Centre for
Economic and
Business Research

CEBR

CEBR STUDENT PAPER

2003-04

Who Cares?

Jacob Palsgaard
Tore Olsen
Bertel Schjerning

2003-04

Who Cares?*

Tore Olsen — ToreOlsen@metrics.dk
Jacob Palsgaard — Jpp@metrics.dk
Bertel Schjerning — Bsc@metrics.dk

May 27'th., 2003

Supervisor: Martin Browning

Abstract

Our concern with this paper is to develop a deeper understanding of what we can and cannot say about intra-household allocation based on the household survey data used in Browning and Lechene (2001) [henceforth: BL]. Our humble contribution can be summarised thus: i) Accepting the BL framework, we show that their conclusions on the presence of caring can be strengthened by using a better GMM estimator ii) where BL cite difficulties in coping with endogeneity in a semiparametric framework, we venture to integrate GMM technique into a semiparametric framework. iii) Using bootstrapped confidence intervals of the semiparametric fit to measure spline fits, we show that the data imposes severe limits on how many splines should be allowed in the unrestricted parametric model. iv) Combining the insight from parametric and semiparametric analysis, we find evidence that the way BL divide their data sample in splines favours their conclusion that agents exhibit caring v) Using our semiparametric fit as an unrestricted model against which to test linear restrictions, we show that all versions of the collective model fit well. vi) In the semiparametric framework with optimal and adaptive bandwidth we show, that which collective model fits the data best depends on the choice of adaptor. Our conclusion is, that data is not good enough to discriminate between caring and non-caring. Till a better dataset is found, we are not convinced that anyone cares - especially not men.

*We thank Martin Browning for valuable comments and access to the dataset. The usual disclaimer applies. Gauss programs and Excel applications are available from the authors and will be available at www.metrics.dk.

Contents

1 Introduction:

Maybe for lack of better data, for decades neo-classical economists treated the family as a single utility-maximizing unit. The 1980's and especially 1990's abound with theories attempting to rectify this theoretical blemish in the neo-classical tradition. Several papers have rejected the unitary model (Browning and Chiappori 1998 and Browning, Bourguignon, Chiappori and Lechene 1994), but little work has been done in developing a unifying framework in which to test and compare alternative theories. This is rectified in a seminal paper by Browning and Lechene (2001) (henceforth BL).

Purpose and methodological considerations: Our intention with this paper is to find out, what we can and cannot learn about intra-household allocation studying the data used in BL. Even with at very good sample, this kind of endeavour would always face a difficult methodological challenge: On one hand, what we can observe is behaviour and, on the other hand, what economic theories makes assumptions about, is what people want to do (their preferences) and what they can do (their constraints). Analytically, it is a strength of the economic theory of intra-household allocation that it separates its discussion of preferences, from its discussion of the decision processes. However, looking at observed behaviour (household data), we claim as students of neo-classical economics to see the joint manifestation of *constrained* maximization over rational *preferences*. This begs the question: Given behaviour, what inferences can we make regarding the appropriateness of theories of preferences and of theories of constraints?

One could fear that observable behavior was theoretically overdetermined, in the sense that several different theories would predict the same kind of behavior. Browning (lectures) and BL credit Chiappori for first showing, that an efficient-outcome requirement imposes enough restrictions to allow for discrimination between models. What is ingenious about the BL framework, is that it provides a unified framework for the identification of and discrimination between: a) the best fitting models of caring and non-caring preferences, b) the best efficient-outcome models and the best non-efficient outcome models. We use the term 'efficient-outcome models' here, because BL use the term 'collective model' in two different ways:

First BL define collective models as a group of models, only characterized by assuming efficiency - in this sense bargaining with a non-cooperative breakdown point is a collective model (BL p.4). This formulation makes collective models entirely compatible with the existence of Warr (and Becker) regions.

The second way BL use 'collective model', they add the extra assumption, that the Pareto-weight is strictly increasing in the distribution factor (BL p.6). This definition is incompatible with the bargaining model with a non-cooperative breakdown point. Or, more precisely: it is incompatible with the bargaining model with non-cooperative breakdown point *if* preferences are such, that a Warr region exists in the non-cooperative model. This is because the bargaining model with a non-cooperative breakdown point inherits the Warr region from the non-cooperative model. It does so, because the existence of a Warr region in the non-cooperative model means, that there will be regions over the distribution factor, where a small transfer of income between partners will not change their threat point and so will not change their bargaining power. With no change in bargaining power, outcome (household behaviour) will remain constant. In what follows we use BL's second definition when we refer to the collective model.

Plan: Over the next page we give a short description of the models treated, and give a brief critique of the way BL handle distribution factors. After that, we detail the results of our empirical analysis and finally we will give a conclusion on what we have learned.

1.1 Models of intra-household allocation

In the analysis of intra-household decision processes, the demand model has to incorporate both the individual preferences of the partners and the way these preferences affect the family outcome. Assume a couple consists of two adults: a wife, A, and a husband, B. Define ρ as the wifes share of income. Assume two goods: a private good, q, and a public good, Q. Define q^i as person i's consumption of a private good (i=a,b).

Caring Preferences: BL choose to modelize both egoistic $u^I = v^I\left(q^I, Q\right)$ and caring preferences $u^I = v^I\left(q^A, q^B, Q\right)$. In the latter case a partner benefits from her partners utility from consumption and not from his consumption per se (and vice versa). With caring preferences a richer partner will behave as a benevolent dictator, making transfers to the poorer partner. Caring will manifest itself in outcomes by an unchanged level of consumption over an interval $[0, \rho_1]$ when men have caring preferences, and $[\rho_2, 1]$ when women have caring preferences. The lengths of these segments - the Becker regions - are determined by the preferences and decision structure in conjunction. Such caring segments are compatible with all the following decision models and can be formulated as restrictions on them. (Please see appendix 1). In the rest of this section, we assume away caring for expositional clarity.

Distribution factors and decision processes: In their description of the allocation process BL explicitly treats only one, exogenously determined, distribution factor ρ. This is controversial[1] as the literature documents more distribution factors to be significant, (see Browning, Bourguignon, Chiappori and Lechene (1994)). Define $g_j(\rho)$ as $\rho's$ contribution to the budgetshare of good j.

The first model is the *collective model*, which imposes the restriction that demand is efficient

[1] A fact that BL recognize (BL p.5)

and partners influence on decisions is strictly increasing in the distribution factor[2]. The resultant behaviour can be rationalized as the outcome of maximization of a household utility function. Consequences for data is that $g_j'(\rho)$ is strictly positive or strictly negative for all ρ, if partners have different preferences for a good.

The second model is the *non-cooperative model*. This model has two distinct features: it predicts changing signs of $g_j'(\rho)$ over ρ, and it allows for Warr regions.

The non-cooperative model treats the household allocation decision as if it were a one-shot game of voluntary contributions to a public good. In all but the extremes of ρ, it results in inefficient outcomes. Intuitively, this is not a very appealing approach, as couples must be believed to have all the prerequisites for sustaining an efficient equilibrium: almost common knowledge of preferences and repeated interaction over a non-limited time horizon. What separates the husbands and wife's situation from being single, is that they share a public good. If the wife prefers a public good more than her husband, she would contribute more to it, had she all the income. However, starting from $\rho = 0$,as income begins to be transferred to her, her marginal utility of the private good will be much higher than for the public good. Consequently she will not immediately start to contribute to the public good. As her husband is equalizing marginal utility between the private and the public good, his contribution will fall as his income falls. This gives rise to the interesting conclusion, that as the person who prefers the public good more gains in relative income (from a low starting point), the consumption of the public good will fall over an interval. Eventually it will rise. This prediction is distinct from that of the collective model, and it is easy to test: $g_j'(\rho)$ will change sign over ρ.

BL attribute the other main feature of the voluntary contributions structure to Warr: this is the result, that when a group of agents all contribute to a public good at the same time, small changes in income allocations will not affect outcomes. The flat segment that arises from this on a budgetshare-against-ρ curve is called the Warr region. Again this is a testable prediction.

The third model is the *bargaining model with a non-cooperative breakdown point*. Other breakdown points could have been chosen, but this model has two important testable predictions: i) Since this model attains efficiency its $g(\rho)$ curve is weakly monotonous. ii) The bargaining model inherits the Warr region from the non-cooperative model (if there is a Warr region). Reason: The threat point affects the bargaining power - and if the non-cooperative model has a Warr region, then the threat point is not affected over this region. Since the bargaining power is not changed, neither is the outcome over the Warr region. Note, though, that the Warr region

[2]Distribution factors are factors that do not influence preferences, but do infuence the relative indluence on outcomes among partners.

will be at higher level than in the non-cooperative model.

2 Econometric Issues

2.1 Data, model and method in BL

For our econometric analysis, we use the same data as BL. We do not wish to paraphrase their description of the data, so please refer to BL for a description of the data.

The model used by BL has the partial linear structure $\omega_j = X\beta_j + g_j(\rho) + \varepsilon_j$ where $j = 1, .., M = 6$ is the number of markets, ω_j is the budget share of good j and X is explanatory variables other than ρ[3]. BL cite difficulties in dealing with endogeneity in a semiparametric framework (BL p.15) and so perform a fully parametric spline regression using GMM: $\omega_j = x\beta_j + D_s\Gamma + \varepsilon_j \quad j = 1, .., M$, where D_s is a spline matrix. This is an asymptotically equivalent approach, if the number of splines increases at a lower rate than sample size, guaranteeing consistency of Γ.

BL formulate a 20 spline unrestricted model[4], in terms of which they express all theory models as restricted variants. For each model and each combination of join points, BL calculate a Chi-Squared criterion, measuring the difference between the restricted estimator $\hat{\beta}^R$ and the unrestricted estimator $\hat{\beta}^{UR}$. They grid search to find the lowest criterion for each type of model, and then compare these minimum χ^2-criteria across models. Thus the unrestricted model provides a reference point against which to test, and this makes comparisons between restricted models possible. This is what constitutes their unifying framework for testing models of intra-household allocation.

2.2 What we find working in the BL framework

We show that BL appear to use an indirect expression of $\hat{\beta}^R$ in terms of $\hat{\beta}^{UR}$. We call this the BL-estimator: $\hat{\beta}^R_{BL} = (R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}}R)^{-1}R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}}\hat{\beta}^{UR}$. We also show that while this is a closed form expression for OLS and GLS, it is not necessarily so for IVE, and for GMM it is not a closed form expression. It effectively does not impose restrictions on the instrument variables and does not make the best use of the weighting matrix. But $\hat{\beta}^R_{BL}$ is consistent. We document the differences between $\hat{\beta}^R_{GMM}$ and $\hat{\beta}^R_{BL}$ in the appendix. The differences in the estimators indicate that there are small sample correlation between Z and ε, which should make us cautious when relying on asymptotic results and statistics. Since we also find evidence of non-normality in the

[3]For a discussion of and tests for the error characteristics, look in appendix 1.2.

[4]Each spline contains 5% of the observations.

GMM-error term, caution should be double.

Applying $\hat{\beta}^R_{GMM}$ instead of $\hat{\beta}^R_{BL}$, we show that the BLs conclusion that both men and women exhibit caring is strengthened. Using $\hat{\beta}^R_{GMM}$ we find the same ranking between different models. But within the collective model with caring, we find support for a longer leftmost Becker-region (we find it to include 10% rather than the 5% of the sample that BL finds).

The choice of number of splines: The placement of the join points will have a crucial impact on the shape of the curve, since a local extrema must be in a join point in order to avoid bias. Secondly the number of splines will determine how volatile the unrestricted curve will be: Too many splines will make the fluctuations arbitrary and too few will impose heavy restrictions on the model. Thus, the number of splines will have an important impact on the minimum chi squared statistics. Even though this can influence which models are accepted, there seems to be little guidance in the literature about the optimal number of splines; BL express indifference to the fact, that the estimation of a large number of spline-parameters in the small sample used, introduces large fluctuations in the unrestricted model (BL p.15). We beg to differ.

3 Semiparametric regression

To deal with these problems, we suggest using a semiparametric approach, which allows us to better balance the problems of bias and variance. This way, we can gain a better idea about the shape of $g(\rho)$. *Ex ante*, a drawback is, that the semiparametric fit will have a larger variance. Thus it will be more difficult to discriminate between some of the theoretical models that we wish to investigate.

In order to make the results comparable with BL, we take an approach that is able to i) control for the other explanatory variables than ρ, ii) adjust for the endogeneity of total expenditure, and iii) implement the covariance structure from the parametric model. To accommodate this, we integrate a semiparametric approach suggested by Robinson(1988) into a GMM framework. We can summarize the estimation procedure in the three following steps[5] i) Estimate $E(\omega_j|\rho)$, $E(X|\rho)$ and $E(Z|\rho)$ nonparametrically and obtain $\hat{m}_h^{\omega_j}(\rho)$, $\hat{m}_h^X(\rho)$ and $\hat{m}_h^Z(\rho)$, ii) Estimate β_j by GMM on the transformed variables $\tilde{\omega}_j = \omega_j - \hat{m}_h^{\omega_j}(\rho)$, $\tilde{X} = X - \hat{m}_h^X(\rho)$ and $\tilde{Z} = Z - \hat{m}_h^Z(\rho)$, and iii) Calculate $\hat{g}_j(\rho) = \hat{m}_h^{\omega_j}(\rho) - \hat{m}_h^X(\rho)\hat{\beta}_j^{GMM}$, and normalize by subtracting the mean[6].

For the estimation of the conditional means $E(\omega_j|\rho)$, $E(X|\rho)$ and $E(Z|\rho)$ we apply the Nadaraya-Watson estimator with Gaussian kernels and optimal global bandwidth adjusted for

[5] A detailed derivation of the estimator is given in appendix 2.

[6] When $g(\rho)$ is estimated we effectively ignore the intercept in the regression. Therefore we assume that $g(\rho)$ integrates to zero in order to identify $g(\rho)$. See appendix 2 for discussion.

the varying density of ρ. In words, this adjustment means that the nonparametric estimators attain a higher degree of smoothness in areas with sparse data. This is done in order to correct for the larger magnitude of variance relative to bias in these regions.

In order to make inferences, we need a measure of the precision of the semiparametric fit. For this purpose we simulate confidence intervals based on bootstrap samples. Since we have already seen indications of small sample problems, we have reason to believe that the asymptotic theory provides a poor guide to the precision of the estimator. Therefore, we regard the bootstrap technique as a desirable alternative to the asymptotics, since it is tractable and may provide a better finite sample approximation. The chosen bootstrap algorithm is the so called *Naive bootstrap* (see Härdle(1990)), where bootstrap samples are drawn with replacement from the original data. This algorithm has the advantage, that it is robust to heteroscedasticity, which is particular relevant in our case.

4 Spline vs. semiparametric estimation

As indicated above the number of splines is crucial for the parametric analysis. The semiparametric analogue to the number of splines is the choice of bandwidth, since this choice determines the contribution of the surrounding observations in the estimation of the curve at a given point. In the semiparametric model the bandwidth is chosen optimally by minimizing the Mean Squared Error of the estimator, whereas the spline method in BL has no such procedure.

Therefore we believe that the semiparametric regression provides a reasonable benchmark for analyzing the spline models. Hence it gives a powerful tool to evaluate i) how to decide the number of splines in the unrestricted model, and ii) how the restricted parametric model fits against an unrestricted semiparametric alternative.

By setting the bandwidth much lower than the optimal value, we show that the semiparametric curve can give a fit much like the unrestricted spline model in BL. Moreover the spline model is shown to lie outside the confidence intervals of the semiparametric model with optimal bandwidth[7].We take this as an indication that the unrestricted model in BL is suffering from spurious fluctuations, caused by an excessive number of splines.

In order to evaluate the distance between the spline regression and the semiparametric fit, we suggest a convenient goodness of fit test. The test statistic is based on the Weighted Average Squared Error between the semiparametric fit and the predicted values of $g(\rho)$ from the spline regression. We simulated a distribution of the test statistic under the null that the restricted parametric model lies within the confidence bands of the semiparametric fit. This

[7]See appendix figures A3.1 and A3.2

gives a desirable supplement to the minimum chi square method, that allow us to rank different parametric models' fit relative to the semiparametric alternative.

