Helsinki Center of Economic Research

Discussion Papers

Finding the Functional Forms for Slowly Changing Dependencies

Olli Ropponen
University of Helsinki and HECER

Discussion Paper No. 253
February 2009

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781, E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 253

Finding the Functional Forms for Slowly Changing Dependencies*

Abstract

This study introduces a two step method for finding the functional form for the model to be estimated. The proposed method allows us to get rid of the restrictions of the parallel model even when there are no interaction terms available. It is applicable in the cases where the dependencies change slowly. The method is applied to the data of Finnish Household Surveys.

JEL Classification: C14, C50, E21

Keywords: semiparametric methods, modelling, consumption

Olli Ropponen

HECER
Department of Economics
University of Helsinki
P.O. Box 17
FI-00014 University of Helsinki
FINLAND

e-mail: olli.ropponen@helsinki.fi

* HECER (Helsinki Center of Economic Research), University of Helsinki, P.O. Box 17, FIN-00014 University of Helsinki. olli.ropponen@helsinki.fi. I thank Valtteri Ahti, Mika Kortelainen, Timo Kuosmanen, Markku Lanne, Henri Nyberg, Heikki Pursiainen, Marja Riihelä, Heikki Räisänen, Tuomo Suhonen, Risto Sullström and Yrjö Vartia for the constructive comments and discussions we have had concerning this paper. I also thank the Finnish Government Institute for Economic Research for providing me the possibility to use the data.

1 Introduction

In a number of studies the researchers estimate a model that is linear both in parameters and in variables in order to capture the systematic dependencies. The results can then be summarized by the parameter estimates each of them having a nice interpretation of being an average effect. A drawback of this linear regression model is that it can only detect linear dependencies. In some special cases the dependence between the variables may truly be a linear one, but in most of the cases it is unlikely to be and the linearity is then considered as an approximation for the true dependence. This approximation typically performs well with two points close to each other, but becomes worse as the distance grows. This study introduces a method for finding the functional form for the model to be estimated. It thus provides you a way to proceed in the cases where the linear regression model does a bad job in approximating the true dependence.

The method is composed of two steps. As the first step we follow the footsteps of Hausman and Newey (1995), Schmalensee and Stoker (1999) and Yatchew and No (2001) and estimate a partially parametric model. The second step is concerned about transforming the dependence estimated in the first step. As an application we study the relations between the consumption expenditures and the age, that is the age profiles for consumption. There are two specific research questions we are after. First we want to know the way the consumption expenditures are distributed over the life-cycle. This can be answered by using the results from the first step estimation. Second, we want to know the ways the age profiles for consumption differ between different generations. The answer for this is got from the second step estimation.

For the illustrative purposes we plot the (averages of the) logarithm of consumption expenditures as a function of age for the big generation in figure 1. Here we do not have observations for this generation as old. However, with the method introduced we are able to estimate the regression curve relating the age and the consumption expenditures for the whole life-cycle of the generation. The capability for giving the prediction to the last part of the regression curve arises from the information of the previous generation.

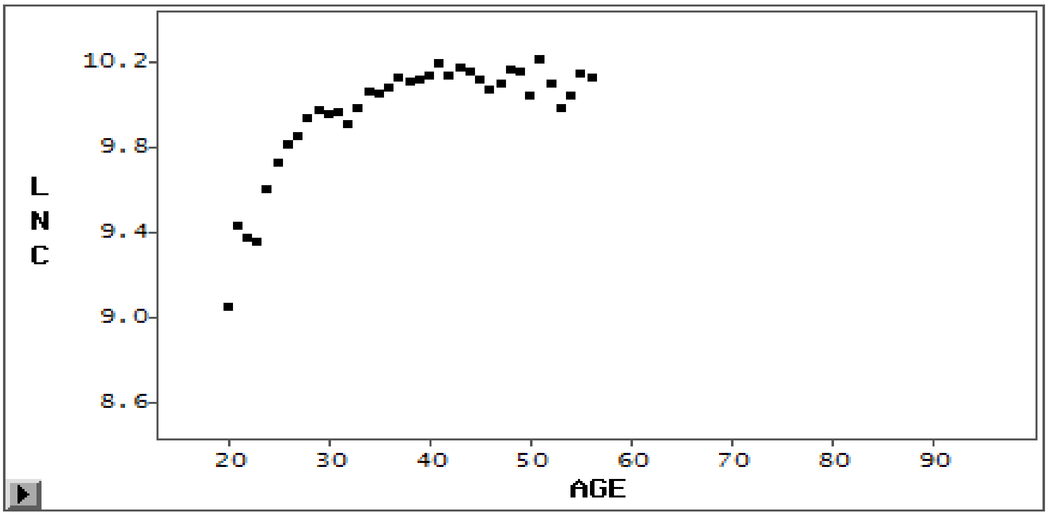

Figure 1: The averages of the observations of the big generation.

In many cases a researcher might have been estimating the parallel model with respect to the generations. In these cases the regression curves for different generations would just be vertical shifts from each others. One way to get rid of this restriction would be by including the interaction terms between ages and generations. In this study, we provide another way to get rid of the restriction. This is applicable even when there are no interaction terms available.

Section 2 illustrates the method by applying it for the Finnish Household Survey data to study the age profiles for consumption. In section 3 we then derive the method in the general. Section 4 provides discussion and section 5 concludes.

2 Aging and Consumption in Finland

2.1 The Data

The data set includes five independent cross-sections of Finnish Household Surveys[1] in 1985, 1990, 1994-1996, 1998 and 2001. The numbers of observations for the survey years are 8200, 8258, 6743, 4359 and 5495 respectively making the total number of observations 33055. The data resembles the Family Expenditure Surveys, which are widely used in studies concerned with the consumption.[2] The studies typically employ these data from about 15 to 20 year time period as is the case also here. The Finnish Household Surveys differ from the Family Expenditure Surveys in the sense that these do not have data from every consecutive year.

The study uses the information on four variables in the data. Three of these, the total consumption expenditures, age and year of birth (cohort), are household specific variables and the fourth one we are using is the survey year. The total consumption expenditures are given in terms of 2001 euro.[3] The age of the household is taken to be the age of the household head and the year of birth of the household is taken to be the one of the household head. The combined data from 1994 to 1996 will be referred from now on as the data from 1995.

[1]The more detailed information about the data can be found in Statistics Finland 2001 and 2004 (Tilastokeskus 2001 and 2004).

[2]For example Lewbel (1991), Härdle and Mammen (1993), Blundell et al. (1994), Kneip (1994), Attanasio and Browning (1995), Banks et al. (1997), Banks et al. (1998), Deaton (1998), Blundell and Duncan (1998), Blundell et al. (1998), Pendakur (1999), Blundell et al. (2003), Stengos et al. (2006) and Blundell et al. (2007) have used these data in their studies.

[3]The data in 1985, 1990, 1995 and 1998 the consumption expenditures are originally in the former currency of Finland, mark. These are first turned into euro and then the nominal values are transformed into real ones in 2001 by the Consumer Price Index.

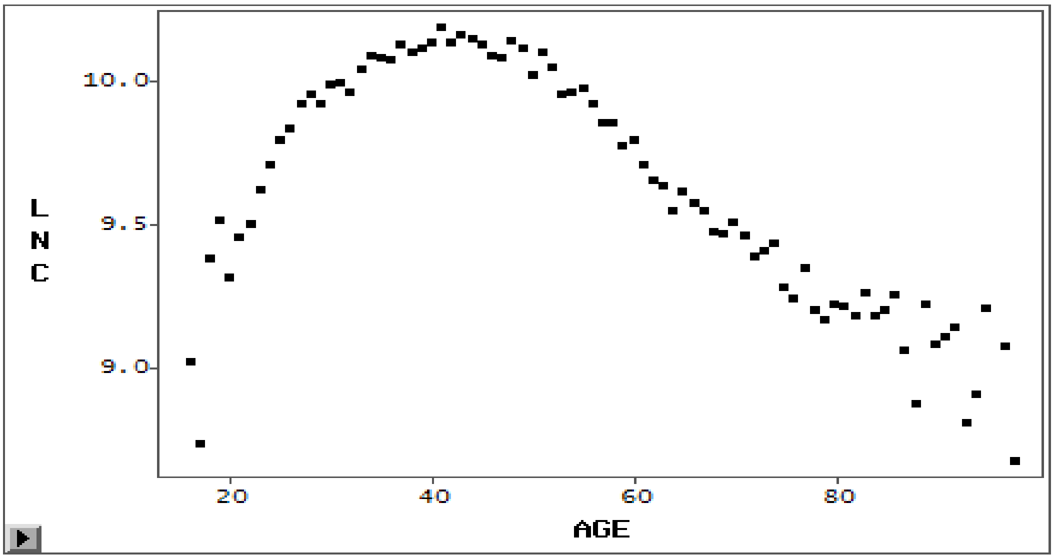

Figure 2: The averages for the logarithms of the consumption expenditures.

In figure 2, we plot the averages of the whole data for the Finnish Household Surveys being used. Here we see that until age about 40 or 45 the consumption expenditures of a household increase and after that they decrease. The age profile for consumption has thus a hump or the inverted U shape. This is also observed in Family Expenditure Surveys. According to Blundell et al. (1994) the consumption initially rises and then falls after mid-forties, which is similar to the case in the data of Finnish Household Surveys used here. Attanasio and Browning (1995) find that the observed shape of the age profile can be explained by the family composition and income over the life-cycle.[4] In figure 3 we plot the averages of the data for each of the survey years. According to this, the age profiles seem to have the hump shape also in every survey year. This shape of the profile seems to be changing slowly in time and the changes might even be close to parallel shifts. In addition to

[4]In this paper we are not trying to find the reasoning behind, but stick in providing good approximations for the age profiles for consumption.

