines at higher ends of the distribution. This means that the negative effect of PRAIRIES is even worse for households with the lowest demand for diversity already. A similar pattern is observed for households in ONTARIO. Reasons for the diversity differences across provinces may be different economic conditions, seasons as well as cultural differences between the provinces. However, more research is necessary to clearly identify those conditions.

The reason for running a combined regression over all years is to observe in which year the demand for food diversity has been highest using yearly dummy variables. OLS results show that compared to the reference year 1984, the demand for diversity is indeed 11% lower in 1996 and 8% lower in 2001 ($p<0.01$) confirming Figure 2. For 1996, the year effect is higher in the middle parts of the distribution (13% each at 0.25q and 0.5q), but is lower at the ends of the distribution (11%, 11% and 9% at 0.1q, 0.75q, and 0.9q, respectively). It is not true that the decrease in diversity is highest for households with the lowest demand for diversity as was hypothesised in section 4.1. Instead and unlike the effect of other independent variables, the decrease in diversity is higher among the households located in the middle of the diversity distribution (8% each for the 0.25q, 0.5q and 0.75q), whereas the decrease reduces to 7% at the 0.9q. In summary, the Berry-Index of households in the middle quantiles degrades even more in 1996 compared to households located at the tails of the diversity distribution. In 2001, the decline in diversity compared to 1984 is also highest in the middle quantiles and the decline is lower only at the upper tails.

In addition to the quantile regressions presented in Table 2, simultaneous-quantile regressions have been estimated for the 0.1q, 0.25q, 0.5q and the 0.75q. Due to space constraints, the results of these estimates are not shown but are available from the authors upon request. The results of the simultaneous quantile regressions vary only slightly to the quantile regressions presented above. According to ARIAS, HALLOCK and SOSA-ESCUERDO (2001) and FATTOUH, SCARAMOZZINO and HARRIS (2005) the bootstrapping method from the simultaneous quantile regression can be used to build a joint distribution which allows setting up F-Tests for equality of coefficients across different pairs of quantiles. Results of the equality tests are given in Table 3.

Table 3: Transformed Berry-Index: F-Tests for equality of coefficients across quantiles

	0.1q	0.25q	0.5q	0.75q
0.1q				
0.25q	F-value: 8.22**			
0.5q	F-value: 19.84**	F-value: 13.83**		
0.75q	F-value: 27.47**	F-value: 21.87**	F-value: 5.10**	
0.9q	F-value: 32.18**	F-value: 25.21**	F-value: 8.92**	F-value: 2.83**

Note: ** = p<0.01, * = p<0.05. F-Test and p-value based on bootstrapped standard errors with 500 replications.

Based on the F- and corresponding p-values the null hypothesis of equal coefficients among quantiles is always rejected at 1% level. This means that for all comparative pairs of quantiles, the impact of the independent variables is different at different points of the distribution of the dependent variable. The results presented validate the use of the quantile regression approach for studying food diversity.

To summarize, it can be stated that the determinants of Canadian demand for food diversity correspond to a large extent to those found for other developed countries. However, results of the quantile regressions highlight that the marginal effects of independent variables are different at decentral points of the diversity distribution as compared to the mean. Especially remarkable is that the independent variables tend to decline in magnitude as one moves up the diversity distribution. This pattern is different only for the variables SINGLE and the yearly dummy variables. The (negative) effect of SINGLE is higher not only at the lower but also at the higher tails of the distribution. The yearly dummy variables exhibit that the negative effect of 1996 and 2001 is higher at the medium quantiles than at the tails.

Thus, quantile regressions of food diversity give a more complete picture of the demand for diversity for different subgroups of the population and should be preferred or at least complete OLS estimations.

5 Conclusions

This study observes determinants of Canadian household demand for food diversity and examines, for the first time, whether the demand for food diversity has changed over time using a pooled sample of three years of the Canadian Family Food Expenditure Survey (1984, 1996, and 2001). The traditional approach of analysing food diversity at the conditional mean is furthermore extended with quantile regressions. In addition to the traditional OLS regression, food diversity is studied at the two symmetric quantiles 0.1q/0.9q and 0.25q/0.75q as well as at the median (0.5q).

Diversity is measured using the Berry-Index. Confirming previous diversity studies, the household level demand for diversity is positively influenced by (the logarithm of) income, age, and household size among others across all years and quantiles. If the reference person is male or single the demand for food diversity is lower as compared to females and married Canadians.

More interestingly however, quantile regressions show that for most cases the independent variables have a larger influence at the lowest quantile (0.1q) whereas the influence is smaller at the highest quantile (0.9q). For example, looking at the impact of real income in the OLS regression: A 1% increase in real income increases the demand for the Berry-Index about 13%. Quantile regression at 0.1q reveals that this effect has a magnitude of 22% compared to only 7% at the 0.9q. Thus, low diversified households, especially those located at the lower 10% of the distribution are much more affected by increasing real income. This pattern is repeated for almost all independent variables. As an exception, the variable SINGLE and the yearly dummy variables reveal a more complex pattern. The negative effect of living in a single household on diversity is higher not only at the lowest but also at the highest quantile under consideration. The yearly dummy variables reveal that the demand for diversity decreased even more compared to 1984 if the household is located in the middle quantiles. There is a moderate decrease in diversity at the ends of the diversity distribution.