The goodness of fit tests supports the evidence form the graphical inspection of the unrestricted spline regression. Regardless of the chosen adopter we see, that the model with 20 splines is strongly rejected by the goodness of fit test. Actually, all models using more than 6 splines are rejected. We interpret this as substantial evidence for our hypothesis: given the sample size there is a upper bound for the number splines to be used in the unrestricted model. In this case the upper bound is six. Consistent with other studies, the unitary model is rejected by the test, while models having 1-6 splines can't be rejected.

The question is now: How does this affect the conclusion in BL? We performed the minimum chi-square grid search with fewer splines and found that the conclusions is highly sensitive to the number of splines in the unrestricted parametric model. In fact with less than 9 splines the flat-slope-flat model can be reduced to slope-flat, and with less than 7 splines slope-flat can be reduced to slope.

As the Goodness-of-fit test rejects models with more than 6 splines no support for Becker regions can be found in a fully parametric setup. However, this setup still provides substantial evidence in favor of the monotone curves which form the core of the collective model.

When the collective models are tested *directly* against the semiparametric specification, we find that the results are sensitive to the degree of smooth adaption. Specifically, as the adaptor increases, the test favors the hypothesis of Caring. But still, with a reasonable choice of smooth adaption ($\xi = 0.5$), we find that flat-slope-flat is the collective model that gives the best fit.

4.1 Conclusion

As we have argued, the parametric specification in BL favors the hypothesis of caring. Generally the evidence for Becker regions is shown to be quite sensitive to the applied econometric method. However, in line with BL, we find substantial evidence in favor of efficient decision process within the household. Our conclusion is, that data is not good enough to discriminate between caring and non-caring. Till a better data set is found we are not convinced that anyone cares - especially not men.

5 Appendix 1: Introducing the spline model.

5.1 The genesis of an unrestricted model specification

5.1.1 The very general non-parametric model of demand

Testing theory models of intra-household allocation involves accounting for both inter- and intra-household allocation. Making different models commensurable involves testing against a sufficiently general unrestricted consumption model. In very general terms the unrestricted model can be formulated *non-parametrically.*

$$\omega_j = f_j\left(x_{EN}, x_{EX}, \rho\right) + \varepsilon_j \quad j = 1, .., M \tag{1}$$

where ω_j is budget share of good j, ρ is the wife's share of income[8], while x_{EN} and x_{EX} are a complete set of other possibly endogenous and exogenous variables theoretically relevant in explaining the budget share of good j.

5.1.2 The separable and semi-parametric model

Testing the relevant models involves restricting $f_1'\left(\rho\right) = ... = f_M'\left(\rho\right) = 0$ for intervals of ρ - the possible Becker and Warr regions. If no weighty theoretical considerations can be brought to hold against linearity in parameters of X_{EN} and X_{EX}[9] and separability between these and ρ, then we can formulate the relevant regression model *semiparameticaly*,

$$\omega_j = \left(x_{EN}, x_{EX}\right)\beta_j + g_j\left(\rho\right) + \varepsilon_j \quad j = 1, .., M \tag{2}$$

where β controls for the *inter-household* allocation (e.g. that the total expenditure, prices faced and demographic factors affects what and how much families consume) while $g_j\left(\rho\right)$ explains the *intra-household* allocation (e.g. that, *ceteris paribus* and for some values of ρ, the budget share of women's clothing is increasing in ρ). This formulation (2) is practical, because restrictions can be formulated in terms of $g(\rho)$ only.

5.1.3 The fully parametric model

Testing will be significantly simplified if a model can be formulated, that is asymptotically identical to (2) and yet has a *fully parametrical* specification. This is the route chosen by BL.

[8]Which is is taken to be exogenous here. This is possibly problematic given the data source used - FAMEX is based on interviews. Thus the data are subject to recall bias (BL2001, p29).

[9]Or transformations of X_{EN} and X_{EX}.

They choose to divide ρ in intervals, each containing the same number of observations. Let D_s be a spline dummy matrix with s splines, then

$$\omega_j = (x_{EN}, x_{EX})\,\beta_j^x + D_s\Gamma + \varepsilon_j \quad j = 1, .., M \tag{3a}$$

is asymptotically equivalent to (2) [10] when the number of splines, s, goes to infinity at a slower rate than the number of observations does, such that consistency of Γ is assured.

BL use $(3a)$ as the unrestricted model specification against which they test different theory models using Minimum Chi-squared technique. The model is attractive because it is purely parametric and easy to restrict.

Small samples always warrant carefulness in relying on asymptotic results, and this is doubly so in this case. For a given number of observations, every spline added reduces the information used in estimating its slope. Specifying too many splines in the unrestricted model will introduce spurious variation (Browning, conversation April 2003) and contrary to BL(p.15) we contend that the number of splines in the unrestricted model does matter. We find that the model results are highly sensitive to the number of splines chosen. For given number of splines, the following table shows the probability that the best fitting flat-slope-flat model can be reduced to the slope-flat model with the best fit.

Table A1.1: Probability that flat-slope-flat model can be restricted to slope-flat

# splines in unrestricted model	6	7	8	9	10	11	12	20
Probability in %:	28.02	11.13	6.32	4.38	3.86	3.62	3.14	1.65

Of course this might be an unfair comparison as the length of the first flat parts might be different for different choices of the number of splines in the unrestricted model. However, as the next table shows, the number of splines in the unrestricted model matter. Here the splines have exactly equal length.

Table A1.2: χ^2-tests with different unrestricted models

# splines		10 splines	20 splines	rho-value
Flat-slope-flat	join point 1	1	2	0.254
	join point 2	6	12	0.433
	chi2 value	52.6	150.3	
slope-flat	join point	6	12	0.433
	chi2 value	56.6	156.2	
f-s-f reduces to s-f ?		4.5%	1.5%	

[10] We take x_{EX} to include a constant term.

We will take up the discussion about the choice of the number of splines and its relation to bandwidth in Appendix 3. For now we will continue assuming that the 20 splines used by BL is the optimal number.

5.2 The choice of estimation technique

Having decided the model, one needs to choose the estimation technique. It would be easy just to go ahead with GMM, because it is the least restrictive technique we know. However, doing so we might loose efficiency, so we want to use it only if we have to. Before we look at the data, let's first look at the model. It is an identical regressor system of equations.

$$\omega = \begin{pmatrix} [(x_{EN}, x_{EX}), D_s] & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & [(x_{EN}, x_{EX}), D_s] \end{pmatrix} * \begin{pmatrix} \begin{pmatrix} \beta_1^x \\ \Gamma_1 \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \beta_M^x \\ \Gamma_M \end{pmatrix} \end{pmatrix} + \varepsilon$$

For simplicity in the following we define

$$x \quad = \quad (x_{EN}, x_{EX}, D_s), \quad X = (I_M \otimes x), \quad \beta_j = \begin{pmatrix} \beta_j^x \\ \Gamma_j \end{pmatrix} \quad and \quad \beta = \begin{pmatrix} \beta_1 \\ : \\ \beta_M \end{pmatrix}$$

$$so \quad : \quad \omega = \begin{pmatrix} x & 0 & \cdots & 0 \\ 0 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x \end{pmatrix} \begin{pmatrix} \beta_1 \\ : \\ \beta_M \end{pmatrix} + \varepsilon = X\beta + \varepsilon$$

The dimension of x is $n \times (k+s)$, where n is number of observations, k is the total number of variables in (x_{EN}, x_{EX}), and s is the number of splines. The dimension of β_j is thus $(k+s) \times 1$. For the data used in BL n=2029 observations, $M = 6$ markets, $k = 31$ variables including a constant term and 2 possibly endogenous variables and $s = 20$ is the number of splines

If $\varepsilon \sim N(0, \sigma^2 I_{n*M})$ and $E(X'\varepsilon) = 0$ then, obviously, we should use an OLS estimator of β and we would be able to rely on various t, χ^2, and F statistics. However, the alternatives to such error term characteristics are plenty.

We will test the explanatory variables for endogeneneity and the error term for both normality and homoscedasticity before deciding on an estimator.

5.2.1 *Ex ante* expectations

Ex ante, we expect to find *endogeneity* of total expenditure as well as a rich covariance structure. We do not have ex ante expectations about the normality of the error term.

Endogeneity: Browning, Bourguignon, Chiappori and Lechene (1994) work with similar data for Canadian households with no children and stress the need to address endogeneity of total expenditure in these kind of Data. Likewise Blundell, Browning, and Crawford (2003) and Blundell, Duncan and Pendakur (1998) find endogenity of total expenditure in similar British data and we expect to find it here as well.

Heteroscedasticity: We expect to reject homoscedasticity against:

- *Groupwise homoscedasticity.* We are estimating several markets in one go and have no reason to assume that the variance of the error term is the same across markets. As such we would expect to reject homoscedasticity against groupwise homoscedasticity in an Extended Goldfeld Quandt test.

- *Between-market covariances.* In demand systems we have every reason to expect correlation across markets. Were this not so, we would be at least as well off estimating the markets separately. Thus, we expect non-zero covariances between the error term on one market with that on another market. This would be a typical SUR specification. We will modify the Extended Goldfeld Quandt test to test between (H0) groupwise homoscedasticity with zero error covariances and (H1) groupwise homoscedasticity with between-market covariances. We expect this test to reject the assumption of simple groupwise homoscedasticity.

- *Within market heteroscedasticity.* The join points are organized so that each spline interval has the same proportion of observations. We have no reason to assume that this organizing principle also guides the local extrema of the 'true model' (pardon our Bayesian). This has at least two consequences. Our unrestricted model will be biased if the local extrema of the 'true model' are not exactly co-located with join points of the model we use . And this bias will lead to heteroscedasticity even if the true model has 20 splines and homoscedasticity. (Note that expanding the number of splines to overcome this bias will lead to a higher variance of the estimator. The proper number of splines is a compromise between bias and variance of the estimator)

Ex ante we have no particular expectation of the error term for one family on one market being correlated with that of an other family on the same or another market. We will not test for it.

Normality:... makes things so much easier. The Breusch Pagan test for heteroscedasticity is sensitive to the assumption of normality, and the Extended Goldfeld Quandt Likelihood Ratio hinges on the assumption of normality. But actually we do not know enough about the Data Generating Process to assume normality. We can test jointly for homoscedasticity and normality using a Jarque-Bera test. However, as we expect to reject homoscedasticity, this test can not conclusively reject normality. We will look for other indicators of non-normality, such as differences between the Breusch-Pagan test statistic and the Koenker Basset statistic. Of course we will do some residual plots as well. This uncertainty as to the distribution alone could justify using a method of moments based estimator.

We will address all these concerns in the following. Let's look at the data.

5.2.2 Testing for endogeneity

We have two possible instruments summarized by z_{EN}. They are: ln(total income) and [ln (total income)]^2 Define $z = (z_{EN}, x_{EX}, D_h)$ and $Z = (I_M \otimes z)$. We need to test if Z and X are correlated, and we would like to test if Z and e are uncorrelated. However, the latter is not possible to test, because we do not have an overidentified case. Thus we cannot do Sargans overidentification test.

Before we can test for endogeneity, we must test if instruments are weak. We have every reason to expect that income is strongly correlated with total expenditure via the budget constraint. We will test for weak instruments nevertheless.

We test for weak instruments by doing OLS on $x_{EN} = z_{EN}\pi + (x_{EX}, D_h)\phi + E$. Then we test if $\hat{\pi}$ is significantly different from zero. π is of dimension (2×2) and we can do F-tests that a column in π is zero. If a column in $\hat{\pi}$ is not significantly different from zero, it would indicate that our instruments jointly do a poor job of explaining the variation in a possibly endogenous x variable. We find F-statistics of 505.9 for ln(total expenditure) and 509.1 for [ln(total expenditure)]^2 These are significant by a large margin since $F_{0,95}$ (2, 2029-51) = 3.000. We are not interested in leaving any of the instruments out, so we will not dwell at the t-values.

Now we can proceed to testing for endogeneity by a residual augmented regression: $\omega = X\beta + (I_M \otimes \hat{E})\vartheta + u$. Our $H0$ is: $\vartheta = 0$, where ϑ is a $(M * 2) \times 1$ vector. The relevant F-statistic is 6.044 and significant, as $F_{0,95}(6 * 2, 6 * (2029 - 51 - 2) = 1.753$. We conclude that the total expenditure transformations, x_{EN}, are endogenous and proceed instrumenting. Effectively we are testing OLS against IVE, which can be argued to be a 'cheap' solution and maybe we ought to have tested GLS against GMM, as we will see that there is heteroscedasticity in the model. We also recall that the F-statistic we use is based on the relation of two squared standardized normal distributions, making normality crucial in its interpretation. However, all the literature we know also suggests that expenditure is endogenous in this setting [see Blundell, Browning, and Crawford (2003) and Blundell, Duncan and Pendakur (1998)]. We proceed with confidence in our conclusion.

Table A1.3: Tests for weak instruments and endogeneity

Test for weak instruments	
ln exenditure jointly explained by *ln income* and *(ln income)* ^{2}	
F	505.85
df1 (# possibly endogenous variables)	2
df2 (observations-# estimated parameters)	2029-51
prob F	0.00%
(ln exenditure) ^{2} jointly explained by *ln income* and *(ln income)* ^{2}	
F	509.05
df1 (# of possibly endogenous variables)	2
df2 (observations-# estimated parameters)	2029-51
prob F	0.00%

Test for endogeneity of expenditure variables	
- using a residual augmented regression	
F:	6.044
df1 (# of possibly endogenous variables)	6*2
df2 (observations-# estimated parameters)	6*(2029-51-2)
prob F	0.00%

5.2.3 Testing for heteroscedasticity

As we have ascertained that there is endogeneity in the model, we will test using IVE estimates. First, we test for homoscedsticity, $\varepsilon \sim (0, \sigma^2 I_{M*n})$, against a very general specification of heteroscedasticity, $\varepsilon_i \sim (0, \sigma_i^2)$. We use a Breusch-Pagan/Godfrey Langrange Multiplier test. This test needs a specific $H1$ and we choose $H1 : \sigma_i^2 = Z_i\delta, \;\; i = 1, ..., M*n$. Under $H0$ and assuming normality the variance of $\varepsilon_i^2 = 2\sigma^4$ and the LM statistic becomes:

$$LM_{Breusch-Pagan} = \frac{1}{2}(ESS_{\frac{\hat{\varepsilon}_i^2}{\hat{\varepsilon}'\hat{\varepsilon}/(M*n)} = Z_i\delta + u}) \sim \chi^2(M*(k+s)-1)$$

Testing for all markets jointly the $LM_{Breusch-Pagan}$-statistic is 3030 and $\chi^2_{0,95}(6*51-1)$ is 346.73, so we can reject homoscedasticity. As described in Greene (2002) the statistic is sensitive to the assumption of $var(\hat{\varepsilon}_i^2) = 2\sigma^4$. We only know that this holds if $\hat{\varepsilon}_i$ is normally distributed. Greene attributes a more robust test for heteroscedastiticy to Koenker and Basset.

$$LM_{Koenker-Basset} = \frac{1}{V}(ESS_{\hat{\varepsilon}_i^2 = Z_i\delta + u}) \sim \chi^2(M*(k+s)-1), \quad \text{where } V = \frac{1}{M*n}\sum_{i=1}^{M*n}\left(\hat{\varepsilon}_i^2 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{M*n}\right)$$

The $LM_{Koenker-Basset}$-statistic is 1701 - a little more than half that of $LM_{Breusch-Pagan}$. However, both are significant by a large margin and we reject homoscedasticity as we expected to do. We will discuss the large difference between the two estimators when we discuss the normality assumption below.