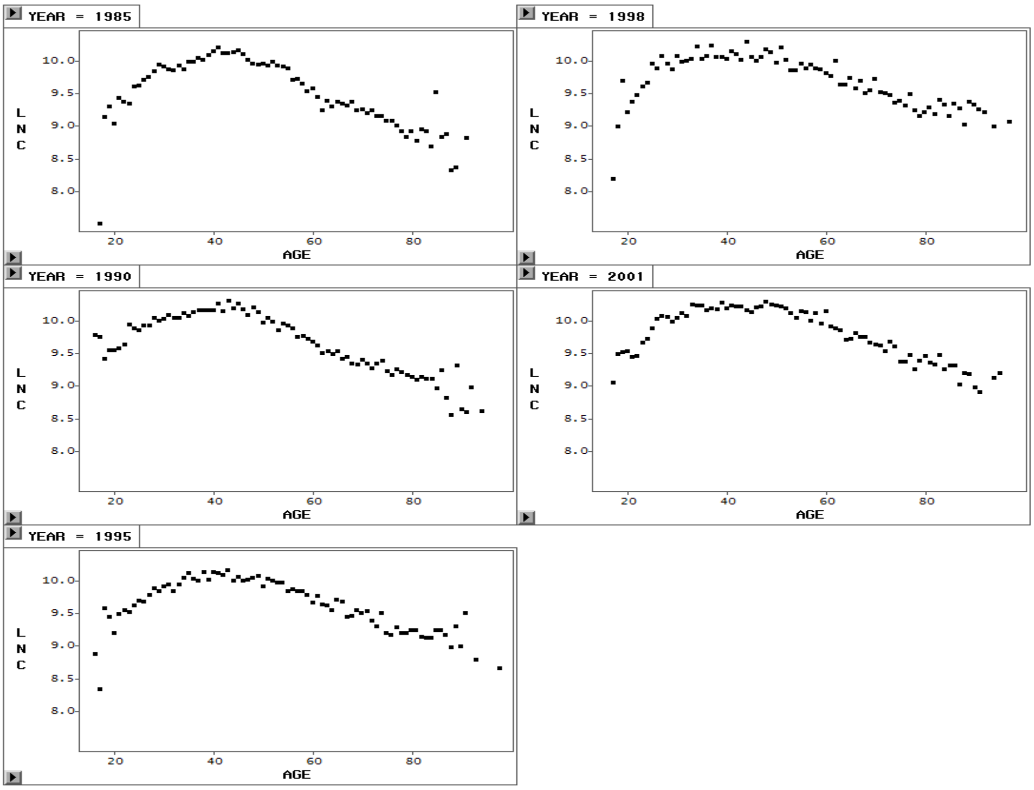

Figure 3: The averages for the logarithms of consumption expenditures for the survey years.

the shape, we also observe the smoothness of the profiles. The small jumps appear just because we have a sample.

We have two research questions to be studied in this paper. The first one is about how the consumption expenditures are distributed over the life-cycle. The second of these is concerned about how do the age profiles for consumption differ between the consecutive generations. Especially we want to know how do the consumption expenditures of the big generation differ from the previous generation as old.[5] For this purpose we define the three

[5]Some studies concerned with the different cohorts use the cohort averages. Here we are not taking that route. Also many of the papers use the adult equivalent scales (see for

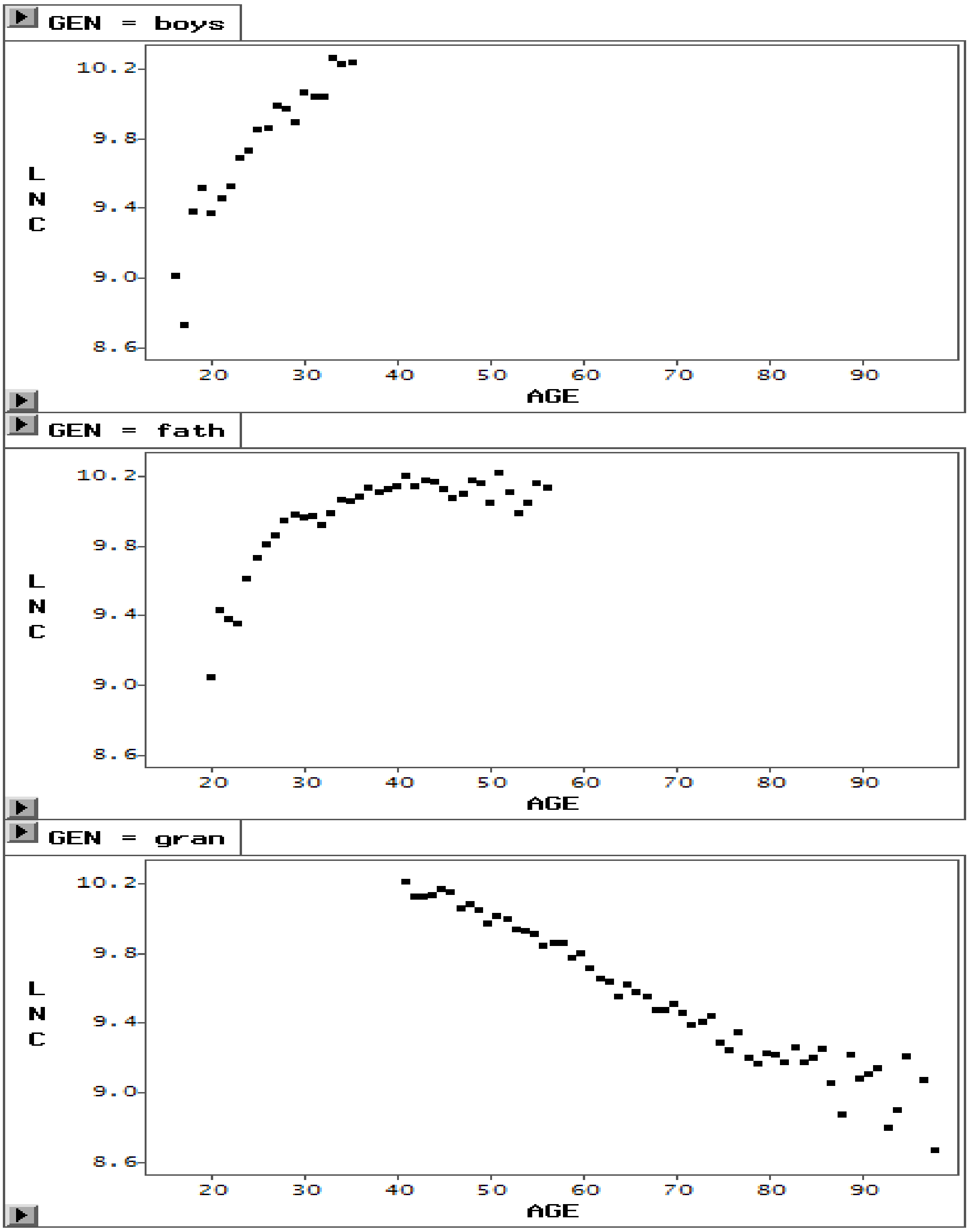

Figure 4: The averages for boys, fathers and grandfathers.

example Lewbel (1989) and Banks (1994) for the underlying reasoning). We are neither following those tracks.

generations that we will call boys, fathers and grandfathers. The households with year of birth from 1965 on are defined to belong to boys, the ones with year of birth from 1945 to 1965 belong to fathers and the ones with year of birth before 1945 belong to grandfathers. The numbers of observations for boys, fathers and grandfathers in the data are 3375, 15646 and 14036 respectively. The averages for the generations are depicted in figure 4. Here we see that we do not have observations for the whole life-cycle for any of the generations, but we can observe only some part of the whole age profile for consumption. Despite of that we still want to compare the whole age profiles between the generations - the comparison is just being done with this imperfect information.

2.2 Testing for the Linearity

In many cases when there is no a priori information about the functional form for the model to be estimated, a researcher estimates a linear regression model and hopes that it provides a good approximation for the true underlying dependence. Let us first check how likely is it that the data would appear from the process that is linear. The tests for the linearity of the regression curve can be done with the specification test[6] with the test statistics

$$V = \frac{\sqrt{n}(s_{res}^2 - s_{diff}^2)}{s_{diff}^2}, \tag{1}$$

where[7]

$$s_{res}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln c_i - \hat{\alpha} - \hat{\beta} age_i)^2 \text{ and}$$

[6]There exists a vast literature on specification testing. Ellison and Ellison (2000) provides a nice summarization and a large number of references on the issue until that time.

[7]Here the observations are ordered by variable age in a way that $age_1 \leq age_2 \leq \ldots \leq age_n$.

$$s^2_{diff} \quad = \quad \frac{1}{2n}\sum_{i=2}^{n}(\ln c_i - \ln c_{i-1})^2. \qquad (2)$$

By the test statistics we compare the residual variance from the linear regression model to the one from the smooth underlying function for the dependence.[8] In addition to the smoothness of the dependence we only need ages to be dense in the domain to be able to perform the test. Under the null hypothesis of the linear model being the correct one the test statistics has asymptotically standard normal distribution, i.e. $V \overset{as}{\sim} N(0,1)$. Under the alternative hypothesis s^2_{res} will overestimate the residual variance and thus the large positive values of the test statistics are the ones that reject the null hypothesis, i.e. we have a one-sided test.

First we test the null hypothesis of linearity for the relation depicted in figure 2, that is for the whole data. The test statistics gets a value 56.7573 and thus we reject the null hypothesis of the functional form being linear. Testing the null hypothesis of linearity for each of the survey years, the dependencies depicted in figure 3, gives us V-statistics 27.5386, 25.1543, 23.9473, 15.8628 and 25.2315 respectively. Thus, the null hypothesis about the linearity at every point in time is rejected. Third we test the linearity of the dependence for our three generations depicted in figure 4 and here we get the V-statistics 3.9970, 6.4562, 7.8823 respectively. So the null hypothesis about the linearity is rejected for each of the generations even now when we do not have observations over the whole life-cycle, but only from some part of that for each generation. Because the linearity does not seem to be the case in some part of the life-cycle, it is unlikely that this would hold for the whole life-cycle. As the linearity is highly rejected it is not wise to estimate the linear regression model. How to proceed then?

[8]By the smoothness we mean here that the first derivative is bounded.