The results presented in this paper highlight the importance of quantile regressions. In addition, this paper clearly shows that the demand for diversity is not stable over time. Diversity is lower in 1996 and 2001 compared to 1984. Given that the demand for diversity is highly correlated with income, and the fact that the demand for diversity decreased from the 1980s to the new century, questions are raised about the income situation of Canadian households. This study's results should encourage future studies on the reasons for the decline in diversity.

Acknowledgements: The authors gratefully acknowledge financial support from the H. Wilhelm Schaumann-Stiftung, Hamburg, Germany and the Co-operative Program in Agricultural Marketing and Business, Canada. We wish to thank two anonymous reviewers and the participants of seminars in Edmonton, Munich, Orlando and in particular Thomas Herzfeld for helpful comments.

References

ARIAS, .O, HALLOCK, K. F. and W. SOSA-ESCUDERO (2001): Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression using Twins Data. In: Empirical Economics 26:7-40.

BERRY, C. H. (1971): Corporate Growth and Diversification. In: Journal of Law & Economics 14: 371-383.

BUCHINSKY, M. (1998): Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. In: Journal of Human Resources 33(1):88–126.

CLEMENTS, K. W., WU, Y. and J. ZHANG (2006): Comparing International Consumption Patterns. In: Empirical Econonomics 31: 1-30.

DRESCHER, L. (2007): Healthy Food Diversity as a Concept of Dietary Quality: Measurement, Determinants of Consumer Demand, and Willingness to Pay. Dissertation. Cuvillier Verlag, Göttingen.

DRESCHER, L., THIELE, S., ROOSEN, J. and G. B. M. MENSINK, (2009): Consumer Demand for Healthy Eating Considering Diversity – An Economic Approach for German Individuals. In: International Journal of Consumer Studies 33: 684 -696.

FATTOUH, B., SCARAMOZZINO, P. and L. HARRIS (2005): Capital Structure in South Korea: A Quantile Regression Approach. In: Journal of Develeopment Economics 76: 213-250.

GREENE, W.H. (1997): Econometric Analysis. Prentice Hall, Englewood, Cliffs, NJ.

HAO, L. and D.Q. NAIMAN (2007): Quantile Regression. London, Sage.

JACKSON, L. F. (1984): Hierarchic Demand and the Engel Curve for Variety. In: Review of Economics and Statistics 66: 8-15.

KANDPAL, E. and P. E. MCNAMARA (2009): Determinants of Nutritional Outcomes of Children in India: A Quantile Regression Approach. Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29, 2009

KOENKER, R. and G. BASSETT (1978): Regression Quantiles. In: Econometrica 46:33–50.

KOENKER, R. and G. BASSETT (1982): Robust tests for heteroscedasticity based on regression quantiles. In: Econometrica 50:43–61.

KOENKER, R. and HALLOK, K. F. (2000): Quantile Regression. In: The Journal of Economic Perspectives 15(4): 143-156.

LEE, J. (1987): The Demand for Varied Diet with Econometric Models for Count Data. In: American Journal of Agricultural Economics 1987: 687-691.

MOON, W., FLORKOWSKI, W. J., BEUCHAT, L. R., RESURRECCION, A. V., PARASKOVA, P., JORDANOV, J. and M. S. CHINNAN (2002): Demand for Food Variety in an Emerging Market Economy. In: Applied Economics 34: 573-581.

SCHINDLER RANGVID, B. (2007): School Composition Effects in Denmark: Quantile Regression Evidence from PISA 2000. In: Empirical Econonomics 33: 359-388.

STATISTICS CANADA (2006): Canadians Spending More on Eating Out. Available via DIALOG. http://www41.statcan.ca/2006/0163/ceb0163_002_e.htm. Acessed 24 Apr 2008.

STATISTICS CANADA (2008): Family Food Expenditure Survey. Available via DIALOG. http://www.statcan.ca/english/Dli/Data/Ftp/foodex.htm. Accessed 13 Aug 2008.

STEWART, H. and J. M. HARRIS (2005): Obstacles to Overcome in Promoting Dietary Variety: The Case of Vegetables. In: Review of Agricultural Economics 27: 21-36.

THEIL, H. and R. FINKE (1983): The Consumer´s Demand for Diversity. In: European Economic Review 23: 395-400.

THIELE, S. and C. WEISS (2003): Consumer Demand for Food Diversity: Evidence for Germany. In: Food Policy 28: 99-115.

VARIYAM, J., BLAYLOCK, J. and D. SMALLWOOD (2002): Characterizing the Distribution of Macronutrient Intake among U.S. Adults: A Quantile Regression Approach. In: American Journal of Agricultural Economics 84(2): 454-466.

ZIETZ, J., ZIETZ E: N. and G. S. SIRMANS (2008): Determinants of House Prices: A Quantile Regression Approach. In: The Journal of Real Estate Finance and Economics 37: 317-333.