The heteroscedasticity might be more structured than $\varepsilon_i \sim (0, \sigma_i^2)$ and for efficiency reasons we want to test for this before going ahead with the very general specification. So first we test for homoscedasticity, $\varepsilon \sim N(0, \sigma^2 I_{M*n})$, against a hybrid: groupwise homoscedasticity - that is: $H1 : \varepsilon \sim N(0, \sum \otimes I_n)$ with

$$\sum = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_M^2 \end{pmatrix}$$

Note that this is not the typical SUR variance specification, as it does not allow for covariance between markets. We use an Extended Goldfeld-Quandt test as described in Johnston and Dinardo(1997). The test is a likelihood ratio test with

$$LR = (M*n) * \ln\hat{\sigma}^2 - \sum_{j=1}^{M} n * \ln\hat{\sigma}_j^2 \sim^a \chi^2(M-1)$$

This test relies on splitting data up in smaller groups. We split X by the M markets. What is normally problematic about this test, is that $\hat{\sigma}_j = \left(\omega_j - x_j\hat{\beta}\right)'\left(\omega_j - x_j\hat{\beta}\right)/n_j$, with $\hat{\beta} = \left(X'\left(\hat{\Sigma}\otimes I_n\right)^{-1}X\right)X'\left(\hat{\Sigma}\otimes I_n\right)^{-1}\omega$. In other words $\hat{\sigma}_j$ depends on $\hat{\beta}$, which depends on $\hat{\Sigma}$ which is constituted by the $\hat{\sigma}_j$'s.... Thus, one normally needs to do several iterations to get a good estimate. However, as we have an identical regressor model, iterations are uneccessary. We show this below:

$$
\begin{aligned}
\hat{\beta}_{IVE} &= \left((I_M \otimes x'z)\left((I_M \otimes z)'\left(\sum \otimes I_n\right)(I_M \otimes z)\right)^{-1}(I_M \otimes z'x)\right)^{-1} \\
&\quad *(I_M \otimes x'z)\left((I_M \otimes z)'\left(\sum \otimes I_n\right)(I_M \otimes z)\right)^{-1}(I_M \otimes z)'y \\
&\Updownarrow \\
\hat{\beta}_{IVE} &= \left((I_M \otimes x'z)\left(\sum \otimes z'z\right)^{-1}(I_M \otimes z'x)\right)^{-1}(I_M \otimes x'z)\left(\sum \otimes z'z\right)^{-1}(I_M \otimes z)'y \\
&\Updownarrow \\
\hat{\beta}_{IVE} &= \left(\sum^{-1} \otimes x'z\,(z'z)^{-1}\,z'x\right)^{-1}(\sum^{-1} \otimes x'z\,(z'z)^{-1}\,z')y \\
&\Updownarrow \\
\hat{\beta}_{IVE} &= (I_M \otimes (x'z\,(z'z)^{-1}\,z'x)^{-1}x'z\,(z'z)^{-1}\,z')y
\end{aligned}
$$

Thus $\hat{\beta}$ does not depend on the variance structure. What this means, is that we can go ahead and calculate the Goldfeld-Quandt Likelihood Ratio in one go. We find that it is 3012. Since $\chi^2_{0,95}(5) = 11.07$ we also reject this specification of the error term. Recall however, that this is a maximum likelihood statistic and as such is based on the assumption of normality. We expect the statistic to be quite sensitive to this assumption but have not studied it further.

We now do our own extension of the Goldfeld-Quandt test. We take our last H1, and make it our new H0 for testing the typical SUR specification of between-market covariance.

$$
H0 : \varepsilon \sim N\left(0, \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_M^2 \end{bmatrix} \otimes I_n\right) \qquad H1 : \varepsilon \sim N\left(0, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \sigma_{M1} & \cdots & \cdots & \sigma_{MM} \end{bmatrix} \otimes I_n\right)
$$

Our new test statistic is

$$
LR = \left[\sum_{j=1}^{M} n \ln \hat{\sigma}_j^2\right] - n \ln \left|\hat{\Sigma}_{H1}\right| \sim^a \chi^2(M-1)
$$

We get a test statistic of 434 and reject the H0. The same caveat as above applies here - the statistic is sensitive to the normality assumption.

We have showed that there are strong indications of between market covariances and the next logical thing to do is investigating for within-market heteroscedasticity. So we return to the Koenker-Basset statistic, now for the individual markets. The results can be seen in the table A1.4 below.

Table A1.4: Tests for Heteroscedasticity

Breusch Pagan and Koenker Basset - testing for general heteroscedasticity			
	Test statistic	Degrees of freedom	Prob
Breusch Pagan	3029.97	305	0.000%
Koenker Basset	1701.17	305	0.000%
H0: Homoscedasticity We regress $e_{t,IVE}^2$ on all instrument variables			

Extended Goldfeld-Quandt (Johnston & Dinardo p 170) - testing for homoscedasticity against marketwise homoscedasticity		
Likelihood Ratio	Degrees of freedom	Prob
3012.36	6-1=5	0.000%
H0: Homoscedasticity H1: Marketwise Homoscedasticity		

Extended Goldfeld-Quandt (our extension) - testing for marketwise homoscedasticity agaist marketwise covariance		
Likelihood Ratio	Degrees of freedom	Prob
433.79	6*6-6=30	0.000%
H0: Marketwise Homoscedasticity H1: Marketwise Homoscedasticity and Between Market Covariance		

Koenker Basset - testing for intra-market heteroscedasticity			
	Test statistic	Degrees of freedom	Prob
Food at Home	111.20	50	0.000%
Household Operations	102.10	50	0.002%
Women's Clothing	120.48	50	0.000%
Men's Clothing	136.45	50	0.000%
Children's Clothing	129.28	50	0.000%
Vices	90.51	50	0.040%
H0: within-market homoscedasticity, H1: within-market heteroscedasticity We regress $e_{t,IVE}^2$ on all instrument variables			

We reject the hypothesis that within market errors are homoscedastic, so we have every reason to go forward with GMM.

5.2.4 Normality

As we saw in the last figure there is a large deviation between the Breusch-Pagan statistic and the Koenker-Basset statistic. This might well indicate that the error terms do not satisfy normality. A Jacques Bera test of normality on the GMM residuals corroborates this. The Jacques Bera statistics we calculated in Gauss are reproduced in the table A1.5 along with a Shapiro-Francia W' test for normal data and a Skewness/Kurtosis test for normality from Stata.

Table A1.5: Tests for Normality

Jarque Bera test for normality and homoscedasticity				
Market	Statistic	Prob	Skewness	Kurtosis
All Markets	4587.29	0.000	0.686	5.676
Food at Home	43.99	0.000	0.191	3.611
Household Operations	468.04	0.000	0.844	4.639
Women's Clothing	450.85	0.000	0.943	4.334
Men's Clothing	651.10	0.000	0.926	5.067
Children's Clothing	860.57	0.000	1.077	5.353
Vices	462.01	0.000	1.001	4.207

Shapiro-Francia W' test for normal data					
Market	Obs	W'	V'	z	Prob>z
Food at Home	2029	0.99296	7.531	4.205	0.00001
Household Operations	2029	0.96188	40.778	6.586	0.00001
Women's Clothing	2029	0.95227	51.060	6.844	0.00001
Men's Clothing	2029	0.95633	46.714	6.743	0.00001
Children's Clothing	2029	0.94467	59.194	7.007	0.00001
Vices	2029	0.94397	59.937	7.020	0.00001

Skewness/Kurtosis test for normality		
Market	Pr(Skewness)	Pr(Kurtosis)
Food at Home	0.000	0.000
Household Operations	0.000	0.000
Women's Clothing	0.000	0.000
Men's Clothing	0.000	0.000
Children's Clothing	0.000	0.000
Vices	0.000	0.000

Again, we stress that the Jarque Bera test is not purely a test of normality. H0 is normality *and homoscedasticity* (Greene 2002). Since we have found strong indications of heteroscedasticity we cannot conclusively use it to reject normality. So let's use an old fashioned but useful technique: looking at a residual plot. Sometimes the eye is a very good calibrator. And our eyes confirm what the Shapiro-Francia W' test tells us: GMM-errors do not appear to be normally distributed. We present normal fractile plots and kernel density estimates from Stata below.

Figure A1.1: Normal probability plots and density plots

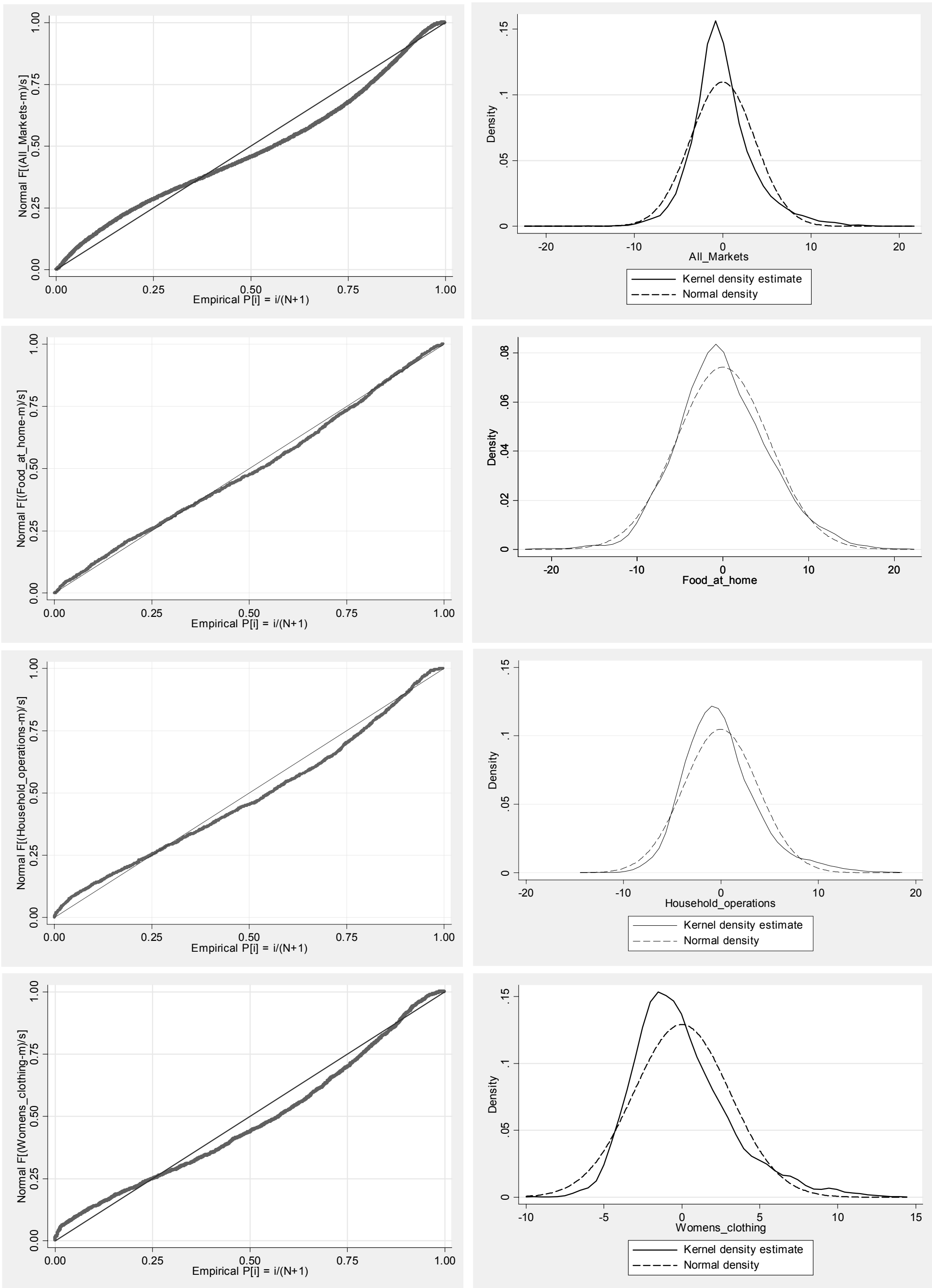
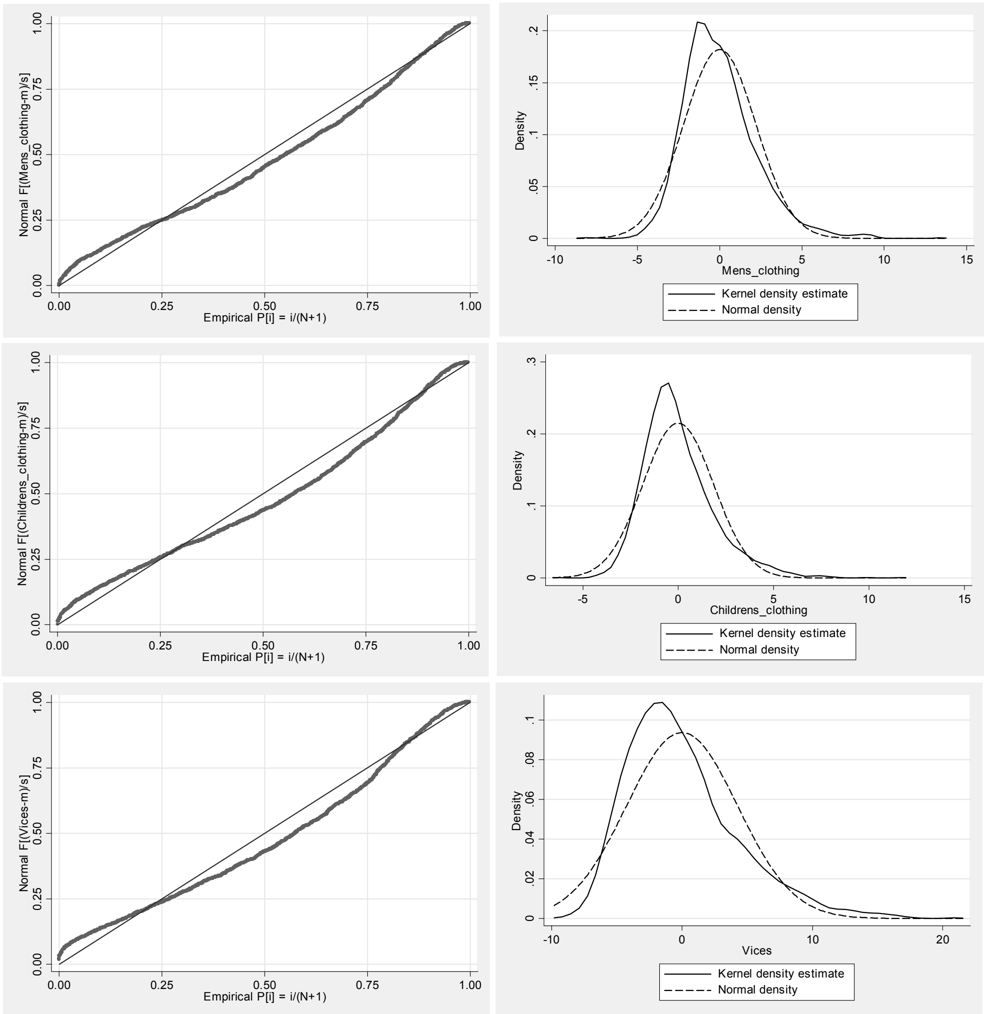


Figure A1.2: Normal probability plots and density plots



The tendency we see in these plots is, that the most dense areas are below zero and that the positive-residual tails are longer. Is this enough to render our tests useless? We don't know. But we feel that together with the indications of endogeneity and heteroscedasticity we have every reason to go ahead with GMM, because it is the least restrictive method of estimation we know. We leave the issue of normality now and we will neither study the distribution of the errors nor the consequences for the minimum chi-squared test statistics further in this appendix, even if it is well worth further study. But we will be cautious when interpreting results. We will use efficient GMM as our method of estimation, with a Lars Peter Hansen efficient and feasible weighting matrix $\hat{W}_{lph} = \widehat{Z'\Omega Z}$. We construct it in $M \times M$ blocks of $(k+s) \times (k+s)$ matrices relating market j with market l : $\hat{W}_{j,l} = \frac{1}{n}\sum_{i=1}^{n} Z_i' Z_i \hat{\varepsilon}_{j,i}^{IVE} \hat{\varepsilon}_{l,i}^{IVE}, \quad j,l = 1,...,M.$

5.3 Splines and Restrictions

Spline smoothing is the central feature of the fully parametric model ($3a$) and we will use this paragraph to discuss the way we have implemented it and the way restrictions can be imposed over the splines. First we show two ways of constructing splines and argue why we have chosen the method we use. Second we reconstruct the results of BL and show four things:

- BL appear to apply restrictions to the splines of the X variables only
- While this is an asymptotically valid approach it yields results that are quite different (if not significantly so) from the proper GMM results
- The reason for this difference is that the BL-method does not restrict the instrument variable splines.
- We propose an estimator which works in both the exactly identified case and with overidentification. It is computationally much faster than full GMM, yet it can be used for the Minimum χ^2 grid search with virtually identical results.