2.3 Step I: The Estimation of the Partially Parametric Model

As the linearity of the age profile is rejected we follow the ideas similar to Hausman and Newey (1995), Schmalensee and Stoker (1999) and Yatchew and No (2001) and estimate the partially parametric model. The specification we are using is

$$\ln c_i = f(age_i) + \gamma_{year_i} + \epsilon_i, \tag{3}$$

where $\ln c_i$ refers to the logarithm of the total consumption expenditures, age_i to the age, $year_i$ to the survey year and ϵ_i to the error term of the household i and f is some smooth function. The data $\{(\ln c_i, age_i, year_i)\}$ are first reordered in an increasing order, that is $age_1 \leq age_2, \ldots, age_{33055}$. Then we take the difference to get[9]

$$\begin{aligned} \ln c_i - \ln c_{i-1} &\approx \gamma_{1985}(\delta_{year_i,1985} - \delta_{year_{i-1},1985}) + \\ &+ \gamma_{1990}(\delta_{year_i,1990} - \delta_{year_{i-1},1990}) + \\ &+ \gamma_{1995}(\delta_{year_i,1995} - \delta_{year_{i-1},1995}) + \\ &+ \gamma_{1998}(\delta_{year_i,1998} - \delta_{year_{i-1},1998}) + \\ &+ \gamma_{2001}(\delta_{year_i,2001} - \delta_{year_{i-1},2001}) + \epsilon_i - \epsilon_{i-1}. \end{aligned} \tag{4}$$

The estimation of this gives us parameter estimates $\hat{\gamma}_{1985} = -0.28556$, $\hat{\gamma}_{1990} = -0.13869$, $\hat{\gamma}_{1995} = -0.00816$, $\hat{\gamma}_{1998} = -0.00246$ and $\hat{\gamma}_{2001} = -0.00001$.

From now on we treat the γ_{year}'s as if they were known[10] and turn to the estimation of a pure nonparametric model

$$\ln c_i - \hat{\gamma}_{year_i} = f(age_i) + \epsilon_i. \tag{5}$$

[9]Here we use the approximation $f(age_i) - f(age_{i+1}) \approx 0$. In the cases when $age_i = age_{i+1}$ this approximation becomes exact, that is $f(age_i) - f(age_{i+1}) = 0$. This happens in all but 81(/33055) cases.

[10]To be the ones we estimated them to be.

Our task is to find a good approximation for the smooth function f. The estimation of f is performed with multiple nonparametric regression techniques combined with multiple choices of the weight functions. This is to guarantee the robustness of our results. The estimations include spline, kernel and loess estimations.[11] The weights employed obey normal, triangular, quadratic and tri-cube distributions. The most crucial choice regarding the performance of the regression curve is the choice of smoothing parameter that tells about how much the dependence is smoothed. If the dependence is smoothed too much the important features of the dependence are eliminated whereas with too little smoothing the data are followed too closely and the predictions for the new data are not that good. A common feature for all the techniques being employed is that the value for the smoothing parameter is chosen by the cross-validation.

16 different estimations for f are performed.[12] The results are depicted one by one in figures 6 and 7 and all of these are pooled in figure 5. From this pooled figure we see that all the regression curves give us very similar results from age about 25 to age about 80 and the differences arise only at the ends of the regression curves, where we do not have as much observations as in the middle. The results from the spline estimation are depicted in the graph in the upper left corner of figure 6. Here we see that it performs like most of the regression curves. The results from the kernel estimations are depicted in the three bottom graphs on the left hand side of the figure 6. From the graphs we see that these regression curves start to wiggle at the right end. This kind of

[11] We also performed the estimation of f by using the local polynomial fitting with 12 different specifications. The results from these are given in appendix A. These give the same qualitative results than the spline, kernel and loess estimations. The same holds also for the 8 parametric specifications employed. These results are given for a request. The reasoning behind the performed nonparametric regressions are given in appendix B.

[12] One estimation of the regression curve is performed by the spline estimator, three with the kernel estimators and 12 with the loess estimators.

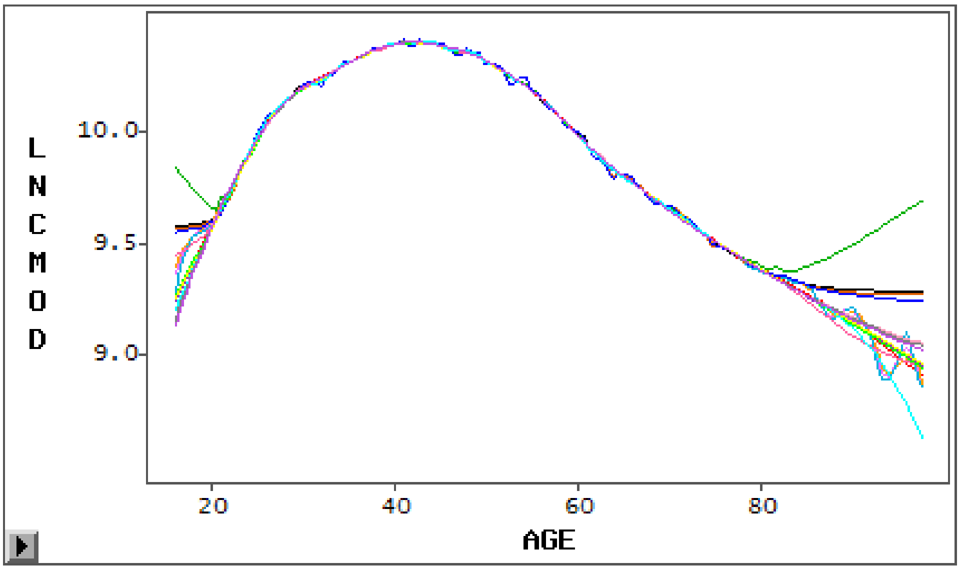

Figure 5: The result from the spline, kernel and loess estimations for f in equation 5.

behavior is not typical for our profiles and arises due to the estimator, not due to the true profile being like that. There exists adjusting kernel estimators that can get rid of this wiggling. We are not going to use them here, but take the behavior of the estimator to be a confirmation about the already known property appearing from time to time in the kernel estimation. The results from the loess estimations using zeroth order local polynomial are depicted on the right hand side in figure 6. The first of these deviates from all the other estimates by having much higher tails and the three other by giving us bumpy estimates for f. In figure 7 we plot the regression curves from the rest of the loess estimations. The graphs on the left hand side use the first order polynomial and the ones on the right hand side use the second order polynomials. Except the graph in the upper right corner all the regression

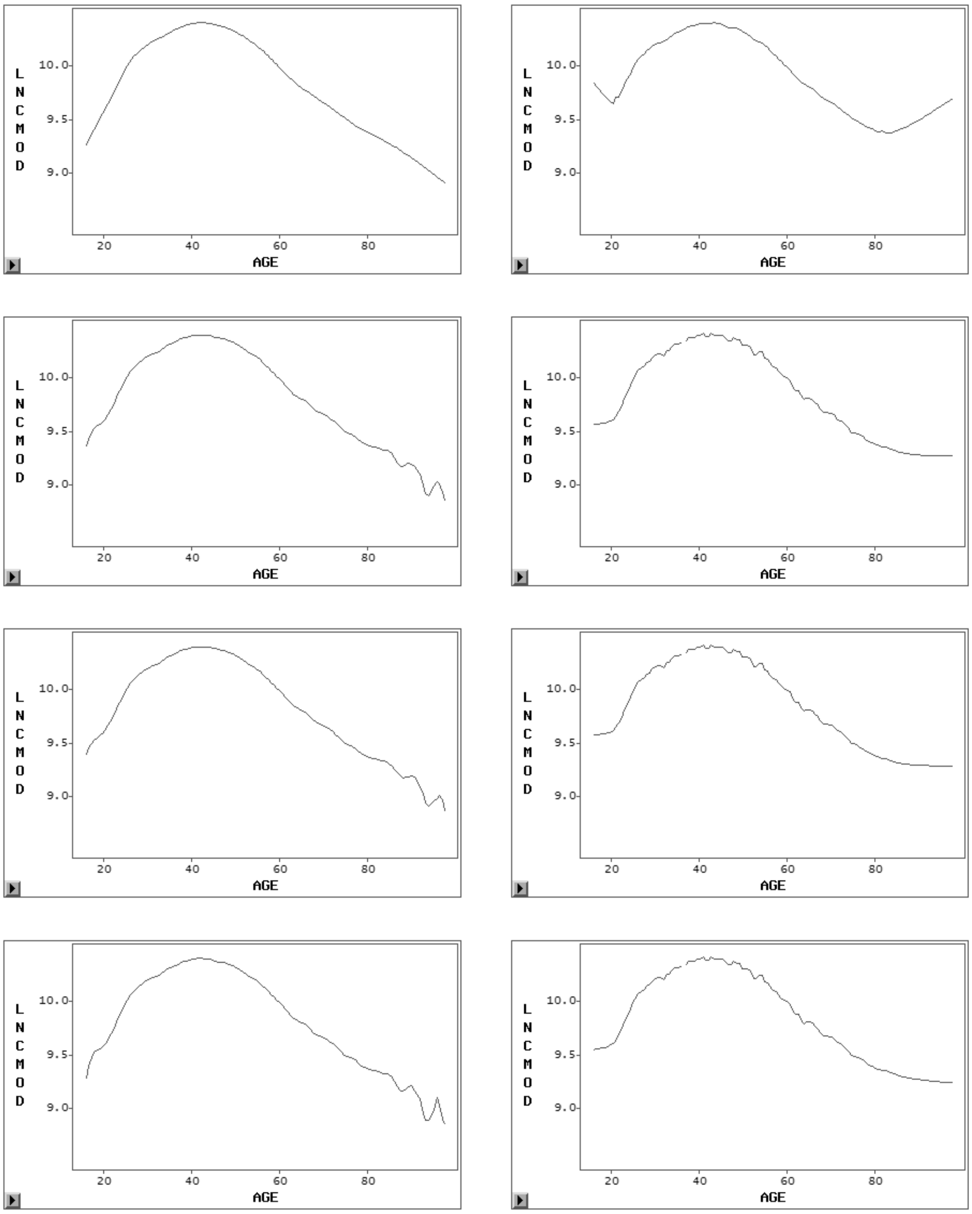

Figure 6: Estimates for f in equation 5. The graph in the upper left corner results from the spline estimation. The last three graphs on the left hand side are from the kernel estimations with normal, triangular and quadratic weights being employed. The right hand side graphs use the loess estimation. Here the means are used and the weights employed are normal, triangular, quadratic and tri-cube respectively.

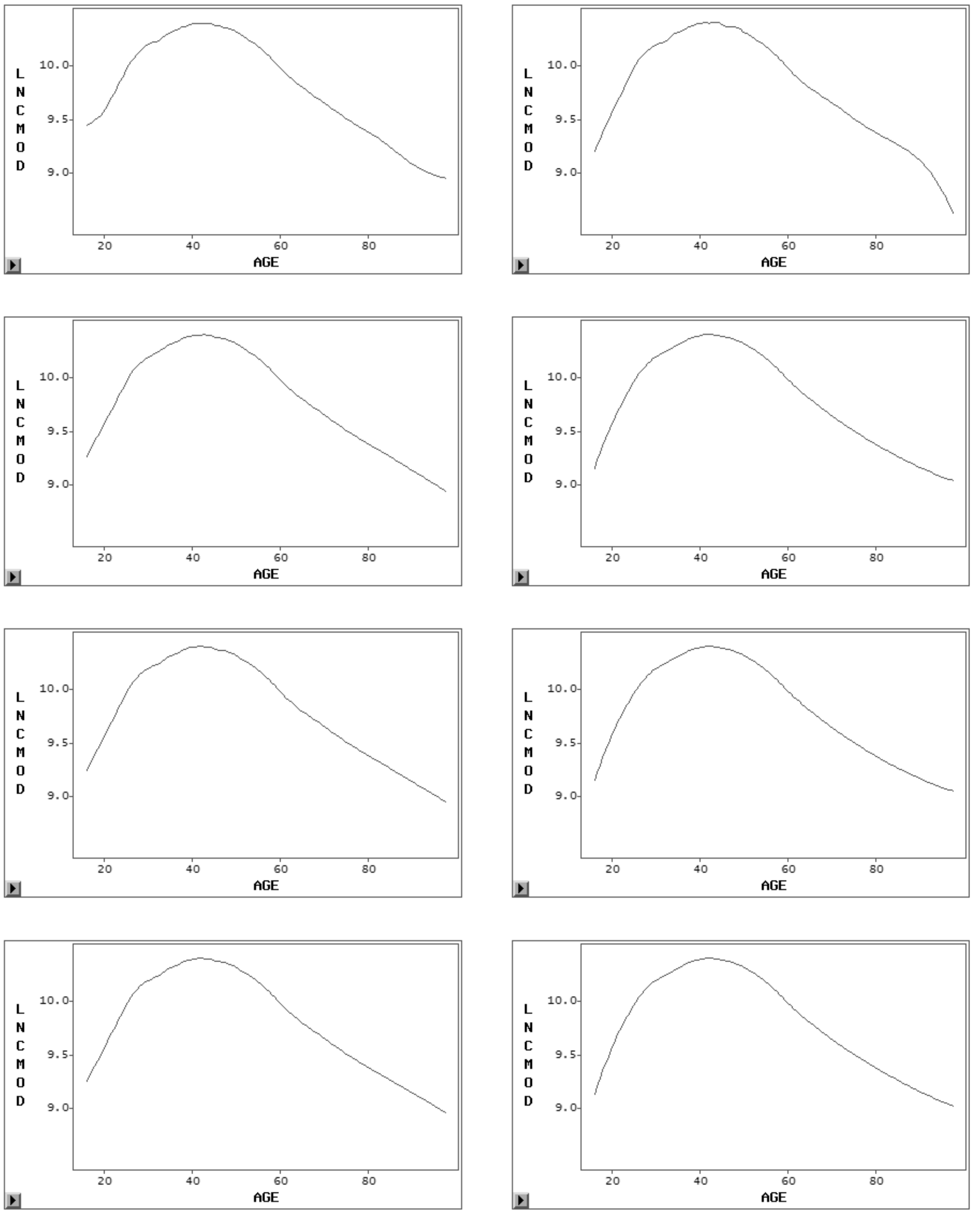

Figure 7: Estimates for f in equation 5. The left hand side graphs use loess estimation with linear dependencies and the ones on the right hand side use the quadratic dependencies. The weights follow normal, triangular, quadratic and tri-cube distributions respectively.

curves here share the typical features with the other regression curves. The deviating one has the right tail lower than that of the other regression curves.

There are (at least) seven regression curves that share almost identical behavior. These arise from the spline estimation and the loess estimations using the first and the second order polynomials with weights being triangular, quadratic and tri-cube. These are depicted in figure 8. Each of these provides a close approximation for the true underlying age profile for consumption, but we will from now on concentrate on the one arising from the loess estimation using first order polynomials with the quadratic weights.

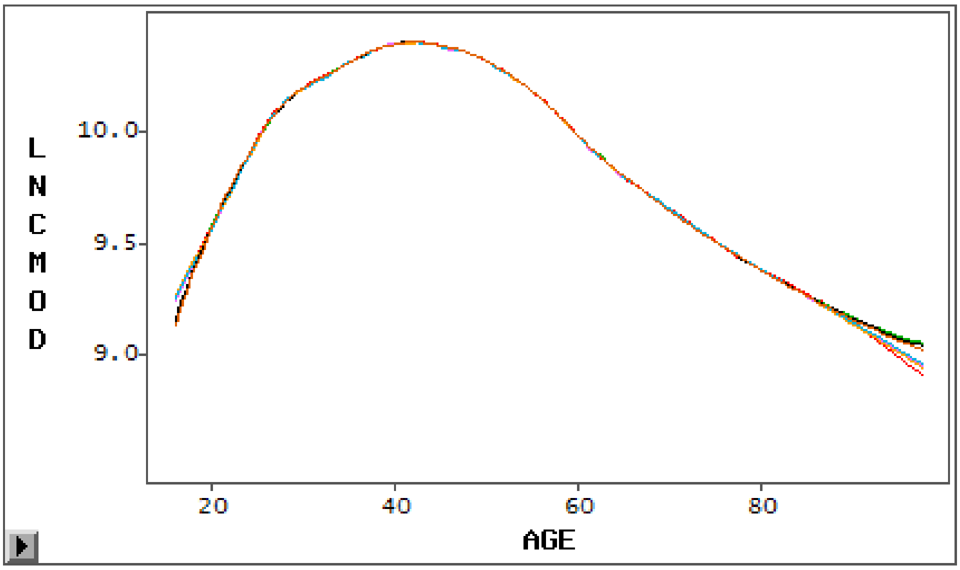

Figure 8: The result from the spline and 6 loess estimations for f in equation 5.

2.4 Step II: The Estimation of the Age Profiles for Consumption for Different Generations

So far we have performed some familiar semiparametric estimations. Now in the second step we deviate from these by giving the regression curve a freedom to be transformed. How the transformation is done is described below.

The regression curves above seem to perform reasonably well in their task of describing the age profiles for consumption. Despite of that there is still an obvious restriction in these. They give the same predictions for every generation. In order to get the individual profiles for each of the generations we allow the general profile appearing from the first step, to transform. As this can be performed in (infinitely) many ways the next question arising is about how do we allow this to happen. What we argued already in the beginning of the study was that the age profiles seem to share the similar looks at every observed time period. Especially, the maximum of the profile stays at the same place at about age 40 or 45.[13] The linear transformation that only scales[14] and vertically shifts the general profile has the property of keeping the maximum at the same age.[15] For this reason we perform in the second step the estimation of the regression model

$$y_i = f_{gen_i}(x_i) + \nu_i = \psi_{gen_i} + \phi_{gen_i}\hat{f}(age_i) + \beta_{year_i} + \nu_i \tag{6}$$

where $\hat{f}(age_i)$ are the predicted values from the first step and ψ_{gen}'s, ϕ_{gen}'s and β_{year}'s are the parameters to be estimated. The parameter estimates are given in table 1. Here we see that both $\hat{\phi}_{fath} < \hat{\phi}_{boys}$ and $\hat{\phi}_{fath} < \hat{\phi}_{grand}$ and

[13]The maximum is achieved in the first step at the age of 42.

[14]By scaling we just mean multiplying by some number.

[15]There already are models that use the idea of preserving some similarity like in Härdle and Marron (1990), Pinkse and Robinson (1995), Pendakur (1999) and Lewbel (2008). There the differences are allowed to be composed of two shifts only - a vertical and a horizontal one (see figure 1 in page 6 in Pendakur for illustration).