5.3.1 Splines

Greene (2002) demonstrates an approach to making a piecewise linear and continuous regression function. It involves selecting a number of threshold points and defining dummies for them. Suppose we have a variable of interest $\rho' = (\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8, \rho_9)$ and three threshold points $\rho^* = (\rho^*_{s_1}, \rho^*_{s_2}, \rho^*_{s_3})$. Now define three dummy variables $d_p = 1$ if $\rho \geq \rho^*_p$ for $p = s_1, s_2, s_3$. Instead of doing the linear regression on $\omega_j = (x_{EN}, x_{EX})\beta_{x,j} + \Gamma_j \rho + u_j, \quad j = 1, .., M$ we can supplement the ρ variable with spline vectors $D_p = d_p * (\rho - \rho^*_p)$. Define $D_s = (\rho, D_1, D_2, D_3)$. Our regression of interest becomes: $\omega_j = (x_{EN}, x_{EX})\beta_{x,j} + D_s\Gamma_j + u_j$. This could look like:

$$\begin{pmatrix} \omega_{j1} \\ \omega_{j2} \\ \omega_{j3} \\ \omega_{j4} \\ \omega_{j5} \\ \omega_{j6} \\ \omega_{j7} \\ \omega_{j8} \\ \omega_{j9} \end{pmatrix} = (x_{EN}, x_{EX})\beta_{x,j} + \begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ \rho_2 & 0 & 0 & 0 \\ \rho_3 & \rho_3 - \rho^*_{s_1} & 0 & 0 \\ \rho_4 & \rho_4 - \rho^*_{s_1} & 0 & 0 \\ \rho_5 & \rho_5 - \rho^*_{s_1} & 0 & 0 \\ \rho_6 & \rho_6 - \rho^*_{s_1} & \rho_6 - \rho^*_{s_2} & 0 \\ \rho_7 & \rho_7 - \rho^*_{s_1} & \rho_7 - \rho^*_{s_2} & 0 \\ \rho_8 & \rho_8 - \rho^*_{s_1} & \rho_8 - \rho^*_{s_2} & \rho_8 - \rho^*_{s_3} \\ \rho_9 & \rho_9 - \rho^*_{s_1} & \rho_9 - \rho^*_{s_2} & \rho_9 - \rho^*_{s_3} \end{pmatrix} \Gamma_j + u_j$$

The disadvantage of this method is, that to get an idea of the slope for a certain value of ρ, one has to add all the Γ coefficients up to the relevant spline. The slope in point ρ_7 is thus $\Gamma_{j,1} + \Gamma_{j,2} + \Gamma_{j,3}$ on market j. And if one for example wants to impose the restriction that the slope is zero between $\rho^*_{s_2}$ and $\rho^*_{s_3}$, then this implies restricting $\Gamma_{j,1} + \Gamma_{j,2} + \Gamma_{j,3} = 0$. We find this cumbersome. Therefore we use an equivalent approach which is easier to interpret and much easier to impose restrictions on. Instead of the regression above we would do the following regression.

$$\begin{pmatrix} \omega_{j1} \\ \omega_{j2} \\ \omega_{j3} \\ \omega_{j4} \\ \omega_{j5} \\ \omega_{j6} \\ \omega_{j7} \\ \omega_{j8} \\ \omega_{j9} \end{pmatrix} = (x_{EN}, x_{EX})\beta_{x,j} + \begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ \rho_2 & 0 & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \rho_3 - \rho^*_{s_1} & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \rho_4 - \rho^*_{s_1} & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \rho_5 - \rho^*_{s_1} & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \rho_6 - \rho^*_{s_2} & 0 \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \rho_7 - \rho^*_{s_2} & 0 \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \boldsymbol{\rho}^*_{s_3} - \rho^*_{s_2} & \rho_8 - \rho^*_{s_3} \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \boldsymbol{\rho}^*_{s_3} - \rho^*_{s_2} & \rho_9 - \rho^*_{s_3} \end{pmatrix} \Gamma_j + u_j$$

Here the $\Gamma_{j,p}$ coefficient represents the full slope of spline p on market j. This makes testing restrictions on individual splines much easier. Recall $\beta_j = (\beta'_{x,j}, \Gamma'_j)'$. Let's say we want to estimate β_j under the restriction of a slope-flat-slope model where

- The slope over the second spline is zero and
- The slope is the same over the third and fourth spline,

To estimate under these restrictions we could just manually remove the second spline and aggregate the third and fourth splines. Then we would do the regression and get $\hat{\beta}^R_j$. We can instead do this restriction of the splines by post-multiplying $x = (x_{EN}, x_{EX}, D_s)$ by r

$$\begin{pmatrix} \omega_{j1} \\ \omega_{j2} \\ \omega_{j3} \\ \omega_{j4} \\ \omega_{j5} \\ \omega_{j6} \\ \omega_{j7} \\ \omega_{j8} \\ \omega_{j9} \end{pmatrix} = \underbrace{\left[x_{EN}, x_{EX}, \begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ \rho_2 & 0 & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \rho_3 - \rho^*_{s_1} & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \rho_4 - \rho^*_{s_1} & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \rho_5 - \rho^*_{s_1} & 0 & 0 \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \rho_6 - \rho^*_{s_2} & 0 \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \rho_7 - \rho^*_{s_2} & 0 \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \boldsymbol{\rho}^*_{s_3} - \rho^*_{s_2} & \rho_8 - \rho^*_{s_3} \\ \boldsymbol{\rho}^*_{s_1} & \boldsymbol{\rho}^*_{s_2} - \rho^*_{s_1} & \boldsymbol{\rho}^*_{s_3} - \rho^*_{s_2} & \rho_9 - \rho^*_{s_3} \end{pmatrix} \right]}_{x} \underbrace{\begin{pmatrix} I_k & 0_{k\times 1} & 0_{k\times 1} \\ 0_{1\times k} & 1 & 0 \\ 0_{1\times k} & 0 & 0 \\ 0_{1\times k} & 0 & 1 \\ 0_{1\times k} & 0 & 1 \end{pmatrix}}_{r} \boldsymbol{\beta}^R_j + u_j$$

$$\Updownarrow$$

$$\begin{pmatrix} \omega_{j1} \\ \omega_{j2} \\ \omega_{j3} \\ \omega_{j4} \\ \omega_{j5} \\ \omega_{j6} \\ \omega_{j7} \\ \omega_{j8} \\ \omega_{j9} \end{pmatrix} = \underbrace{\left[x_{EN}, x_{EX}, \begin{pmatrix} \rho_1 & 0 \\ \rho_2 & 0 \\ \rho^*_{s_1} & 0 \\ \rho^*_{s_1} & 0 \\ \rho^*_{s_1} & 0 \\ \rho^*_{s_1} & \rho_6 - \rho^*_{s_2} \\ \rho^*_{s_1} & \rho_7 - \rho^*_{s_2} \\ \rho^*_{s_1} & \rho_8 - \rho^*_{s_2} \\ \rho^*_{s_1} & \rho_8 - \rho^*_{s_2} \end{pmatrix} \right]}_{x^R} \boldsymbol{\beta}^R_j + u_j \qquad j = 1, ..., M$$

Here r has dimension $(k+s) \times (k + \#\textit{sloping parts in the restricted model})$. Having obtained the restricted estimator, $\hat{\beta}^R_j$, we face the problem, that it is not immediately row-by-row commensurable with the unrestricted β^{UR}_j as they *differ in dimension*. Thus we can NOT calculate $\left(\hat{\beta}^{UR} - \hat{\beta}^R\right)$ directly, as we want to for the Minimum Chi Squared estimator (see below). We easily solve this by premultiplying β^R_j by r.

$$\beta_j^{R-slope-flat-slope} = \begin{pmatrix} \beta_{x,j}^R \\ \Gamma_{j,\mathbf{1}}^R \\ \Gamma_{j,\mathbf{2}}^R \end{pmatrix} \text{ is } \textit{not} \text{ directly commensurable with } \beta_j^{UR} = \begin{pmatrix} \beta_{j,x}^{UR} \\ \Gamma_{j,\mathbf{1}}^{UR} \\ \Gamma_{j,\mathbf{2}}^{UR} \\ \Gamma_{j,\mathbf{3}}^{UR} \\ \Gamma_{j,\mathbf{4}}^{UR} \end{pmatrix}$$

$$r\ \beta_j^{R-slope-flat-slope} = \begin{pmatrix} \beta_{x,j}^R \\ \Gamma_{j,\mathbf{1}}^R \\ 0 \\ \Gamma_{j,\mathbf{2}}^R \\ \Gamma_{j,\mathbf{2}}^R \end{pmatrix} \textit{ is} \text{ row-by-row commensurable with } \beta_j^{UR} = \begin{pmatrix} \beta_{j,x}^{UR} \\ \Gamma_{j,\mathbf{1}}^{UR} \\ \Gamma_{j,\mathbf{2}}^{UR} \\ \Gamma_{j,\mathbf{3}}^{UR} \\ \Gamma_{j,\mathbf{4}}^{UR} \end{pmatrix}$$

The method is obviously easy to scale to any number of splines and markets. In a system of identical regressors the relevant restriction matrix is $R = (I_m \otimes r)$, so the relevant restricted regression is $\omega = XR\beta^R + \varepsilon$.

5.3.2 Restricted estimators

When we first did our estimations, we did not get the same results as BL. Our β estimates looked different and our grid search indicated different optimal locations of spline points. It took a fair bit of reverse engineering and the kind advice of Martin Browning to find out why this was the case. Below, we present our understanding of the BL-estimator and then reconstruct it.

To understand the BL-estimator we need to understand its *raison-d'etre.* The Minimum Chi squared statistic is $\left(\hat{\beta}^{UR} - \tilde{\beta}^R\right)' C_{\hat{\beta}^{UR}}^{-1} \left(\hat{\beta}^{UR} - \tilde{\beta}^R\right)$, where $\tilde{\beta}^R = R\hat{\beta}^R$ is the commensurable version of $\hat{\beta}^R$. Under the restriction that the marginal effect of ρ on ω is zero for some splines or that it is equal across some adjacent splines, estimating β^R 'from the ground up' means: i) transform X to $X^R = XR$ and Z to $Z^R = ZR$, ii) do IVE estimation using these restricted variables, iii) use the residuals to calculate the weighting matrix, and then iv) perform the GMM estimation. Very loosely speaking, this gives a 'Likelihood Ratio' type statistic in the sense, that we estimate both the unrestricted and the restricted estimator from the ground up..

The problem with Grid searching over this statistic is, that it can require thousands of estimations and is quite time consuming even on fast computers. An attractive alternative would be to use a 'Wald-type' statistic, with $\hat{\beta}^R$ expressed as some function of the unrestricted $\hat{\beta}^{UR}$. To find this function and to understand the BL estimator, we first look at implementing restrictions in an OLS and a GLS setup.

5.3.3 Introducing the $\hat{\beta}^R_{BL}$ estimator

Estimating β^R in an OLS or GLS setup is simple. Minimizing squared residuals in $\omega = X\ R\ \beta^R + \varepsilon$ we get:

$$\begin{aligned}
\hat{\beta}^R_{OLS} &= (R'X'XR)^{-1}R'X'\omega = (R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{OLS}}R)^{-1}R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{OLS}}\hat{\beta}^{UR}_{OLS} \\
\hat{\beta}^R_{GLS} &= (R'X'\Omega^{-1}XR)^{-1}R'X'\Omega^{-1}\omega = (R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{GLS}}R)^{-1}R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{GLS}}\hat{\beta}^{UR}_{GLS}
\end{aligned}$$

This is very smart indeed, as is provides us with a simple closed form expression for $\hat{\beta}^R$. Can we always use this relationship between $\hat{\beta}^R$ and $\hat{\beta}^{UR}$, even if we use IVE or GMM? We tried to apply it using our unrestricted GMM estimators and got results almost identical to those of BL. We concluded, that this was their method of estimation (tables are attached behind appendix 3) and refer to it as the BL-estimator, $\hat{\beta}^R_{BL} = (R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}}R)^{-1}R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}}\hat{\beta}^{UR}$, even if it has a different name. Next, we asked ourselves three questions:

5.3.4 Is $\hat{\beta}^R_{BL}$ consistent when based on the unrestricted GMM estimator ?

$$\begin{aligned}
p\lim(\hat{\beta}^R_{BL}) &= p\lim\left[\frac{1}{M*n}R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{GMM}}R\right]^{-1} * p\lim\frac{1}{M*n}\left[R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{GMM}}\hat{\beta}^{UR}_{GMM}\right] \\
&= p\lim\left[\frac{1}{M*n}R'X'Z(Z'\Omega Z)^{-1}Z'XR\right]^{-1} * p\lim\left[\frac{1}{M*n}R'X'Z(Z'\Omega Z)^{-1}Z'\omega\right]
\end{aligned}$$

assuming $\beta^{UR} = R\ \beta^R$ holds, then $\omega = X\beta^{UR} + \varepsilon = X\ R\ \beta^R + \varepsilon$ and we get

$$\begin{aligned}
p\lim(\hat{\beta}^R_{BL}) &= p\lim\left[\frac{1}{M*n}R'X'Z(Z'\Omega Z)^{-1}Z'XR\right]^{-1} p\lim\left[\frac{1}{M*n}R'X'Z(Z'\Omega Z)^{-1}Z'(RX\beta^R + \varepsilon)\right] \\
&= \beta^R + p\lim\underbrace{\left[\frac{1}{M*n}R'X'Z(Z'\Omega Z)^{-1}Z'XR\right]^{-1}}_{\substack{\neq 0 \\ \text{if } \Sigma_{z'x} \neq 0}} p\lim\left[\underbrace{\frac{1}{M*n}R'X'Z(Z'\Omega Z)^{-1}}_{\substack{\neq 0 \\ \text{if } \Sigma_{z'x} \neq 0}}\quad\underbrace{Z'\varepsilon}_{\substack{= 0 \\ \text{if } \Sigma_{z'e} = 0}}\right] \\
&= \beta^R
\end{aligned}$$

We see that $\hat{\beta}^R_{BL}$ is a consistent estimator for β^R.

5.3.5 Is $\hat{\beta}^R_{BL}$ a closed form expression for $\hat{\beta}^R$ under IVE and GMM?

To check this we minimize squared residuals in $(ZR)'\omega = (ZR)\prime XR + (ZR)'\varepsilon$ and get

$$\hat{\beta}^R_{IVE} = (R'X'ZR(R'Z'ZR)^{-1}R'Z'XR)^{-1}R'X'ZR(R'Z'ZR)^{-1}R'Z'\omega \tag{4}$$

This does not immediately look equivalent to the BL estimator. But we can rewrite it. Let $X_R = RX$ and $Z_R = RZ$. Let $\hat{X}_R^{Z_R}$ be the predicted value of X_R from the regression on the *restricted* instrument variables Z_R: $X_R = Z_R\Pi + u$ and let $\hat{X}_R^Z$ be the predicted value of X_R from the regression on the *un*-restricted instrument variables Z: $X_R = Z\gamma + \xi$. Now, we can express(4) in a compact form

$$\hat{\beta}^R_{IVE} = (\hat{X}_R^{Z_R\prime}\hat{X}_R^{Z_R})^{-1}\hat{X}_R^{Z_R\prime}\omega \tag{5}$$

To find out if this is identical to the $\hat{\beta}^R_{BL}$ we need to understand how $\hat{X}_R^{Z_R}$ by $\hat{X}_R^Z$ relate? This is easily understood by looking at the following

$$\mathbf{x}^R = \left[x_{EN}, x_{EX}, \begin{pmatrix} \rho_1 & 0 \\ \rho_2 & 0 \\ \rho^*_{s_1} & 0 \\ \rho^*_{s_1} & 0 \\ \rho^*_{s_1} & 0 \\ \rho^*_{s_1} & \rho_6 - \rho^*_{s_2} \\ \rho^*_{s_1} & \rho_7 - \rho^*_{s_2} \\ \rho^*_{s_1} & \rho_8 - \rho^*_{s_2} \\ \rho^*_{s_1} & \rho_8 - \rho^*_{s_2} \end{pmatrix} \right], \quad \mathbf{z} = \left[z_{EN}, x_{EX}, \begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ \rho_2 & 0 & 0 & 0 \\ \rho^*_{s_1} & \rho_3 - \rho^*_{s_1} & 0 & 0 \\ \rho^*_{s_1} & \rho_4 - \rho^*_{s_1} & 0 & 0 \\ \rho^*_{s_1} & \rho_5 - \rho^*_{s_1} & 0 & 0 \\ \rho^*_{s_1} & \rho^*_{s_2} - \rho^*_{s_1} & \rho_6 - \rho^*_{s_2} & 0 \\ \rho^*_{s_1} & \rho^*_{s_2} - \rho^*_{s_1} & \rho_7 - \rho^*_{s_2} & 0 \\ \rho^*_{s_1} & \rho^*_{s_2} - \rho^*_{s_1} & \rho^*_{s_3} - \rho^*_{s_2} & \rho_8 - \rho^*_{s_3} \\ \rho^*_{s_1} & \rho^*_{s_2} - \rho^*_{s_1} & \rho^*_{s_3} - \rho^*_{s_2} & \rho_9 - \rho^*_{s_3} \end{pmatrix} \right]$$

In terms of splines, we see that in this case z includes a perfect instrument (spline 1), a bogus instrument (column 2), and two instruments (spline 3 and 4) that jointly explain all the variation in the second spline of x^R. Since the Z variable contains all the disjunct splines used in constructing the spline part of X_R, Z can explain 100% of the spline variation in X_R. Accidental correlation between the unrestricted splines and the endogenous variables aside, this means that $\hat{X}_R^{Z_R} = \hat{X}_R^Z$ and so they can be interchanged. We can now rewrite (5)

$$\begin{aligned} \hat{\beta}^R_{IVE} &= (\hat{X}_R^{\mathbf{Z_R}\prime}\hat{X}_R^{\mathbf{Z_R}})^{-1}\hat{X}_R^{\mathbf{Z_R}\prime}\omega = (\hat{X}_R^{Z\prime}\hat{X}_R^Z)^{-1}\hat{X}_R^{Z\prime}\omega \\ &= (R'X'Z(Z'Z)^{-1}Z'R)^{-1}R'X'Z(Z'Z)^{-1}Z'\omega \\ &= (R'C^{-1}_{\hat{\beta}^{UR}_{IVE}}R)^{-1}R'C^{-1}_{\hat{\beta}^{UR}_{IVE}}\hat{\beta}^{UR}_{IVE} \\ &= \hat{\beta}^R_{BL} \end{aligned}$$

So the $\hat{\beta}^R_{BL}$ is a closed form estimator for IVE also[11]. We were quite impressed with the general applicability of the $\hat{\beta}^R_{BL}$ estimator, but curious if it would work with the GMM estimator.

Given $\varepsilon \sim (0, \Omega)$, define $P\prime P = (R'Z'\Omega ZR)^{-1}$. Minimizing squared residuals in the regression $(ZRP)'\omega = (ZRP)\prime XR + (ZRP)'\varepsilon$ we get

$$\begin{aligned}
\hat{\beta}^R_{GMM} &= \left[R'X'ZRP(P'R'Z'\Omega ZRP)^{-1}P'R'Z'XR\right]^{-1} R'X'ZRP(P'R'Z'\Omega ZRP)^{-1}P'R'Z'\omega \\
&= \left[R'X'ZR(R'Z'\Omega ZR)^{-1}R'Z'XR\right]^{-1} R'X'ZR(R'Z'\Omega ZR)^{-1}R'Z'\omega \\
&= \left[R'X'Z\mathbf{R}(\mathbf{R'}\mathbf{W}^{UR}_{lph}\mathbf{R})^{-1}\mathbf{R}'Z'XR\right]^{-1} R'X'Z\mathbf{R}(\mathbf{R'}\mathbf{W}^{UR}_{lph}\mathbf{R})^{-1}\mathbf{R}'Z'\omega \qquad (6) \\
&= \left[X'_R Z_R(W^R_{lph})^{-1}Z'_R X_R\right]^{-1} X'_R Z_R (W^R_{lph})^{-1} Z'_R \omega
\end{aligned}$$

And when the dimensions of X_R and Z_R are the same we get

$$\hat{\beta}^R_{GMM} = (X'_R Z_R)^{-1} Z'_R \omega = \left(I_m \otimes (x'_R z_R)^{-1} z'_R\right)\omega$$

Can we make practical use of this? In the present model, obviously we can, as we have an exactly identified model when we restrict splines on both X and Z. Thus $\hat{\beta}^R_{GMM} = \hat{\beta}^R_{IVE} = \hat{\beta}^R_{2SLS}$ which is very fast to calculate. We're finally ready to check if the $\hat{\beta}^{BL}$estimator is identical to the $\hat{\beta}^R_{GMM}$ estimator. We recall

$$\begin{aligned}
\hat{\beta}^R_{BL} &= (R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{GMM}} R)^{-1} R' C^{-1}_{\hat{\beta}^{\mathbf{UR}}_{GMM}} \hat{\beta}^{UR}_{GMM} \\
&= \left[R'X'Z(\mathbf{W}^{UR}_{lph})^{-1}Z'XR\right]^{-1} R'X'Z(\mathbf{W}^{UR}_{lph})^{-1}Z'\omega
\end{aligned}$$

As is now obvious, the $\hat{\beta}^R_{BL}$ estimator has a $(\mathbf{W}^{UR}_{lph})^{-1}$ in place of $\mathbf{R}\left(\mathbf{R'}\mathbf{W}^{UR}_{lph}\mathbf{R}\right)^{-1}\mathbf{R'}$ in (6) or $R\left(W^R_{lph}\right)^{-1}R'$. Thus we see, that leaving out the Rs around the weighting matrix is problematic as it gives incorrect weight to different spline variances.

We conclude that while $\hat{\beta}^R_{BL}$ is a closed form expression of $\hat{\beta}^R_{IVE}$ under some assumptions, it is not a proper closed form expression for $\hat{\beta}^R_{GMM}$.

5.3.6 Why use $\hat{\beta}^R_{BL}$?

Why was $\hat{\beta}^R_{BL}$ used by BL in a GMM framework when it differs from the very fast and simple expression for the restricted and exactly identified GMM estimator ?

We assume that BL used the β^R_{BL} estimator because of its general applicability - even if new instruments were to be implemented (causing exact identification to break down) the computer algorithm would still work. Or because it provides a fast way of estimating the $var\left(\hat{\beta}^R_{BL}\right) = (R'C^{-1}_{\hat{\beta}^{\mathbf{UR}}}R)^{-1}$.

[11]under the assumption that the unrestricted splines do not explain any variation in the possibly endogenous variables. We discuss this below.

In the overidentified case, with GMM we face the problem that we do not know the Ω part of W. Thus, we do not know the W either, and our estimator, $\hat{W}$, makes use of the restricted IVE residuals. If we have to recalculate these residuals imposing restrictions for every point in the Grid Search, the search will be dreadfully slow. Asymptotically $R'\hat{W}^{UR}R = R'Z'\widehat{\Omega^{UR}}ZR = R'Z'\Omega ZR = Z_R'\widehat{\Omega^R}Z_R = \hat{W}^R$. But the estimators might differ. Here is why: If unrestricted splines happen to be correlated with the endogenous variables, then $\hat{X}_{\mathbf{EN}}^{Z_R}$ and $\hat{X}_{\mathbf{EN}}^{Z}$ might differ. If they differ, then the residuals used in constructing $\hat{W}^{UR}$ and $\hat{W}^R$ will differ. We couldn't help wonder if the estimation of $\hat{W}$ from restricted IVE residuals rather than the *un*restricted IVE residuals only has a second order effect, so we cut corners and defined a sloppy GMM estimator $\hat{\beta}_{Z_R}^{X_R}$

$$\hat{\beta}_{Z_R}^{X_R} = (X_R'Z_R(R'\hat{W}_{lph}^{\mathbf{UR}}R)^{-1}Z_R'X_R)^{-1}X_R'Z_R(R'\hat{W}_{lph}^{\mathbf{UR}}R)^{-1}Z_R'\omega \tag{7}$$

This estimator imposes restrictions on the instruments in the GMM step, but recycles the unrestricted IVE residuals in the W estimator. Let,s compare the results. We did grid searches over the minimum Chi squared statistics for restricted models with up to 3-join-point models. The differences between the $\hat{\beta}_{Z_R}^{X_R}$ and the $\hat{\beta}_{GMM}^{R}$ estimates never reached more than 10^{-4} and the same goes for the Minimum Chi Squared statistics. The standard deviations differed with up to 0.1. The $\hat{\beta}_{Z_R}^{X_R}$ estimates were thus much closer to the GMM results than the β^{BL} estimator. Yet, on our computers it was faster than GMM by a factor of more than 30.

The BL estimator is still the speed champ in the overidentified cases, but we feel uncomfortable with the β^R estimates it yields and the Chi squared values that result from them. In table A1.6 below, we give an example of how the GMM estimation and BL-estimation differ in χ^2 values and in the ranking of sets of join-points for the flat-slop-flat model. Please note that, by doubling the number of observations included in the left-most Becker-region, GMM estimation strengthens the BL-conclusion, that there is evidence for men exhibiting caring

Table A1.6: Ranking of Join Points

GMM-estimation	Best	2nd best	3rd best	4th best
Join point 1	2	2	2	1
Join point 2	12	10	11	12
Chi Squared	150.31	150.38	150.41	150.44

BL-estimation	Best	2nd best	3rd best	4th best
Join point 1	1	2	2	2
Join point 2	12	10	12	11
Chi Squared	128.69	128.76	128.77	128.84

The closeness of the $\hat{\beta}_{Z_R}^{X_R}$ and the $\hat{\beta}_{GMM}^{R}$ estimates lends evidence to our point: that the defect of the β^{BL} estimator is, that it does not restrict the instrument variables or, equivalently, that it does not weigh the weighting matrix properly.

We propose using the 2SLS estimator, $\hat{\beta}^{R}_{2SLS} = (X'_R Z_R)^{-1} Z'_R \omega$ in the just identified case given by the present model. We propose this, because it is a closed form expression and because it is the fastest way to estimate β^R. We also propose to use the β^{XR}_{ZR}−estimator in the overidentified case, because it seems a fair compromise between speed and precision.

The β−estimates and preferred models, that follows from applying various estimation techniques discussed, are detailed over the next six pages.

Flat-slope-flat Parameter Estimates Food at home	BL results			BL Estimator			GMM estimator			**GMM - Preferred Model**		
	join 1	0.21		join 1	0.21		join 1	0.21		join 1	**0.25**	
	join 2	0.43		join 2	0.43		join 2	0.43		join 2	0.43	
	chi2	127.80		chi2	128.69		chi2	150.44		chi2	**150.31**	
	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value
Constant	37.94	27.26	1.39	38.35	27.27	1.41	19.73	31.20	0.63	20.29	31.18	0.65
ln (Total expenditure)	17.74	13.70	1.30	17.13	13.70	1.25	27.81	15.94	1.75	27.33	15.90	1.72
ln (Total expenditure) Squared	-4.25	2.00	-2.13	-4.17	2.00	-2.08	-5.67	2.32	-2.44	-5.60	2.32	-2.42
Age	1.43	0.34	4.18	1.43	0.34	4.17	1.66	0.35	4.68	1.66	0.35	4.69
Age squared	0.08	0.23	0.35	0.07	0.23	0.32	0.06	0.24	0.27	0.06	0.24	0.26
Number of young children	1.52	0.29	5.17	1.53	0.29	5.20	1.58	0.30	5.21	1.58	0.30	5.22
Number of medium children	2.56	0.17	15.42	2.56	0.17	15.47	2.51	0.17	14.80	2.52	0.17	14.85
Spouse age	0.38	0.38	1.02	0.38	0.38	1.02	0.20	0.39	0.51	0.19	0.39	0.50
Husband fracophone	0.34	0.49	0.69	0.34	0.49	0.70	0.48	0.51	0.94	0.47	0.51	0.94
Husband allophone	1.13	0.39	2.92	1.12	0.39	2.89	1.15	0.41	2.83	1.15	0.41	2.83
House owner	-0.78	0.34	-2.33	-0.79	0.34	-2.34	-0.82	0.35	-2.35	-0.81	0.35	-2.34
City residence	-0.03	0.28	-0.11	-0.04	0.28	-0.12	-0.07	0.29	-0.22	-0.07	0.29	-0.22
Education beyond High School	0.17	0.25	0.68	0.17	0.25	0.66	0.25	0.26	0.94	0.25	0.26	0.93
Region Atlantic	-1.90	1.10	-1.72	-1.92	1.10	-1.75	-1.91	1.13	-1.69	-1.90	1.13	-1.68
Region Quebec	0.19	0.65	0.29	0.19	0.65	0.30	0.19	0.68	0.28	0.20	0.68	0.29
Region Prairies	-1.17	0.67	-1.74	-1.17	0.67	-1.75	-1.27	0.69	-1.85	-1.27	0.69	-1.84
Region Brtish Columbia	-0.26	0.65	-0.41	-0.26	0.75	-0.34	-0.43	0.79	-0.55	-0.42	0.79	-0.54
ln (Price of food)	12.37	5.21	2.38	12.29	5.21	2.36	13.52	5.43	2.49	13.59	5.43	2.50
ln (Price of Household operations)	-0.23	7.68	-0.03	-0.25	7.68	-0.03	-0.94	7.87	-0.12	-0.97	7.88	-0.12
ln (Price of Tobacco)	-4.43	1.69	-2.62	-4.55	1.69	-2.69	-4.32	1.76	-2.45	-4.31	1.76	-2.45
ln (Price of Alcohol)	18.06	7.92	2.28	18.28	7.92	2.31	18.01	8.12	2.22	17.92	8.12	2.21
ln (Price of women's clothing)	-3.81	2.95	-1.29	-3.84	2.96	-1.30	-4.77	3.08	-1.55	-4.72	3.08	-1.53
ln (Price of Restarant)	-4.09	5.49	-0.74	-4.01	5.49	-0.73	-4.33	5.70	-0.76	-4.41	5.70	-0.77
ln (Price of Gas)	0.36	2.25	0.16	0.28	2.25	0.12	0.10	2.36	0.04	0.09	2.36	0.04
ln (Price of Care)	-3.79	2.16	-1.75	-3.86	2.16	-1.79	-4.42	2.24	-1.98	-4.41	2.24	-1.97
ln (Price of Transpotation)	-7.72	3.01	-2.57	-7.79	3.01	-2.59	-7.66	3.10	-2.47	-7.63	3.10	-2.46
ln (Price of Services)	3.28	3.64	0.90	3.26	3.64	0.90	3.90	3.78	1.03	3.91	3.78	1.03
ln (Price of Suppl)	-1.90	8.50	-0.22	-1.65	8.50	-0.19	-0.54	8.75	-0.06	-0.52	8.75	-0.06
ln (Price of Recreation)	8.10	6.83	1.19	8.12	6.82	1.19	6.91	7.07	0.98	6.96	7.07	0.98
ln (Price of Furniture)	-12.53	8.52	-1.47	-12.61	8.51	-1.48	-11.60	8.78	-1.32	-11.68	8.78	-1.33
ln (Price of Carp)	-9.57	6.93	-1.38	-9.43	6.93	-1.36	-9.63	7.22	-1.34	-9.57	7.22	-1.33
Rho	-3.89	1.71	-2.27	-3.87	1.71	-2.27	-3.81	1.79	-2.13	-3.91	1.99	-1.97