$\hat{\psi}_{boys}$	$\hat{\psi}_{fath}$	$\hat{\psi}_{grand}$	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$
−0.1656	0.5831	0	1.0168	0.9441	1.0004
$\hat{\beta}_{1985}$	$\hat{\beta}_{1990}$	$\hat{\beta}_{1995}$	$\hat{\beta}_{1998}$	$\hat{\beta}_{2001}$	
−0.0377	0.0080	−0.0385	−0.0054	0.0727	

Table 1: The results from the second step estimation.

thus the fathers have the most gentle age profile for consumption of the three generations. Combining the general profile and the parameter estimates from the second stage regression allows us to construct the age profiles for each of the generations for each of the survey years. In figure 9 we plot the ones for the survey year 2001. Here the profiles for boys and grandfathers are close to indistinguishable from each others whereas the one for fathers has higher tails. Visually you can observe the difference between the age profiles for the

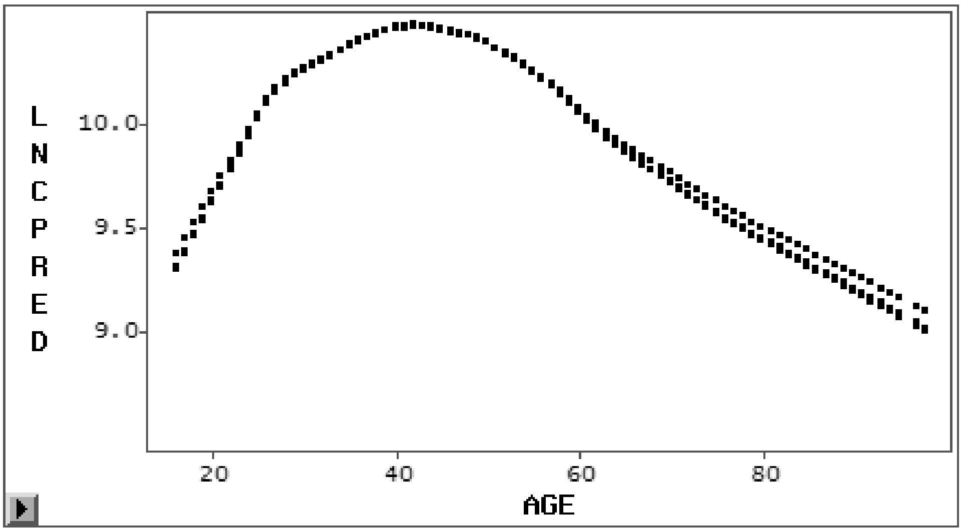

Figure 9: The 2001 age profiles for boys, fathers and grandfathers.

fathers and the grandfathers from the figure. The next question is whether this observed difference is statistically significant. The answer is yes. We test the similarity between ϕ_{fath} and ϕ_{grand} and reject the null.[16] Obviously for each of the generations there are ages that cannot be realized in 2001. Thus the age profiles for consumption for those ages for different generations can answer to counterfactual question about how would these consume if they had that age. This property allows us to compare the different generations. In order to get the idea of how does the difference between fathers and grandfathers show up in euro[17] we plot the difference in figure 10.

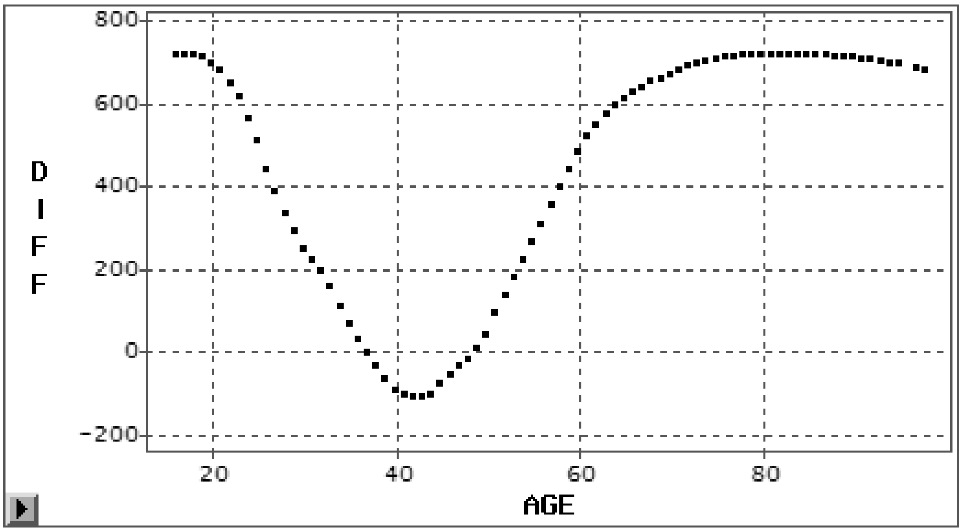

Figure 10: The difference between the age profiles for fathers and grandfathers.

[16]The standard errors for $\hat{\phi}_{fath}$ and $\hat{\phi}_{grand}$ are 0.0322 and 0.01278 respectively. The t-statistics for testing the $H_0 : \phi_{fath} = \phi_{grand}$ gives us 2.253 and thus the null hypothesis about the similarity is rejected at the 5% risk level.

[17]These are given in 2001 euro.

3 The Method

Above we illustrated the method for finding the functional form for the model to be estimated by applying it to the data of Finnish Household Surveys. In this section we introduce this method in general. Suppose we are interested in how an independent variable x and a dependent variable y are related. We allow this dependence to differ between G groups, $g = 1, \ldots, G$. In this case the population regression equation reads as

$$y_i = f_g(x_i) + \epsilon_i \; i = 1, \ldots, n, \; g = 1, \ldots, G \tag{7}$$

and our interest is in estimating the functions $f_1, \ldots, f_G$.

3.1 General Profile

There are (too) many ways that the relations, denoted by f_g, may differ across groups (to say something in general). For this reason we concentrate on a particular types of differences. Suppose we can find a variable m[18] that divides the population into M different groups with the property that the relation between x and y may be written as a partially parametric model

$$y_i = f(x_i) + \gamma_{m_i} + \epsilon_i. \tag{8}$$

Here f is a function which does not depend on m, γ_m's are the parameters to be estimated and ϵ_i is the error term. The variable m classifies the population into M categories and the dependence between x and y differs across two groups only by a vertical shift. If f is a smooth function and x_i's are dense in the domain then we can use an approximation to get the estimates for $\gamma_1, \ldots, \gamma_M$. Let us first rearrange the data $\{y_i, x_i, m_i\}$ in a way that $x_1 \leq x_2 \leq \ldots \leq x_n$. Then writing

$$y_i - y_{i-1} \;\; = \;\; f(x_i) - f(x_{i-1}) + \gamma_1(\delta_{m_i,1} - \delta_{m_{i-1},1}) +$$

[18] This gives another partition than the variable g.

$$
\begin{aligned}
&+ \quad \ldots + \gamma_M(\delta_{m_i,M} - \delta_{m_{i-1},M}) + \epsilon_i - \epsilon_{i-1} \approx \\
&\approx \quad \gamma_1(\delta_{m_i,1} - \delta_{m_{i-1},1}) + \ldots + \gamma_M(\delta_{m_i,M} - \delta_{m_{i-1},M}) + \\
&+ \quad \epsilon_i - \epsilon_{i-1},
\end{aligned} \tag{9}
$$

where δ's are dummy variables ($\delta_{j,k} = 1$ when $j = k$ and 0 otherwise). Estimating this gives us parameter estimates $\hat{\gamma}_1, \ldots, \hat{\gamma}_M$. Let us then proceed as if the population parameters $\gamma_1, \ldots, \gamma_M$ were known (to be $\hat{\gamma}_1, \ldots, \hat{\gamma}_M$). By subtraction we get from 8 that

$$
\tilde{y}_i = y_i - \hat{\gamma}_m = f(x_i) + \epsilon_i, \tag{10}
$$

which is a pure nonparametric model and our task is to find an approximation for the function f. The estimation of f can be performed with standard nonparametric techniques including spline, kernel, loess and local polynomial estimation and we will call the estimate for f, $\hat{f}$, the *general profile*. The estimation of this general profile ends the estimation of the first step.

3.2 Profiles for Different Subgroups

Our focus is on finding good approximations for the functions f_g, $g = 1, \ldots, G$ in equation 7. To do that we use the information from the first step estimation. We allow the functions f_g to be transformations from the estimated general profile $\hat{f}$, that is

$$
y_i = f_g(x_i) + \nu_i = h_g(\hat{f}(x_i)) + \beta_{m_i} + \nu_i. \tag{11}
$$

Again there are (infinite) number of possible functional forms for h_g and here we focus on the linear ones. This means that we have

$$
y_i = f_g(x_i) + \nu_i = \psi_{g_i} + \phi_{g_i}\hat{f}(x_i) + \beta_{m_i} + \nu_i \tag{12}
$$

and our task here in the second step is just to get the parameter estimates for $\psi_1, \ldots, \psi_G$, $\phi_1, \ldots, \phi_G$ and $\beta_1, \ldots, \beta_M$. The nice property of these particular types of transformations is that both the maxima and the minima for the $\hat{f}_g$ are reached at the same x_i's, where $\hat{f}$ reaches its maxima and minima. This is because the general profile is in addition to shifts by $\hat{\psi}_g$ and $\hat{\beta}_m$ just multiplied by a scalar $\hat{\phi}_g$.

4 Discussion

This section gives short discussions about the method, the information used and the robustness of the results.

4.1 Discussion about the Method

The method is implemented in two steps, where the first one covers the estimation of a partially parametric model. Here we first estimate the parametric part in order to reduce the estimation into the estimation of the pure nonparametric model. For the nonparametric estimation we can then use the standard techniques like spline, kernel, loess or local polynomial estimation. The parametric part of the partially parametric model includes only dummy variables for different groups m. Thus, the performance of this model is subjected to finding a variable that divides the population into (M) groups such that the dependencies between x_i and y_i in these are (as) close (as possible) to vertical shifts. From the first step estimation we get the general profile $\hat{f}$ and by using this we can give estimate for y_i, given x_i and $\hat{\gamma}_{m_i}$

$$\hat{y}_i = \hat{f}(x_i) + \hat{\gamma}_{m_i}. \tag{13}$$

The restriction that we have in this model is that $\hat{f}$ is common for everyone. This means that if the predicted value for y_i is larger for the group 1 than for 2 for some x', then this is also the case for every other x''.[19]

[19] I.e. the order of the predictions between different subgroups stays the same independently of the value of x.

The second step is concerned with transforming the dependence resulting from the first step estimation. The transformation is done by estimating the linear regression model, where the dependent variable y_i is regressed on the predicted values form the first step estimation, $\hat{f}(x_i)$.[20] The functional form for the second step estimation is thus carried by the general profile. As long as the group specific profiles f_g are close to being just scaled and shifted general profiles f, the second step gives us good approximations for each of the groups we are interested in. The profiles $\hat{f}_g$ do not have the restriction appearing in the first step estimation.[21] Now, if the predicted value for y_i is larger for group 1 than for group 2 with some x', this does not guarantee, that it would be the case with every other x'', i.e. we do not have a parallel model anymore. Despite of the differences in the degree of restrictions between the general profile and the group specific profiles, these both share a nice property of being able to give reasonable approximations for the places we do not have observations on.