Flat-slope-flat Parameter Estimates Household Operations	BL results			BL Estimator			GMM estimator			**GMM - Preferred Model**		
	join 1	0.21		join 1	0.21		join 1	0.21		join 1	**0.25**	
	join 2	0.43		join 2	0.43		join 2	0.43		join 2	0.43	
	chi2	127.80		chi2	128.69		chi2	150.44		chi2	**150.31**	
	Coefficier	Std. Error	t-value	Coeffici	Std. Error	t-value	Coeffici	Std. Error	t-value	Coefficient	Std. Error	t-value
Constant	45.18	20.44	2.21	44.55	20.40	2.18	39.08	24.45	1.60	39.42	24.51	1.61
ln (Total expenditure)	-25.46	10.56	-2.41	-25.16	10.53	-2.39	-19.45	12.99	-1.50	-19.78	13.02	-1.52
ln (Total expenditure) Squared	3.32	1.55	2.14	3.28	1.54	2.12	2.46	1.90	1.30	2.51	1.90	1.32
Age	-0.19	0.27	-0.71	-0.18	0.27	-0.66	-0.12	0.28	-0.44	-0.12	0.28	-0.44
Age squared	-0.13	0.18	-0.72	-0.13	0.17	-0.72	-0.04	0.18	-0.20	-0.04	0.18	-0.20
Number of young children	0.44	0.21	2.10	0.44	0.21	2.09	0.52	0.22	2.40	0.53	0.22	2.41
Number of medium children	0.24	0.11	2.18	0.24	0.11	2.19	0.27	0.12	2.34	0.27	0.12	2.37
Spouse age	0.96	0.28	3.50	0.96	0.27	3.49	1.01	0.28	3.57	1.01	0.28	3.57
Husband fracophone	-0.98	0.36	-2.75	-0.98	0.36	-2.76	-1.04	0.37	-2.84	-1.04	0.37	-2.85
Husband allophone	-0.25	0.26	-0.95	-0.24	0.26	-0.95	-0.18	0.27	-0.66	-0.18	0.27	-0.66
House owner	2.61	0.25	10.53	2.61	0.25	10.51	2.56	0.26	9.71	2.56	0.26	9.71
City residence	-1.12	0.20	-5.51	-1.12	0.20	-5.53	-1.15	0.21	-5.43	-1.15	0.21	-5.43
Education beyond High School	-0.24	0.18	-1.30	-0.23	0.18	-1.27	-0.26	0.19	-1.39	-0.26	0.19	-1.39
Region Atlantic	0.09	0.76	0.12	0.96	0.76	1.26	1.02	0.80	1.28	1.03	0.80	1.29
Region Quebec	-0.25	0.46	-0.54	-0.23	0.46	-0.51	-0.15	0.48	-0.32	-0.14	0.48	-0.30
Region Prairies	-0.44	0.46	-0.96	-0.42	0.46	-0.91	-0.22	0.48	-0.46	-0.22	0.48	-0.45
Region Brtish Columbia	-1.09	0.53	-2.07	-1.06	0.53	-2.02	-1.15	0.55	-2.11	-1.15	0.55	-2.10
ln (Price of food)	-6.00	3.69	-1.62	-5.83	3.70	-1.58	-5.18	3.82	-1.35	-5.11	3.83	-1.34
ln (Price of Household operations)	8.86	5.21	1.70	8.87	5.21	1.70	10.95	5.49	1.99	10.90	5.49	1.99
ln (Price of Tobacco)	-1.29	1.18	-1.09	-1.34	1.18	-1.13	-1.37	1.23	-1.11	-1.35	1.23	-1.10
ln (Price of Alcohol)	2.93	5.59	0.52	2.99	5.59	0.53	5.87	5.83	1.01	5.77	5.84	0.99
ln (Price of women's clothing)	1.22	2.01	0.61	1.18	2.01	0.59	0.13	2.11	0.06	0.17	2.11	0.08
ln (Price of Restarant)	1.85	3.87	0.48	1.67	3.87	0.43	0.43	4.00	0.11	0.38	4.00	0.09
ln (Price of Gas)	-1.91	1.54	-1.24	-1.99	1.54	-1.29	-2.87	1.62	-1.77	-2.88	1.62	-1.78
ln (Price of Care)	-2.12	1.55	-1.37	-2.06	1.55	-1.33	-2.24	1.61	-1.39	-2.23	1.61	-1.39
ln (Price of Transpotation)	-1.75	2.08	-0.84	-1.80	2.08	-0.86	-1.44	2.16	-0.67	-1.41	2.16	-0.65
ln (Price of Services)	3.18	2.46	1.29	3.23	2.46	1.31	3.55	2.58	1.38	3.56	2.58	1.38
ln (Price of Suppl)	-4.74	5.95	-0.80	-4.93	5.94	-0.83	-3.57	6.24	-0.57	-3.56	6.24	-0.57
ln (Price of Recreation)	-6.38	4.89	-1.31	-6.08	4.88	-1.25	-8.07	5.16	-1.56	-8.03	5.16	-1.56
ln (Price of Furniture)	4.44	5.92	0.75	4.38	5.92	0.74	5.06	6.20	0.82	5.02	6.20	0.81
ln (Price of Carp)	4.95	5.05	0.98	5.02	5.05	0.99	1.08	5.24	0.21	1.14	5.24	0.22
Rho	-2.52	1.27	-1.98	-2.51	1.27	-1.97	-2.97	1.35	-2.20	-3.20	1.47	-2.18

Flat-slope-flat Parameter Estimates Women's Clothing	BL results			BL Estimator			GMM estimator			**GMM - Preferred Model**		
	join 1	0.21		join 1	0.21		join 1	0.21		join 1	**0.25**	
	join 2	0.43		join 2	0.43		join 2	0.43		join 2	0.43	
	chi2	127.80		chi2	128.69		chi2	150.44		chi2	**150.31**	
	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value
Constant	6.75	12.94	0.52	7.35	12.94	0.57	12.62	14.13	0.89	12.18	14.02	0.87
ln (Total expenditure)	-1.17	6.24	-0.19	-1.17	6.24	-0.19	-4.10	6.84	-0.60	-3.68	6.75	-0.55
ln (Total expenditure) Squared	0.70	0.93	0.75	0.70	0.93	0.76	1.13	1.02	1.11	1.07	1.01	1.06
Age	-0.89	0.18	-4.86	-0.09	0.18	-0.49	-0.06	0.19	-0.32	-0.06	0.19	-0.32
Age squared	-0.18	0.10	-1.75	-0.18	0.10	-1.73	-0.20	0.11	-1.86	-0.20	0.11	-1.85
Number of young children	-1.26	0.15	-8.37	-1.26	0.15	-8.39	-1.31	0.16	-8.36	-1.31	0.16	-8.37
Number of medium children	-0.78	0.09	-9.00	-0.79	0.09	-9.04	-0.79	0.09	-8.82	-0.80	0.09	-8.88
Spouse age	-0.03	0.29	-0.10	-0.02	0.23	-0.09	0.02	0.24	0.08	0.02	0.24	0.09
Husband fracophone	0.41	0.27	1.54	0.40	0.27	1.51	0.36	0.28	1.27	0.36	0.28	1.27
Husband allophone	0.57	0.21	2.68	0.58	0.22	2.61	0.55	0.23	2.38	0.56	0.23	2.39
House owner	-0.08	0.17	-0.43	-0.08	0.17	-0.46	-0.10	0.18	-0.56	-0.10	0.18	-0.57
City residence	-0.13	0.15	-0.82	-0.12	0.15	-0.81	-0.16	0.16	-0.97	-0.16	0.16	-0.97
Education beyond High School	-0.20	0.15	-1.37	-0.19	0.15	-1.30	-0.17	0.15	-1.10	-0.17	0.15	-1.09
Region Atlantic	-0.58	0.64	-0.89	-0.57	0.64	-0.88	-0.60	0.66	-0.91	-0.61	0.66	-0.93
Region Quebec	0.11	0.35	0.32	0.09	0.34	0.26	0.08	0.36	0.23	0.07	0.36	0.21
Region Prairies	-0.06	0.37	-0.17	-0.04	0.37	-0.11	-0.11	0.39	-0.28	-0.11	0.39	-0.29
Region Brtish Columbia	-0.79	0.43	-1.83	-0.77	0.43	-1.81	-0.88	0.45	-1.97	-0.88	0.45	-1.98
ln (Price of food)	-1.89	3.10	-0.61	-1.80	3.10	-0.58	-3.23	3.23	-1.00	-3.30	3.23	-1.02
ln (Price of Household operations)	-1.34	4.44	-0.30	1.52	4.43	0.34	2.04	4.59	0.44	2.09	4.59	0.45
ln (Price of Tobacco)	-1.50	0.93	-1.62	-1.48	0.92	-1.60	-1.26	0.96	-1.31	-1.27	0.96	-1.33
ln (Price of Alcohol)	5.75	4.38	1.31	5.85	4.38	1.34	5.49	4.55	1.21	5.60	4.55	1.23
ln (Price of women's clothing)	-1.54	1.69	-0.91	-1.49	1.69	-0.88	-1.62	1.76	-0.92	-1.67	1.76	-0.95
ln (Price of Restarant)	1.87	3.10	0.60	1.72	3.09	0.56	2.46	3.21	0.77	2.53	3.21	0.79
ln (Price of Gas)	-1.15	1.25	-0.92	-1.17	1.25	-0.94	-1.04	1.30	-0.80	-1.03	1.30	-0.79
ln (Price of Care)	-1.04	1.21	-0.86	-0.95	1.21	-0.79	-0.87	1.26	-0.69	-0.88	1.26	-0.70
ln (Price of Transpotation)	-0.33	1.69	-0.20	-0.32	1.69	-0.19	-0.56	1.73	-0.32	-0.60	1.73	-0.34
ln (Price of Services)	-1.83	2.08	-0.88	-1.84	2.08	-0.89	-1.49	2.14	-0.70	-1.51	2.14	-0.70
ln (Price of Suppl)	1.07	4.67	0.23	0.78	4.67	0.17	-1.72	4.88	-0.35	-1.74	4.88	-0.36
ln (Price of Recreation)	-1.76	3.92	-0.45	-1.65	3.91	-0.42	-3.22	4.08	-0.79	-3.26	4.08	-0.80
ln (Price of Furniture)	1.44	4.97	0.29	1.10	4.96	0.22	4.09	5.19	0.79	4.14	5.19	0.80
ln (Price of Carp)	-1.30	3.90	-0.33	-1.25	3.90	-0.32	-0.08	4.09	-0.02	-0.15	4.09	-0.04
Rho	3.31	0.89	3.74	3.27	0.88	3.70	3.64	0.94	3.86	3.89	1.06	3.68

Flat-slope-flat Parameter Estimates Men's Clothing	BL results			BL Estimator			GMM estimator			**GMM - Preferred Model**		
	join 1	0.21		join 1	0.21		join 1	0.21		join 1	**0.25**	
	join 2	0.43		join 2	0.43		join 2	0.43		join 2	0.43	
	chi2	127.80		chi2	128.69		chi2	150.44		chi2	**150.31**	
	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value
Constant	26.00	9.53	2.73	26.16	9.53	2.75	27.70	10.08	2.75	27.74	10.06	2.76
ln (Total expenditure)	-12.02	4.82	-2.49	-12.14	4.82	-2.52	-12.73	5.12	-2.49	-12.82	5.11	-2.51
ln (Total expenditure) Squared	2.09	0.73	2.88	2.11	0.73	2.90	2.18	0.77	2.82	2.20	0.77	2.85
Age	0.01	0.13	0.05	0.01	0.13	0.07	0.00	0.14	0.02	0.00	0.14	0.01
Age squared	-0.09	0.07	-1.22	-0.09	0.07	-1.21	-0.10	0.07	-1.41	-0.10	0.07	-1.40
Number of young children	-0.51	0.11	-4.54	-0.51	0.11	-4.56	-0.53	0.12	-4.46	-0.53	0.12	-4.45
Number of medium children	-0.33	0.06	-5.44	-0.33	0.06	-5.46	-0.33	0.06	-5.25	-0.33	0.06	-5.27
Spouse age	-0.11	0.15	-0.69	-0.11	0.15	-0.72	-0.13	0.16	-0.81	-0.13	0.16	-0.79
Husband fracophone	0.32	0.18	1.80	0.32	0.18	1.84	0.31	0.18	1.70	0.31	0.18	1.71
Husband allophone	0.31	0.15	2.08	0.32	0.15	2.11	0.32	0.16	2.02	0.32	0.16	2.01
House owner	0.01	0.13	0.10	0.01	0.13	0.09	0.02	0.14	0.14	0.02	0.14	0.13
City residence	-0.09	0.11	-0.85	-0.09	0.11	-0.81	-0.05	0.12	-0.47	-0.06	0.12	-0.47
Education beyond High School	-0.35	0.11	-3.33	-0.35	0.10	-3.31	-0.37	0.11	-3.34	-0.37	0.11	-3.35
Region Atlantic	0.13	0.47	0.27	0.12	0.47	0.25	0.27	0.49	0.56	0.28	0.49	0.58
Region Quebec	0.34	0.24	1.38	0.33	0.24	1.34	0.34	0.25	1.35	0.35	0.25	1.37
Region Prairies	0.05	0.26	0.20	0.05	0.26	0.18	0.09	0.28	0.33	0.09	0.28	0.32
Region Brtish Columbia	-0.30	0.29	-1.04	-0.31	0.29	-1.07	-0.20	0.30	-0.67	-0.20	0.30	-0.66
ln (Price of food)	1.46	2.02	0.72	1.42	2.02	0.70	2.20	2.12	1.04	2.24	2.12	1.06
ln (Price of Household operations)	-1.33	3.07	-0.43	-1.28	3.07	-0.42	-2.25	3.24	-0.69	-2.30	3.24	-0.71
ln (Price of Tobacco)	0.34	0.68	0.49	0.34	0.69	0.49	0.43	0.73	0.59	0.44	0.73	0.60
ln (Price of Alcohol)	1.75	3.14	0.56	1.75	3.14	0.56	0.96	3.30	0.29	0.88	3.31	0.26
ln (Price of women's clothing)	-1.85	1.26	-1.46	-1.85	1.26	-1.47	-2.20	1.31	-1.68	-2.19	1.31	-1.67
ln (Price of Restarant)	-0.42	2.25	-0.19	-0.35	2.25	-0.16	-0.70	2.33	-0.30	-0.72	2.33	-0.31
ln (Price of Gas)	-1.43	0.88	-1.61	-1.48	0.88	-1.67	-1.37	0.93	-1.48	-1.37	0.93	-1.48
ln (Price of Care)	-0.23	0.87	-0.27	-0.21	0.87	-0.24	-0.15	0.91	-0.16	-0.14	0.91	-0.15
ln (Price of Transpotation)	0.20	1.24	0.16	0.23	1.24	0.18	0.62	1.28	0.49	0.65	1.28	0.51
ln (Price of Services)	-0.05	1.43	-0.04	-0.12	1.43	-0.09	0.24	1.49	0.16	0.24	1.49	0.16
ln (Price of Suppl)	4.94	3.44	1.43	4.97	3.44	1.44	5.94	3.57	1.67	5.95	3.57	1.67
ln (Price of Recreation)	-6.47	2.71	-2.38	-6.54	2.72	-2.41	-6.12	2.83	-2.17	-6.11	2.83	-2.16
ln (Price of Furniture)	5.23	3.42	1.53	5.32	3.42	1.56	4.82	3.55	1.36	4.83	3.55	1.36
ln (Price of Carp)	-3.07	2.68	-1.15	-3.13	2.68	-1.17	-3.44	2.82	-1.22	-3.40	2.82	-1.21
Rho	-1.34	0.69	-1.94	-1.33	0.69	-1.94	-1.47	0.73	-2.00	-1.82	0.82	-2.21