4.2 Discussion about Using Only Four Variables

In this study we use the information of four variables only. These are the survey year, the total consumption expenditures of a household in a year, the age and the year of birth of the household head. Someone might even ask whether we can perform a reliable study with so few variables. The answer for this potential question is obviously yes. The reasoning is given below.

The researchers have an obvious reason for being afraid of omitting the variables from their regression models - the omitted variable bias. If the problem is met the reliability of the parameter estimates is straightforwardly

[20]The chosen type of transformation retains all the extreme values of $\hat{f}_g$ to be at the same x_i that the ones of $\hat{f}$.

[21]If the parallel model truly is the case then also f_g's share this property, but it is not enforced to hold in general.

subjected to this. The estimator is then not consistent in general and thus fails in its only task of giving good estimates for the true parameter value. The occurrence of the omitted variables takes us away from the safe and easy road of leaning on the laws and theorems based on the asymptotical properties - now even the large sample size is not going to help us, but the problem remains.

Luckily, the omitted variable bias is a problem only when we are interested in the parameter values. Here we are *not* interested in these, but in providing an approximation for the relation between the age and the consumption expenditures. In the first step we want to find an approximation for the general profile. This is done by the nonparametric estimation techniques like spline, kernel and loess estimation. These give under certain conditions the consistent estimators for this general profile f *even if we have some variables omitted*. The second question is about whether there are differences in the age profiles for consumption between the generations. The performance of the estimator arising from the second step estimation is again independent of the omitted variables and is just related to whether the generation specific profiles are close to being just scaled and shifted general profiles. Thus neither of the steps suffers from the possibly omitted variables. Including more variables into the model might even do harm when answering to the relation between the age and the consumption.

4.3 Discussion about the Robustness of the Results

How would a researcher report the results from a study if one was able to choose *any* way? The optimal way would probably be such that the results hold independently of the method and model used. This is obviously something we cannot achieve, because there are already infinite number of slightly different functional forms and all of these cannot be used even in a single study. Despite of this incapability we can still try to get closer to the optimal way. This study takes steps towards that direction by performing a battery of estimations that all give the same qualitative results thus making the results extremely robust. First, multiple estimation techniques - spline, kernel, loess and local polynomial estimation - have been employed.[22] Second, these are combined with multiple choices of weight functions. Third, multiple different definitions for the generations have been performed. In table 1 we give the results from the second step estimation. Here the generations are defined such that boys are born after 1965, fathers from 1945 to 1965 and grandfathers before 1945. The estimations have also been performed with the corresponding pairs of years of birth $(1940, 1960), (1941, 1961) \ldots, (1950, 1970)$ and the results for the scaling parameters are given in table 2 in appendix C. Each of these estimations suggests that the fathers have more gentle age profile for consumption than grandfathers. As this study has taken a lot care of the robustness of the results we believe that it is highly unlikely that the results would appear again and again if they were driven by the choice of the model, by the choice of the weights or by the definition of the generation.

5 Conclusions

The paper gives an empirical study employing the Finnish Household Survey data. Two research questions have been studied. First we model the relation between the age and the consumption expenditures to describe the way the consumption is distributed over the life-cycle. Second, we focus on the question of how the age profiles differ between the different generations. Answering this question gives us also an answer to the most interesting ques-

[22]In addition to these nonparametric estimations also multiple parametric estimations have been performed and again these all give the same qualitative results. These results are available when requested.

tion about how the big generation is going to consume when they are old in comparison with the previous generation when they were old.

On technical point of view there are some special things that have to be taken care of when choosing the method being employed. If we approached our main question purely by the dummy-variables for ages, years of birth and survey years, we would restrict ourselves into a parallel model. The null hypothesis about the parallel model with respect to generations is rejected and thus the estimation of this parallel model would not be the right thing to do. With cross-terms for ages and years of births we would be able to get rid of this restriction. Now we do not have all the combinations of ages and years of births and thus this approach is frustrated. Another thing we have to be able to handle is that a person born at a particular year can be observed at a certain age only at a particular year. Thus we have to cope with the effect of the 'state of the world' for every year. We follow the footsteps of Ehrlich and Becker (1972) as they emphasize that the state of the world should be separated clearly from the tastes. This is handled here by letting the survey year to carry the information about the state of the world and the year of birth carries the information about the tastes (for the consumption).

The paper provides a method that is implemented in two steps. In first of these we estimate the partially parametric model. The results of this estimation tell that the consumption increases as a function of age until about 40 or 45 and decreases after that. The second step is concerned about transforming the dependence got from the first step. After this transformation we have the own age profiles for consumption for each of the generations. The results from the second step estimation indicate that the big generation will probably consume more as old than the previous generation at the same age even if there was no economic growth.

The contributions of this paper to the economic literature are two-folded.

First, we provide a technique[23] whereby we can give reasonable approximations for the population that we do not observe completely - like the one depicted in figure 1. Second, we provide our piece of information about the big generation.

Bibliography

Attanasio O, Browning M. 1995. Consumption over the Life-Cycle and over the Business Cycle. *The American Economic Review* 85: 1118-1137.

Banks J, Blundell R, Lewbel A. 1997. Quadratic Engel Curves and Consumer Demand. *The Review of Economics and Statistics* 79: 527-539. DOI: www.mitpressjournals.org/doi/pdf/10.1162/003465397557015

Banks J, Blundell R, Tanner S. 1998. Is There a Retirement-Savings Puzzle? *The American Economic Review* 88: 769-788.

Banks J, Johnson P. 1994. Equivalence Scale Relativities Revisited. *The Economic Journal* 104: 883-890.

Bierens H, Pott-Buter H. 1990. Specification of Household Engel Curves by Nonparametric Regression. *Economic Reviews* 9: 123-184.

Blundell R, Browning M, Meghir C. 1994. Consumer Demand and the Life-Cycle Allocation of Household Expenditures. *The Review of Economic Studies* 61: 57-80.

Blundell R, Browning M,Crawford I. 2003. Nonparametric Engel Curves and Revealed Preference. *Econometrica* 71: 205-240. DOI: www.blackwell-synergy.com/links/doi/10.1111%2F1468-0262.00394

Blundell R, Chen X, Kristensen D. 2007. Semi-Nonparametric IV Estimation of Shape-invariant Engel Curves. *Econometrica* 75: 1613-1669. DOI: www.blackwell-synergy.com/doi/abs/10.1111/j.1468-0262.2007.00808.x?ai=6pf&mi=483l8&af=A

Blundell R, Duncan A. 1998. Kernel Regression in Empirical Microeconomics. *Journal of Human Resources* 33: 62-87.

[23]The asymptotic properties of the estimator are left to future studies.

Blundell R, Duncan A, Pendakur K. 1999. Semiparametric Estimation and Consumer Demand. *Journal of Applied Econometrics* 13: 435-461. DOI: doi.wiley.com/10.1002/(SICI) 1099-1255(1998090)13:5%3C435::AID-JAE506%3E3.0.CO%3B2-K

Clark R. 1975. A Calibration Curve for Radiocarbon Dates. *Antiquity* 49: 251-266.

Cleveland W. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74: 829-836.

Cleveland W, Devlin S. 1988. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* 83: 596-610.

Deaton A. 1998. The Analysis of Household Surveys. *The Johns Hopkins University Press.*

Deaton A, Muellbauer J. 1980. An Almost Ideal Demand System. *The American Economic Review* 70: 312-326.

Ehrlich I, Becker G. 1972. Market Insurance, Self-Insurance and Self-Protection. *Journal of Political Economy* 80: 623-648.

Ellison G, Ellison S. 2000. A Simple Framework for Nonparametric Specification Testing. *Journal of Econometrics* 96: 1-23. DOI: 10.1016/S0304-4076(99)00048-2

Engel E. 1857. Die Productions- und Consumptionsverhaeltnisse des Koenigsreichs Sachsen. *Zeitschrift des Statistischen Bureaus des Koniglich Sachsischen Ministeriums des Inneren 8 und 9.*

Engel E. 1895. Die Lebenskosten Belgischer Arbeiter-Familien Früher und Jetzt. *International Statistical Institute Bulletin* 9: 1-74.

Fan J, Gijbels I. 1992. Variable Bandwidth and Local Linear Regression Smoothers. *The Annals of Statistics* 20: 2008-2036. DOI: 10.1214/aos/1176348900

Härdle W, Mammen E. 1993. Comparing Nonparametric Versus Parametric Fits. *The Annals of Statistics* 21: 1926-1947. DOI: 10.1214/aos/1176349403

Härdle W, Marron J. 1990. Semiparametric Comparison of Regression Curves. *The Annals of Statistics* 18: 63-89. DOI: 10.1214/aos/1176347493

Hausman J, Newey W. 1995. Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss. *Econometrica* 63: 1445-1476.

Hausman J, Newey W, Powell J. 1995. Nonlinear Errors in Variables: Estimation of Some Engel Curves. *Journal of Econometrics* 65: 205-234. DOI: 10.1016/0304-4076(94)01602-V

Horowitz J. 2006. Testing a Parametric Model Against a Nonparametric Alternative With Identification Through Instrumental Variables. *Econometrica* 74: 521-538. DOI: 10.1111/j.1468-0262.2006.00670.x

Horowitz J, Manski C. 2006. Identification and Estimation of Statistical Functionals Using Incomplete Data. *Journal of Econometrics* 132: 445-459. DOI: 10.1016/j.jeconom.2005.02.007

Horowitz J, Spokoiny V. 2001. An Adaptive, Rate-Optimal Test of a Parametric Mean Regression Model Against a Nonparametric Alternative. *Econometrica* 69: 599-631. DOI: 10.1111/1468-0262.00207

Jorgenson D, Lau L, Stoker T. 1982. The Transcendental Logarithmic Model of Aggregate Consumption Behavior. In Advances in Econometrics, Basman R, Rhodes G (eds). Greenwich: JAI Press.

Kaplan E, Meier P. 1958. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 53: 457-481.

Kneip A. 1994. Nonparametric Estimation of Common Regressors for Similar Curve Data. *The Annals of Statistics* 22: 1386-1427. DOI: 10.1214/aos/1176325634

Leser C. 1963. Forms of Engel Functions. *Econometrica* 31: 694-703.

Lewbel A. 1989. Household Equivalence Scales and Welfare Comparisons. *Journal of Public Economics* 39: 377-391.

Lewbel A. 1991. The Rank of Demand Systems: Theory and Nonparametric Estimation. *Econometrica* 59: 711-730.

Lewbel A. 2008. Shape Invariant Demand Functions. *Working paper*

Loader C. 1999. Local Regression and Likelihood. *Springer-Verlag.*

Mammen E. 1991. Nonparametric Regression Under Qualitative Smoothness Assumption. *The Annals of Statistics* 19: 741-759. DOI: 10.1214/aos/1176348118

Marron J. 1988. Automatic Smoothing Parameter Selection: A Survey. *Empirical Economics* 13: 187-208. DOI: 10.1007/BF01972448

Nadaraya E. 1964. On Estimating Regression. *Theory of Probability and Its Applications* 10: 186-190. DOI: http://dx.doi.org/10.1137/1109020

Pagan A. 1984. Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review* 25: 1-40.

Pagan A. 1986. Two Stage and Related Estimators and Their Applications. *The Review of Economic Studies* 53: 517-538.

Pagan A, Ullah A. 1999. Nonparametric Econometrics. *Cambridge University Press.*

Pendakur K. 1999. Semiparametric Estimates and Tests of Base-independent Equivalence Scales. *Journal of Econometrics* 88: 1-40. DOI: 10.1016/S0304-4076(98)00020-7

Pinkse C, Robinson M. 1995. Pooling Nonparametric Estimates of Regression Functions with Similar Shape. In Advances in Econometrics and Quantitative Economics: Essays in Honour C. R. Rao, Maddala G, Phillips P, Srinivasan T (eds). Cambridge, Mass: Blackwell.

Priestley M, Chao M. 1972. Non-Parametric Function Fitting. *Journal of the Royal Statistical Society, Series B (Methodological)* 34: 385-392.

Racine J, Li Q. 2004. Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data. *Journal of Econometrics* 119: 99-130. DOI: 10.1016/S0304-4076(03)00157-X

Reinsch C. 1967. Smoothing by Spline Functions. *Numerische Mathematik* 10: 177-183. DOI: 10.1007/BF02162161

Reinsch C. 1971. Smoothing by Spline Functions II. *Numerische Mathematik* 16: 451-454. DOI: 10.1007/BF02169154

Robinson P. 1988. Root-N-Consistent Semiparametric Regression. *Econometrica* 56: 931-954.

Schimek M. (eds) 2000. Smoothing and Regression: approaches, computation and application. *John Wiley & Sons, Inc.*

Schmalensee R, Stoker T. 1999. Household Gasoline Demand in the United States. *Econometrica* 67: 645-662. DOI: 10.1111/1468-0262.00041

Schoenberg I. 1964. Spline Functions and the Problem of Graduation. *Proc. Nat. Acad. Sci. U.S.A.* 52: 947-950.

Stengos T, Sun Y, Wang D. 2006. Estimates of Semiparametric Equivalence Scales. *Journal of Applied Econometrics* 21: 629-639. DOI: 10.1002/jae.863

Tilastokeskus 2001. Kulutustutkimus 1998 Laatuselvitys. Julkaisusarjassa Tulot ja kulutus 2001 : 4; *Yliopistopaino.* (english: Statistics Finland 2001. Finnish Household Budget Survey 1998, Quality Report. In Income and Consumption 2001:4)

Tilastokeskus 2004. Kulutustutkimus 2001-2002 Laatuselvitys. Julkaisusarjassa Katsauksia 2004/5; *Multiprint Oy.* (english: Statistics Finland 2004. Finnish Household Budget Survey 2001-2002, Quality Report. In Reviews 2004/5.)

Ullah A. 1988. Nonparametric Estimation and Hypothesis Testing in Econometric Models. *Empirical Economics* 13: 223-249.

Wahba G, Wold S. 1975. A Completely Automatic French Curve: Fitting Spline Functions by Cross-Validation. *Communications in Statistics, Series A* 4: 1-17. DOI: 10.1080/0361091 7508548493

Watson G. 1964. Smooth Regression Analysis. *Sankhya, Series A* 26: 359-372.

Wilke R. 2006. Semi-Parametric Estimation of Consumption-Based Equivalence Scales: The Case of Germany. *Journal of Applied Econometrics* 21: 781-802. DOI: doi.wiley.com/10 .1002/jae.888

Working H. 1943. Statistical Laws of Family Expenditure. *Journal of the American Statistical Association* 38: 43-56.

Yatchew A. 1998. Nonparametric Regression Techniques in Economics. *Journal of Economic Literature* 36: 669-721.

Yatchew A. 1992. Nonparametric Regression Model Tests Based on Least Squares. *Economic Theory* 8: 435-451.

Yatchew A. 2003. Semiparametric Regression for the Applied Econometrician. *Cambridge University Press.*

Yatchew A, Bos L. 1997. Nonparametric Least Squares Regression and Testing in Economic Models. *Journal of Quantitative Economics* 13: 81-131.

Yatchew A, No A. 2001. Household Gasoline Demand in Canada. *Econometrica* 69: 1697-1709. DOI: www.blackwell-synergy.com/doi/abs/10.1111/1468-0262.00264

Appendix A

In figure 11 we give the results from the regression curves arising from the local polynomial fitting with the zeroth order polynomials. These share the wiggling with the regression curves from the kernel estimation. The similar feature is also observed with the first order polynomials on the left hand side of the figure 12, but here the effect is milder. On the right hand side of this figure we have the results from the estimations with the second degree polynomials. Here all but the third one share the typical features of the other regression curves.

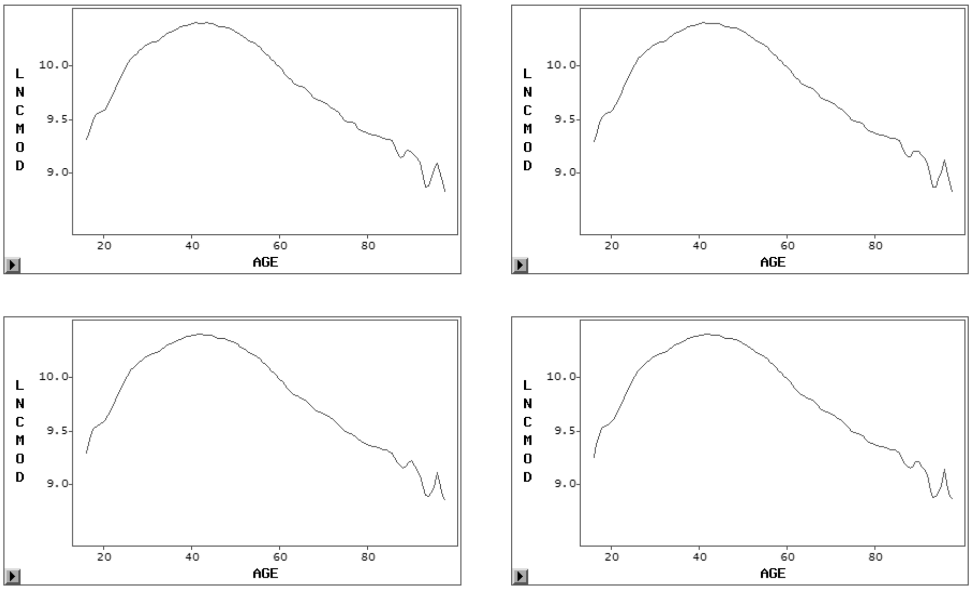

Figure 11: Estimates form local polynomial regressions with the zeroth degree polynomials for f in equation 5. The upper left graph uses normal, the upper right triangular, the lower left quadratic and the lower right tri-cube weight function.