Flat-slope-flat Parameter Estimates Kid's Clothing	BL results			BL Estimator			GMM estimator			**GMM - Preferred Model**		
	join 1	0.21		join 1	0.21		join 1	0.21		join 1	**0.25**	
	join 2	0.43		join 2	0.43		join 2	0.43		join 2	0.43	
	chi2	127.80		chi2	128.69		chi2	150.44		chi2	**150.31**	
	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value
Constant	-0.19	8.33	-0.02	0.15	8.32	0.02	1.60	9.63	0.17	1.48	9.58	0.15
ln (Total expenditure)	-0.17	4.10	-0.04	-0.09	4.09	-0.02	-0.72	4.88	-0.15	-0.59	4.83	-0.12
ln (Total expenditure) Squared	0.00	0.61	0.00	-0.01	0.61	-0.02	0.08	0.72	0.12	0.07	0.71	0.09
Age	0.21	0.12	1.65	0.20	0.12	1.60	0.19	0.13	1.43	0.19	0.13	1.43
Age squared	-0.19	0.08	-2.35	-0.19	0.08	-2.34	-0.20	0.09	-2.32	-0.20	0.09	-2.32
Number of young children	0.18	0.11	1.65	0.19	0.11	1.69	0.17	0.12	1.44	0.17	0.12	1.43
Number of medium children	1.04	0.07	15.48	1.04	0.07	15.44	1.02	0.07	14.40	1.02	0.07	14.39
Spouse age	-0.18	0.14	-1.30	-0.17	0.14	-1.23	-0.18	0.15	-1.20	-0.18	0.15	-1.20
Husband fracophone	0.12	0.16	0.74	0.12	0.16	0.74	0.10	0.16	0.62	0.10	0.16	0.62
Husband allophone	0.03	0.12	0.20	0.03	0.12	0.25	0.00	0.13	-0.01	0.00	0.13	0.00
House owner	-0.07	0.10	-0.70	-0.08	0.10	-0.75	-0.09	0.11	-0.86	-0.09	0.11	-0.86
City residence	-0.25	0.10	-2.57	-0.25	0.10	-2.57	-0.30	0.10	-2.88	-0.30	0.10	-2.88
Education beyond High School	0.04	0.09	0.51	0.05	0.09	0.60	0.00	0.09	-0.04	0.00	0.09	-0.04
Region Atlantic	0.57	0.35	1.61	0.58	0.35	1.64	0.54	0.37	1.46	0.53	0.37	1.45
Region Quebec	0.02	0.20	0.08	0.01	0.20	0.05	0.05	0.20	0.25	0.05	0.20	0.24
Region Prairies	0.19	0.22	0.85	0.21	0.22	0.92	0.19	0.24	0.79	0.19	0.24	0.79
Region Brtish Columbia	-0.34	0.24	-1.45	-0.34	0.24	-1.44	-0.33	0.25	-1.31	-0.33	0.25	-1.32
ln (Price of food)	0.47	1.71	0.27	0.52	1.71	0.31	-0.25	1.82	-0.14	-0.28	1.82	-0.15
ln (Price of Household operations)	-4.57	2.44	-1.87	-4.53	2.44	-1.86	-3.79	2.55	-1.49	-3.76	2.55	-1.48
ln (Price of Tobacco)	-0.86	0.57	-1.51	-0.87	0.57	-1.53	-0.56	0.60	-0.94	-0.57	0.60	-0.95
ln (Price of Alcohol)	-0.08	2.63	-0.03	0.08	2.63	0.03	-0.85	2.76	-0.31	-0.80	2.76	-0.29
ln (Price of women's clothing)	-2.78	1.02	-2.74	-2.86	1.01	-2.81	-2.54	1.06	-2.38	-2.55	1.06	-2.40
ln (Price of Restarant)	2.84	1.78	1.59	2.78	1.78	1.57	3.73	1.89	1.98	3.76	1.89	1.99
ln (Price of Gas)	-0.59	0.71	-0.82	-0.55	0.71	-0.78	-0.73	0.74	-0.99	-0.73	0.74	-0.98
ln (Price of Care)	0.58	0.73	0.79	0.57	0.73	0.79	0.63	0.76	0.84	0.63	0.76	0.83
ln (Price of Transpotation)	-0.09	0.97	-0.09	-0.07	0.96	-0.07	0.17	1.02	0.17	0.16	1.02	0.15
ln (Price of Services)	-0.84	1.13	-0.74	-0.78	1.13	-0.69	-1.47	1.21	-1.22	-1.48	1.21	-1.22
ln (Price of Suppl)	6.63	2.79	2.38	6.50	2.79	2.33	7.03	2.92	2.40	7.03	2.92	2.40
ln (Price of Recreation)	2.59	2.35	1.10	2.61	2.35	1.11	2.05	2.46	0.83	2.04	2.46	0.83
ln (Price of Furniture)	-2.01	2.79	-0.72	-2.11	2.79	-0.75	-1.83	2.93	-0.63	-1.82	2.93	-0.62
ln (Price of Carp)	-0.69	2.40	-0.29	-0.80	2.39	-0.34	-1.17	2.58	-0.45	-1.20	2.58	-0.46
Rho	1.54	0.59	2.64	1.53	0.58	2.62	1.30	0.62	2.11	1.45	0.68	2.13

Flat-slope-flat Parameter Estimates VICES	BL results			BL Estimator			GMM estimator			**GMM - Preferred Model**		
	join 1	0.21		join 1	0.21		join 1	0.21		join 1	**0.25**	
	join 2	0.43		join 2	0.43		join 2	0.43		join 2	0.43	
	chi2	127.80		chi2	128.69		chi2	150.44		chi2	**150.31**	
	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value	Coefficient	Std. Error	t-value
Constant	7.11	23.72	0.30	3.29	23.66	0.14	12.70	25.27	0.50	12.64	25.3016	0.50
ln (Total expenditure)	5.16	12.25	0.42	6.85	12.21	0.56	-0.45	13.19	-0.03	-0.47	13.1949	-0.04
ln (Total expenditure) Squared	-0.84	1.79	-0.47	-1.09	1.79	-0.61	-0.12	1.92	-0.06	-0.12	1.925223	-0.06
Age	-0.13	0.30	-0.45	-0.12	0.30	-0.41	-0.11	0.30	-0.37	-0.12	0.30447	-0.38
Age squared	-0.01	0.18	-0.07	-0.02	0.18	-0.12	-0.08	0.18	-0.42	-0.08	0.184428	-0.42
Number of young children	-0.76	0.23	-3.34	-0.78	0.23	-3.39	-0.87	0.24	-3.67	-0.86	0.23611	-3.66
Number of medium children	-0.24	0.13	-1.86	-0.24	0.13	-1.88	-0.21	0.13	-1.60	-0.22	0.133615	-1.62
Spouse age	-0.33	0.35	-0.96	-0.35	0.35	-1.02	-0.40	0.36	-1.13	-0.40	0.357461	-1.11
Husband fracophone	0.24	0.35	0.70	0.24	0.35	0.70	0.19	0.36	0.53	0.19	0.35842	0.53
Husband allophone	-1.35	0.29	-4.74	-1.35	0.28	-4.74	-1.45	0.29	-4.91	-1.45	0.294344	-4.91
House owner	-1.42	0.28	-5.17	-1.44	0.27	-5.25	-1.38	0.28	-4.85	-1.38	0.284838	-4.86
City residence	0.43	0.22	1.97	0.44	0.22	2.00	0.54	0.23	2.35	0.54	0.229611	2.35
Education beyond High School	1.22	0.21	5.91	1.22	0.21	5.88	1.15	0.22	5.34	1.15	0.215825	5.33
Region Atlantic	-0.68	0.89	-0.77	-0.65	0.89	-0.73	-0.93	0.93	-1.01	-0.92	0.926455	-0.99
Region Quebec	-0.21	0.49	-0.43	-0.19	0.48	-0.40	-0.37	0.50	-0.73	-0.36	0.503834	-0.72
Region Prairies	-0.44	0.53	-0.83	-0.45	0.53	-0.84	-0.67	0.55	-1.22	-0.68	0.550723	-1.23
Region Brtish Columbia	-0.38	0.58	-0.65	-0.37	0.58	-0.64	-0.61	0.60	-1.02	-0.61	0.601506	-1.02
ln (Price of food)	0.83	4.15	0.20	0.79	4.15	0.19	-0.22	4.34	-0.05	-0.17	4.347676	-0.04
ln (Price of Household operations)	0.94	5.77	0.16	0.48	5.76	0.08	1.31	6.06	0.22	1.23	6.06389	0.20
ln (Price of Tobacco)	-0.67	1.27	-0.53	-0.68	1.27	-0.54	-1.10	1.32	-0.83	-1.08	1.318533	-0.82
ln (Price of Alcohol)	7.87	6.40	1.23	7.48	6.40	1.17	7.57	6.69	1.13	7.47	6.694724	1.12
ln (Price of women's clothing)	-4.69	2.39	-1.96	-4.62	2.39	-1.93	-3.37	2.46	-1.37	-3.37	2.459353	-1.37
ln (Price of Restarant)	-3.23	4.38	-0.74	-3.07	4.38	-0.70	-3.63	4.56	-0.79	-3.64	4.564308	-0.80
ln (Price of Gas)	1.15	1.61	0.72	1.16	1.61	0.72	1.15	1.66	0.69	1.15	1.659794	0.70
ln (Price of Care)	-2.40	1.75	-1.37	-2.31	1.75	-1.32	-2.43	1.81	-1.34	-2.42	1.811398	-1.34
ln (Price of Transpotation)	-1.82	2.38	-0.76	-1.77	2.38	-0.74	-2.81	2.50	-1.13	-2.77	2.495185	-1.11
ln (Price of Services)	2.75	2.68	1.03	2.74	2.68	1.02	2.87	2.77	1.04	2.88	2.773084	1.04
ln (Price of Suppl)	-4.30	6.77	-0.64	-3.79	6.77	-0.56	-5.17	7.09	-0.73	-5.17	7.08955	-0.73
ln (Price of Recreation)	-1.29	5.57	-0.23	-1.29	5.57	-0.23	0.70	5.78	0.12	0.70	5.7788	0.12
ln (Price of Furniture)	7.37	7.02	1.05	7.38	7.02	1.05	6.01	7.25	0.83	6.05	7.24912	0.83
ln (Price of Carp)	-4.34	5.66	-0.77	-4.11	5.66	-0.73	-1.56	5.87	-0.27	-1.51	5.87025	-0.26
Rho	-1.42	1.39	-1.02	-1.46	1.38	-1.05	-1.36	1.44	-0.95	-1.87	1.58476	-1.18

5.4 Grid searching: Ranking models with $\hat{\beta}^{BL}$ and $\hat{\beta}^{R}_{GMM}$

How important are the differences in estimators? That depends on the purpose of testing. Between models the $\hat{\beta}^{R}_{BL}$ ranks models just the same as the other estimators, e.g. flat-slope-flat cannot be reduced to slope-flat when testing in a 20-spline framework. But if the interest is with the results for a specific model, $\hat{\beta}^{R}_{BL}$ does a sub-optimal job. Using $\hat{\beta}^{R}_{BL}$ and grid searching flat-slope-flat we find the first flat part of ρ only to reach 0.21 (or 5% of the sample). Using $\hat{\beta}^{R}_{GMM}$ leads us to prefer a join point at 0.25 (or 10% of the sample). As already mentioned this strengthens the BL conclusion that men exhibit caring.

5.4.1 Grid search technique and results.

Here we want to look closer at the results. For every restriction we've programmed an algorithm that runs through all the different join point combinations possible. For each possible set of join points it calculates $\left(\hat{\beta}^{UR} - \tilde{\beta}^{R}\right)' C^{-1}_{\beta^{UR}} \left(\hat{\beta}^{UR} - \tilde{\beta}^{R}\right)$ and stores this value along with $\hat{\beta}^{R}$ and the standard deviation for each element in $\hat{\beta}^{R}$. Finally it ranks all the χ^2 statistics to determine which set of join points gives the best fit with the unrestricted model. Doing this for different restrictions we can produce a list of chi squared values for different kinds of models. See below

Table A1.7: Minimum Chi Squared test results

Model	Number of ρ parameters	GMM-Criterion	Criterion with BL-estimator	BL published criterion
Unitary	0	199.5	167.9	167.4
Collective (no caring)	6	166.7	138.6	138.0
Flat-slope	7	166.8	138.5	138.0
Slope-flat	7	156.1	132.6	131.8
Slope-slope	13	143.4	124.4	123.8
Flat-slope-flat	8	150.3	128.7	127.8
Slope-flat-slope	14	143.0	123.6	123.2
Slope-slope-slope	20	136.9	116.7	116.1
fl-sl-fl-sl	15	137.5	119.8	119.4
sl-fl-sl-fl	15	145.8	123.7	123.2
fl-sl-fl-sl-fl	16	139.8	118.8	NA
sl-fl-sl-fl-sl	22	134.6	115.6	NA
sl-sl-sl-sl-sl	34	114.8	97.4	NA

Now we can compare models and calculate the probability that a specific less-restricted model explains significantly more of the variation in the data[12] than does a more-restricted model.

[12]as represented by the unrestricted fit

The difference between the Minimum Chi Squared criteria is itself χ^2 distributed with degrees of freedom equal to the difference in the degrees of freedom of two statistics. Comparing flat-slope-flat and slope-flat we would thus use a χ^2 table and see if 166.8-156,1=10.7 is significant in a $\chi^2\,(8-7)$ distribution.

The number of ρ parameters reported above stem from taking the degrees of freedom added by the estimation of M*s (6*20=120 in the present case) spline coefficients and subtracting the restrictions imposed. The number of restrictions can be expressed thus: M*(s-the number of sloping segments)-the number of restricted model join points. For slope-flat-slope this is: 6*(20-2)-2=106. Altogether we get: Number of ρ parameters = M* (number of sloping segments) + number of restricted model join points. That is 6*2+2=14 for the slope-flat-slope case.

The models we compare should ofcourse be theoretically meaningful. We start with the list of χ^2 statistics given above and the probability that less restricted models explain significantly more variation than more restricted models The results are presented in table A1.8 below. As they rank the models just like the BL-results we just refer to the BL paper for an analysis of the results. We document results using both BL estimation and GMM estimation.

Unrestriced Model

Results using GMM Estimation (Restricted model)	flat	slope	Flat-slope	Slope-flat	**Flat-slope -flat**	Slope -slope	Slope-flat -slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope -slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
slope				0.11	**0.03**	0.15	0.25	0.06	1.29	0.26	0.81	0.96	0.39
Flat-slope					**0.00**	0.07	0.12	0.03	0.72	0.14	0.50	0.61	0.27
Slope-flat					**1.65**	4.81	7.03	1.77	24.76	6.10	11.96	12.25	3.88
Flat-slope-flat						**22.44**	**29.18**	**7.80**	**71.99**	**22.89**	**34.29**	**33.07**	**10.09**
Slope-slope							54.03	5.46		30.87	49.19	45.87	12.48
Slope-flat-slope								1.97		20.00	41.84	39.58	10.49
fl-sl-fl-sl											98.79	88.91	24.85
sl-fl-sl-fl										1.40	11.47	12.95	4.03
fl-sl-fl-sl-fl											58.78	52.14	12.57
Slope-slope-slope												30.83	7.56
sl-fl-sl-fl-sl													7.10
sl-sl-sl-sl-sl													

Results using BL Estimation (Restricted model)	flat	slope	Flat-slope	Slope-flat	**Flat-slope -flat**	Slope -slope	Slope-flat -slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope -slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
flat		0.01	0.01	0.00	**0.00**	0.00	0.01	0.00	0.01	0.00	0.01	0.03	0.02
slope			75.90	1.42	**0.70**	4.77	5.94	2.72	9.36	3.06	8.00	11.40	5.12
Flat-slope					**0.17**	2.84	3.74	1.67	6.29	1.94	5.79	8.64	4.00
Slope-flat					**4.79**	22.45	25.44	12.03	35.15	12.80	25.36	32.03	13.36
Flat-slope-flat						**51.01**	**53.57**	**26.36**	**66.25**	**27.00**	**44.52**	**52.15**	**21.74**
Slope-slope							37.57	10.17	70.44	12.98	35.72	45.73	17.02
Slope-flat-slope								5.16		8.77	32.61	43.34	15.82
fl-sl-fl-sl										29.88	67.60	75.49	26.26
sl-fl-sl-fl										2.61	21.86	32.52	12.16
fl-sl-fl-sl-fl											72.17	79.19	26.13
Slope-slope-slope												58.94	15.41
sl-fl-sl-fl-sl													10.88
sl-sl-sl-sl-sl													

Table   shows the probability (in %) that a more restricted model explains as much variation as the less restricted model

6 Appendix 2: Semiparametric Estimation in a Model with Endogenous Regressors

6.1 Semiparametric Estimation

In Browning and Lechene (2001), the unrestricted model has the following partial linear structure

$$\omega_j = \alpha_j + X\beta_j + g_j(\rho) + \varepsilon_j \qquad j = 1, ..., M = 6 \tag{8}$$

In order to estimate $g_j(\rho)$ we take our point of departure in the semi-parametric procedure suggested by Robinson (1988), where a simple transformation can be used to give an estimator of β_j and $g_j(\rho)$. In line with Robinson, we take expectations conditional on ρ in the unrestricted model (8). Since ρ is assumed exogenous, in the sense that $E[\varepsilon_j|\rho] = 0$, the resulting expression reduces to

$$g_j(\rho) = \alpha_j + E(\omega_j|\rho) - E(X|\rho)\beta_j \tag{9}$$

In order to estimate $g_j(\rho)$ we replace $E(\omega_j|\rho)$ and $E(X|\rho)$ by their nonparametric estimators[13] $\hat{m}_h^{\omega_j}(\rho)$ and $\hat{m}_h^X(\rho)$ and replace β_j by a consistent estimate of β_j. To achieve an estimator for β_j, we transform the model in deviations from ρ-conditional means, which effectively cancels out the constant α_j and $g_j(\rho)$, since ρ is assumed exogenous

$$\omega_j - E(\omega_j|\rho) = [X - E(X|\rho)]\beta_j + \varepsilon_j$$

Since total expenditure is endogenous, we have to correct for this, when β_j is estimated.

6.1.1 Estimation of β_j

To adjust for endogeneity, one could adapt the popular residual augmented regression technique (see for example Blundell, Duncan and Pendakur (1998)), which effectively is IV-estimation on transformed variables. In order to obtain more efficiency and to make our results comparable to BL2001, we implement Robinson's semiparametric approach in the GMM-framework. The estimation method in the present paper is slightly different in two ways. First, the endogenous variables enter in the linear part of the model, in contrast the endogenous variables enter in the non-linear part in the studies by Blundell, Duncan and Pendakur (1998). Secondly, we allow for a more complex covariance structure, by implementing the estimation of β_j into a GMM-framework.