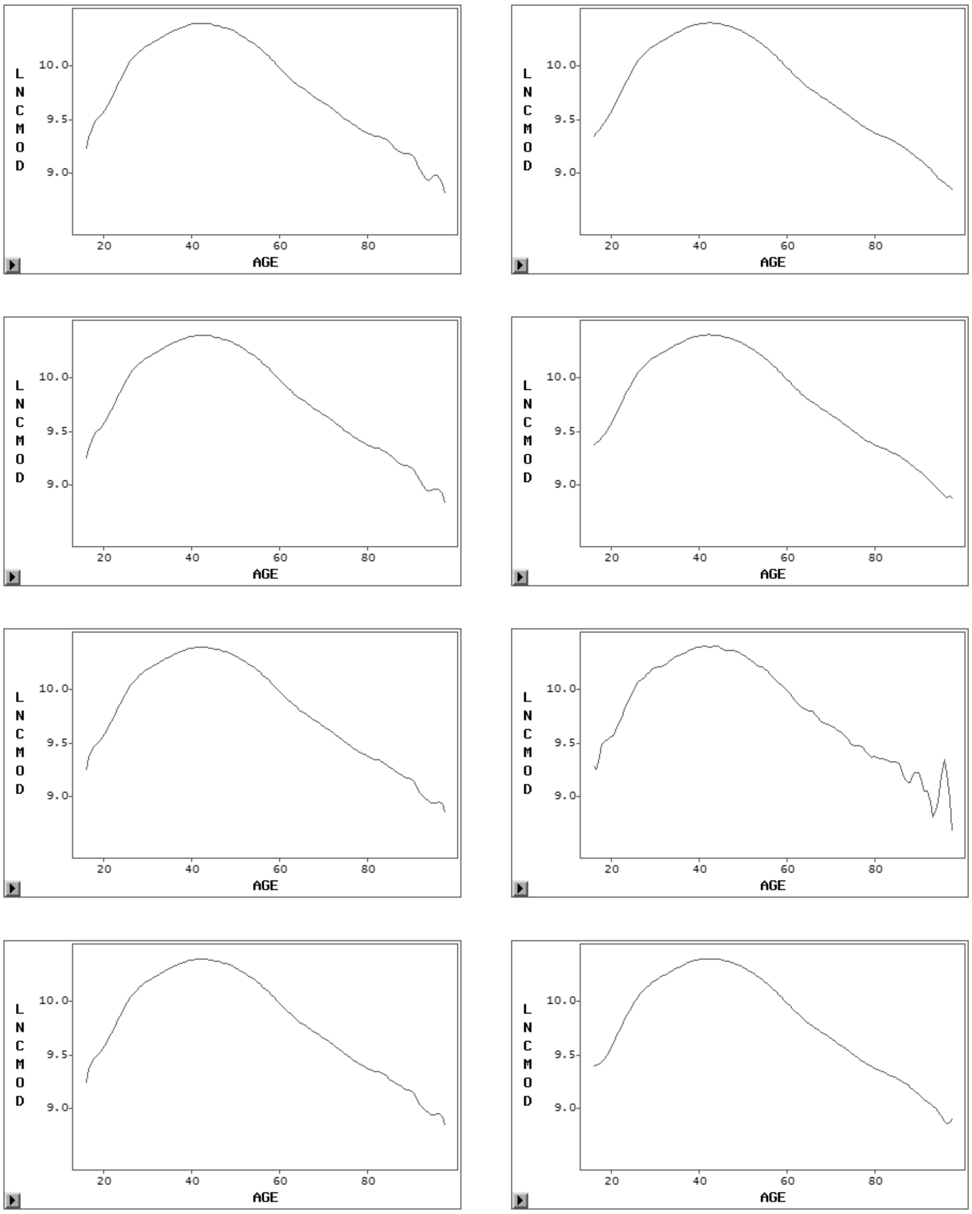

Figure 12: Estimates from local polynomial regressions for f in equation 5. The left hand graphs use the first order polynomials and the ones on the right hand side use the second order polynomials. The weights employed are normal, triangular, quadratic and tri-cube respectively.

Appendix B

Here we provide a short review on the reasoning behind the spline, kernel and loess estimations performed in the first step.

Nonparametric estimation techniques have been in economic researchers toolbox for a long time - the kernel and the spline estimation techniques are introduced already in the 1960s[24] and the loess estimation technique in the 1970s[25]. These techniques have been employed mostly in the estimation of the Engel curves,[26] but otherwise these are, for some reason, not being used that much.[27] The demand for the techniques arose in the context of Engel curves as the parametric models were observed to be insufficient to describe the curves despite the multiplicity of the different specifications.

In the spline estimation we want to find f minimizing

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2 + \eta(f'')^2. \tag{14}$$

Here the the first term tells us about the accuracy the regression curve describes the data and the second term is about the curvature of the function f. The trade-off between them is defined by the parameter η. Finding the solution for the above is equivalent to finding f to minimize

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2 \ s.t. \ (f'')^2 \leq L. \tag{15}$$

[24]Nadaraya (1964), Watson (1964), Schoenberg (1964), Reinsch (1967).

[25]Proposed in Cleveland (1979) and extended in Cleveland and Devlin (1988).

[26]The estimation of the Engel curves was first proposed by Engel (1857) and (1895). After that Working (1943), Leser (1963), Deaton and Muellbauer (1980), Jorgenson et al. (1982), Bierens and Pott-Buter (1990), Lewbel (1991), Härdle and Mammen (1993), Kneip (1994), Hausman et al. (1995), Pinkse and Robinson (1995), Banks et al. (1997), Blundell and Duncan (1998), Blundell et al. (1998), Pendakur (1999), Blundell et al. (2003), Stengos (2006), Wilke (2006) and Blundell et al. (2007) have studied these curves that describe the fractions being consumed to a subcategory.

[27]This is the reason for providing this review.

Now the trade-off is described by L that is chosen by the cross-validation[28]. Here L is chosen to be the one that minimizes the cross-validation function

$$CV(L) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}_{-i}(x_i))^2, \tag{16}$$

where $\hat{f}_{-i}$ arises as a solution from minimization of

$$\frac{1}{n}\sum_{j\neq i}^{n}(y_j - f(x_j))^2 \ s.t. \ (f'')^2 \leq L. \tag{17}$$

Here as L gets bigger the function f is allowed to have more curvature but as a trade-off it has to give better predictions. The results from the spline estimation are depicted in the graph in the upper left corner of figure 6.

Like the spline estimation also the kernel estimation is concerned in finding an approximation for the function f describing the systematic dependence between some variables. The value for this at x_0 is defined to be a weighted sum of the original values neighboring the point x_0 that is

$$\hat{f}(x_0) = \sum_{i=1}^{n} w_i(x_0)y_i, \tag{18}$$

where the weights $w_i(x_0)$ take the form

$$w_i(x_0) = \frac{\frac{1}{\lambda n}K(\frac{x_i - x_0}{\lambda})}{\frac{1}{\lambda n}\sum_{i=1}^{n}K(\frac{x_i - x_0}{\lambda})} \tag{19}$$

The shape of the weight function is driven by the choice of the density function that is here referred as the kernel, K. The kernels being employed are normal, triangular and quadratic ones. The weights for the observation at x, that belong to the neighborhood of x_0 ($N(x_0)$), are thus[29]

$$\begin{cases} \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}(x - x_0)^2), \ x \in N(x_0) \text{ and 0 elsewhere,} \\ a(1 - |x - x_0|), \ x \in N(x_0) \text{ and 0 elsewhere and} \\ a(1 - (x - x_0)^2), \ x \in N(x_0) \text{ and 0 elsewhere} \end{cases}$$

[28] This was first proposed for the spline estimation by Wahba and Wold (1975).

[29] a is to make sure that we have a density function, i.e. the integral over the domain gives the unity. The value for this is dependent on the choice of the bandwidth.

respectively. In addition to the kernel one has to choose the bandwidth λ that tells about the size of neighborhood being encountered at each point. This choice is made by the cross-validation[30], where we choose λ to minimize the cross-validation function

$$CV(\lambda) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}_{-i}(x_i;\lambda))^2. \tag{20}$$

This way we want λ to be chosen such that we get the best possible predictions at x_0 when these are given according to the information of the neighborhood only. The results from the kernel estimations are depicted in the three bottom graphs on the left hand side of the figure 6. From top to bottom the weights employed obey normal, triangular and quadratic distributions.

The third nonparametric technique[31] being employed is the loess estimation. In the estimation we use three different types of local polynomials, zeroth, first and second order. The three estimates for the value of the function f at x_0 are

$$\begin{aligned} \hat{f}(x_0) &= \hat{a}(x_0), \\ \hat{f}(x_0) &= \hat{a}(x_0) + \hat{b}(x_0)x_0 \text{ and} \\ \hat{f}(x_0) &= \hat{a}(x_0) + \hat{b}(x_0)x_0 + \hat{c}(x_0)x_0^2 \end{aligned} \tag{21}$$

respectively. The estimates for a, b and c for the third case come as a solution for

$$min_{a,b,c} \sum_{x_i \in N(x_0)} (y_i - a(x_0) - b(x_0)x_i - c(x_0)x_i^2)^2 w_i(x_0), \tag{22}$$

where $N(x_0)$ denotes the neighborhood of point x_0 and w_i stands for the weights. The size of the neighborhood is chosen by the cross-validation that

[30] This was first proposed for the kernel estimation by Clark (1975).

[31] The nonparametric estimation has also been performed with the fourth technique, the local polynomial estimation. The reasoning of the method can be found in Fan and Gijbels (1992) and the results from the estimations are given in appendix A.

minimizes the out-of-sample prediction error. Four weight functions are employed: normal, triangular, quadratic and tri-cube.[32] The estimations using zeroth order local polynomial are depicted on the right hand side in figure 6 and in figure 7 we plot the regression curves from the rest of the loess estimations. The graphs on the left hand side use the first order polynomial and the ones on the right hand side use the second order polynomials.

Appendix C

Here we provide the values for the scaling parameters in the second step estimation with different definitions for the generations. In the first row in table 2 we have defined the generations such that the boys are the ones that are born after 1960 and fathers are born between 1940 and 1960 and grandfathers are born before 1940. The other definitions for the generations are denoted analogously with this one. What is observed in the table is that with all the used definitions for the generation, the fathers have more gentle age profile for consumption than the one for grandfathers - that is the result we had already with our original definition. The other thing we observe is that the value for the scaling parameter starts to change for boys as we have fewer and fewer boys. For that reason the estimation results for the boys become a bit less robust. This is also seen in the standard errors for the coefficients for boys as these more than doubled from the definition using 1960 compared to the one with 1970. The accuracy of the estimates work other way around for fathers and grandfathers even if the standard errors are not reduced into half between the extremes. This makes the results for fathers and grandfathers to stay very robust even if we change a definition for the generation a bit.

[32] The density function for tri-cube distribution is $a(1-|x-x_0|^3)^3$ in the neighborhood of x_0 and zero elsewhere. Here the choice of a depends on the choice of the smoothing parameter and the purpose of this is to guarantee that we have a density function.

	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$		$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$
$(1940, 1960)$	1.019	0.937	0.994	$(1946, 1966)$	1.024	0.937	0.998
	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$		$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$
$(1941, 1961)$	1.016	0.937	0.992	$(1947, 1967)$	1.021	0.955	0.994
	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$		$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$
$(1942, 1962)$	1.008	0.916	0.996	$(1948, 1968)$	0.993	0.964	0.994
	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$		$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$
$(1943, 1963)$	1.018	0.930	0.997	$(1949, 1969)$	0.950	0.981	0.996
	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$		$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$
$(1944, 1964)$	1.021	0.944	1.000	$(1950, 1970)$	0.934	0.982	0.999
	$\hat{\phi}_{boys}$	$\hat{\phi}_{fath}$	$\hat{\phi}_{grand}$				
$(1945, 1965)$	1.017	0.944	1.000				

Table 2: The results from the second step estimation with different definitions for the generations.