[13]The estimators are described in the subsequent section "Nonparametric Estimation"

Now consider the unrestricted model in (8) and suppose a part of X is endogenous, in the sense that

$$E[\varepsilon_j|X] \neq 0$$

To obtain consistent estimates of β_j and to allow for a richer covariance structure, we estimate β by GMM. The moment condition are

$$E\left(Z'\varepsilon_j\right) = 0 \tag{10}$$

Moreover, there must be some correlation between the instruments and the endogenous variables, i.e.

$$E\left(Z'X\right) \neq 0$$

where Z is a matrix of instruments and $\varepsilon_j = \omega_j - \alpha_j - X\beta_j - g_j(\rho)$ is the error term of the model. Since ε_j includes the unknown nonlinear function $g_j(\rho)$, we need to transform the model in order to identify β_j. Following Robinson (1988) we transform the variables into deviations from $\rho-$conditional means. Define the transformed variables as

$$\begin{aligned}
\tilde{\omega} &= \omega - E(\omega|\rho) \\
\tilde{X} &= X - E(X|\rho) \\
\tilde{Z} &= Z - E(Z|\rho) \\
\tilde{\varepsilon}_j &= \varepsilon_j - E(\varepsilon_j|\rho)
\end{aligned}$$

Inserting the theoretical counterpart of ε_j, the transformed errorterm can be expressed as

$$\tilde{\varepsilon}_j = \underbrace{\omega_j - E(\omega_j|\rho)}_{\tilde{\omega}_j} - \underbrace{(X - E(X|\rho))}_{\tilde{X}}\beta_j$$

and the intercept α_j cancels out together with the unknown nonlinear part of the model, $g_j(\rho)$.

What remains to clarify, is that the orthogonality conditions in (10), are equivalent to $E\left[\tilde{Z}'\tilde{\varepsilon}_j\right] = 0$ and $E\left[\tilde{Z}'\tilde{X}_j\right] \neq 0$. If this holds, we can estimate the model characterized by (10), using GMM on the transformed variables. In order to se this, rewrite the orthogonality condition as

$$E\left(Z'\varepsilon_j\right) = E\left[\left(\tilde{Z} + E(Z|\rho)\right)'(\tilde{\varepsilon}_j + E(\varepsilon_j|\rho))\right] = 0$$

Since ρ is assumed exogenous, $E(\varepsilon_j|\rho) = 0$, this reduces to

$$E\left[\tilde{Z}'\tilde{\varepsilon}_j\right] + E\left[E(Z|\rho)'\,\tilde{\varepsilon}_j\right] = 0$$

Again we use the exogeneity of ρ plus the fact that Z is assumed to be exogenous

$$\begin{aligned} E\left[\tilde{Z}'\tilde{\varepsilon}_j\right] + E\left(Z|\rho\right)' E\left[\tilde{\varepsilon}_j\right] &= \\ E\left[\tilde{Z}'\tilde{\varepsilon}_j\right] &= 0, \quad j = 1, ..., M \end{aligned} \tag{11}$$

It is seen, that moment conditions in (11) and (10) are equivalent. Now insert the theoretical counterpart of $\tilde{\varepsilon}_j$ in (10)

$$E\left[\tilde{Z}'\left(\underbrace{\omega_j - E(\omega_j|\rho)}_{\tilde{\omega}_j} - \underbrace{(X - E(X|\rho))}_{\tilde{X}}\beta_j\right)\right] = 0, \quad j = 1, ..., M$$

That is β_j can be estimated simply by applying the GMM-estimator on the model reformulated in transformed variables $\tilde{\omega}_j$, $\tilde{X}$ and $\tilde{Z}$, where $E(\omega_j|\rho)$, $E(X|\rho)$ and $E(Z|\rho)$ are replaced by their nonparametric estimators $\hat{m}_h^{\omega_j}(\rho)$, $\hat{m}_h^X(\rho)$ and $\hat{m}_h^Z(\rho)$. The GMM-estimator of $\beta = (\beta_1, .., \beta_M)'$ can now be expressed as

$$\hat{\beta}^{GMM} = \left[\tilde{Z}'\tilde{X}\left(\widehat{\tilde{Z}'\Omega\tilde{Z}}\right)^{-1}\tilde{X}'\tilde{Z}\right]^{-1}\tilde{Z}'\tilde{X}\left(\widehat{\tilde{Z}'\Omega\tilde{Z}}\right)^{-1}\tilde{X}'\tilde{\omega}$$

where the weighting matrix $\widehat{\tilde{Z}'\Omega\tilde{Z}}$ has the same structure as in the linear parametric model given in appendix 1.

Robinson (1988) established the powerful result, that even though the nonparametric estimators that enters the regression are converging to the asymptotic distribution at a rate lower than $\sqrt{n}$, under some regularity conditions β is $\sqrt{n}$ consistent. The Robinson paper primarily focuses on the simplest settings, such as seemingly unrelated regressions. Fortunately, the model can be transformed, such that the results in Robinson (1988) still applies and thus β^{GMM} is $\sqrt{n}$ consistent too.

6.1.2 Estimation and Identification of $g_j(\rho)$

Given the nonparametric estimators of the conditional means and the GMM estimator $\hat{\beta}_j^{GMM}$, the estimator of $g_j(\rho)$ can now be expressed as

$$\hat{g}_j(\rho) = \hat{m}_h^{\omega_j}(\rho) - \alpha_j - \hat{m}_h^X(\rho)\,\hat{\beta}_j^{GMM} \tag{12}$$

One drawback of the semiparametric estimation is that $g_j(\rho)$ is only identified up to the constant, α_j. The constant simply cancels out, when the model is expressed in the deviation from conditional means. Hence, if the constant is ignored when $g_j(\rho)$ is estimated semiparametrically, we effectively estimate $g_j(\rho)+\alpha$, such that we have to chose between calculating $g_j(\rho)$ given the constant or calculating the constant given $E[g_j(\rho)]$. Therefor we must impose more structure on the model, in order to identify $g_j(\rho)$.

Fortunately, in this context, the level of $g_j(\rho)$ is irrelevant for the purpose of discriminating between different types of models, since the restrictions are imposed on the derivatives of $g_j(\rho)$. Hence, without any loss of generality, we can normalize $g_j(\rho)$ by assuming that it integrates to zero[14].

$$E(g_j(\rho)) = 0$$

With the assumption above we can now identify $g_j(\rho)$ by taking deviations from the mean

$$g_j(\rho) + \alpha_j - E[g_j(\rho) + \alpha_j] = g_j(\rho) - E[g_j(\rho)] = g_j(\rho)$$

Thus, $g_j(\rho)$ can now be identified as

$$\hat{g}_j(\rho_i)^{Normalized} = \hat{g}_j(\rho_i) - \frac{1}{n}\sum_{i=1}^{n} \hat{g}_j(\rho_i) \ , \quad i = 1, .., n$$

This normalization is important for at least to purposes. First of all, when we calculate the confidence intervals, we are not interested in the variance of the constant. Since the variance of $g_j(\rho)$ is computed on the basis of bootstrap samples, we effectively overestimate the variance of the curve, if we do not exclude the constant. Secondly, if the semiparametric estimates of $g_j(\rho)$ is used as an alternative against which to test a linear parametric null (e.g. flat-slope-flat), again the normalization of $g_j(\rho)$ is crucial.

We can now summarize the estimation procedure in the three following steps

1. Estimate $E(\omega_j|\rho)$, $E(X|\rho)$ and $E(Z|\rho)$ nonparametrically and obtain $\hat{m}_h^{\omega_j}(\rho)$, $\hat{m}_h^X(\rho)$ and $\hat{m}_h^Z(\rho)$

2. Estimate β_j by GMM on the transformed variables $\tilde{\omega}_j = \omega_j - \hat{m}_h^{\omega_j}(\rho)$, $\tilde{X} = X - \hat{m}_h^X(\rho)$ and $\tilde{Z} = Z - \hat{m}_h^Z(\rho)$

3. Calculate $\hat{g}_j(\rho)$ as in (12) (ignoring the constant), and normalize by subtracting the mean.

6.2 Non-parametric estimation of conditional means

6.2.1 The kernel method

For the estimation of the conditional means $E(\omega_j|\rho)$, $E(X|\rho)$ and $E(Z|\rho)$ we apply the Nadaraya-Watson estimator with Gaussian kernels. For a univariate regression of y on x the Nadaraya-

[14]When the semiparametric estimate of $g_j(\rho)$ is compared to $g_j(\rho)$ from the spline regression, we normalize the latter estimate as well.

Watson estimator is expressed as

$$\hat{m}_h^y(x) = \frac{\frac{1}{n}\sum_{i=1}^n y_i K_h(x_i - x)}{\frac{1}{n}\sum_{i=1}^n K_h(x_i - x)}$$

where $K_h(z) = h^{-1}\frac{1}{\sqrt{2\pi}}\exp\left[-(z/h)^2\right]$ is the Gaussian kernel function with bandwidth h. Under standard regularity conditions, this estimator is consistent and asymptotic normal, with a rate of converge lower than $\sqrt{n}$ (see Härdle (1990)). Formally we have

$$\sqrt{nh}\left[\frac{\hat{m}_h^y(x) - m^y(x) - h^2 B(x)}{\sqrt{V(x)}}\right] \xrightarrow{D} N(0,1) \tag{13}$$

where $V(x) = \hat{\sigma}(x) c_K / f(x)$ is the variance and $B(x) = d_K\{m_h^{\prime\prime y}(x) + 2m_h^{\prime y}(x) f'(x)/f(x)\}$ is the nondisappearing bias, with the kernel constants $c_K = \int K_h(z)\,dz$ and $d_K = \int z^2 K_h(z)\,dz$.

6.2.2 Bandwidth selection

As indicated in (13) the choice of bandwidth is crucial for the rate of convergence to the true regression curve $m^y(x)$ and thereby for the interpretation of the results. Therefore the bandwidth must be chosen to balance the variance versus the squared bias, since the variance of the nonparametric estimator is decreasing in the bandwidth while the Bias is increasing. The presence of the bias suggest that the precision of the curve should be measured in terms of the mean squared error (MSE)

$$MSE(h) = \left[h^2 B(x)\right]^2 + V(x)/(nh)$$

The optimal bandwidth is chosen to minimize $MSE(h)$, the first order conditions are

$$\frac{\partial MSE}{\partial h} = \left[4B(x)^2 h^3 - \frac{V(x)}{h^2 n}\right] = 0$$

with solution

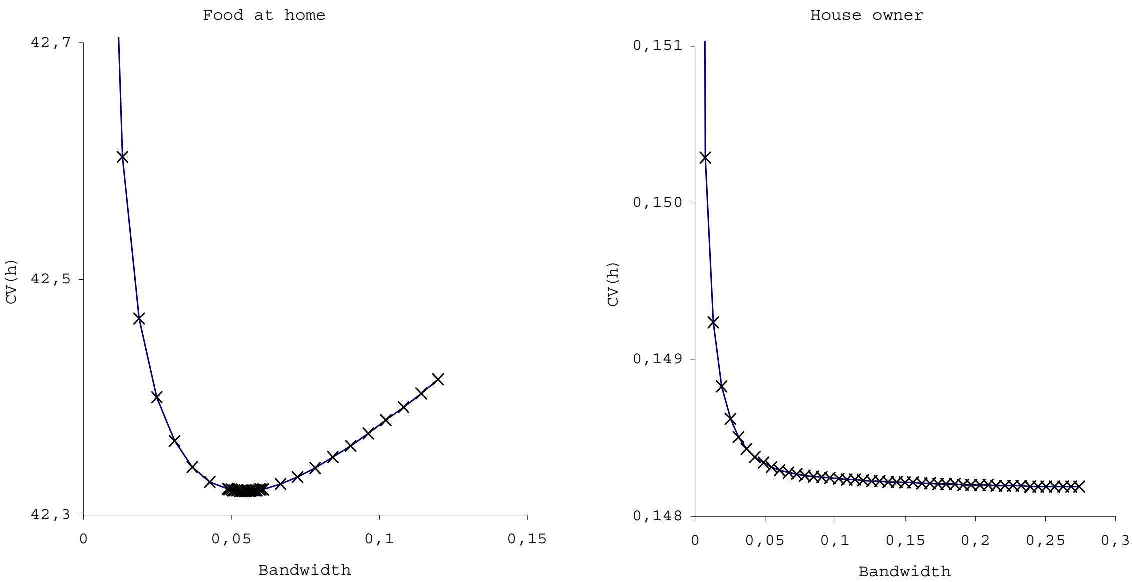
$$h^* = n^{-\frac{1}{5}}\sqrt[5]{\frac{V(x)}{4B(x)^2}} \text{ if } B(x)^2 \neq 0$$

The distance measure $MSE(h)$, can be approximated by a cross-validation function

$$CV(h) = \frac{1}{n}\sum_{j=1}^n w(x_i)\left[y_i - \hat{m}_{h,j}^y(x_i)\right]^2$$

where $w(x)$ is some trimming function and $\hat{m}_{h,j}^y(x_i)$ is a nonparametric estimate of $E[y|x]$ leaving out the i'th observation. The optimal bandwidth was chosen by minimizing the cross-validation function, performing a grid search over h. In most cases the algorithm actual found a minimum. For some of the explanatory variables however, the cross validation function was monotone decreasing in bandwidth with a horizontal asymptote (see figure A2.1, right panel).

Figure A2.1: Cross-validation functions for "Food at home" and "House owner"



In these cases we take it as an indication of the fact that the dependent variable and ρ are uncorrelated I.e. $E(y|\rho) = E(y)$ which gives $m_h^y(\rho) = E(y)$. Hence there will no optimal h, or put differently h must be infinitely large. The practical implication is that when h goes to infinity the kernel output approaches the same value (zero) for every observation. $\lim_{h\to\infty} K_h(z) = 0$ Therefor the Nadaraya-Watson estimator will be calculated as the mean of the dependent variable i.e.

$$\lim_{h\to\infty} \hat{m}_h^y(x) = \frac{K_{h\to\infty}(z)\,\frac{1}{n}\sum_{i=1}^{n} y_i}{\frac{1}{n} n K_{h\to\infty}(z)} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

For most variables an optimal bandwidth h^* were found. However, using a global bandwidth may be a bad approximation for several reasons . Different types of variation in the data or function can simply make a global h^* too restrictive even though it is optimal. To accommodate this problem it is attractive to use an adaptive bandwidth, where each point on the curve has its own bandwidth. When selecting the adaptation one has to decide on the basis of the assumed variation.

- If the observations and therefore the design points are sparse in some areas then the adaptive bandwidth has to accommodate this by expanding the bandwidth in these areas and vice versa.

- If there is heteroscedasticity then there must be more smoothing i.e. wider bandwidth in areas with high variance.
- On the other hand the variation can be in the true curve i.e. some parts are more flat than others. In this situation the influence of the bias will vary. Therefor the bandwidth will have to be very low in the wiggling areas.

Since ρ is approximately normally distributed, rather uniform, the most serious problem found in data is the first of the three. Blundell and Duncan (1998) and Blundell, Browning and Crawford (2003) face an equivalent problem with a varying density in the explanatory variable. They choose the weight λ_i.

$$\lambda_i = \left[\frac{\hat{f}_{h^*}(x_i)}{\prod_{i=1}^{n} \hat{f}_{h^*}(x_i)^{1/n}} \right]^{-\xi}, \ \xi \in [0,1]$$

If the density in a design point is lower than the geometric average of the density function (i.e. the data is sparse) then $\lambda_i > 1$ and the adaptive bandwidth $h_i^A = h^* \lambda_i$ is larger than the global bandwidth h^*. In words, this means that in sections with sparse data the nonparametric fit entails a higher degree of smoothness, in order to correct for the larger magnitude of the variance relative to the bias.

The parameter ξ controls the degree of adaption, such that the bandwidth is made proportionally wider in areas with sparse data, when ξ is increasing. In Jennen-Stenmetz and Gasser (1988) the question of how to choose the weight parameter ξ, is examined. As pointed in this paper, obviously, if $\xi = 0$, the adaptive bandwidth is equal to the global bandwidth and at the other extreme, $\xi = 1$ the kernel estimator approximates the $k - NN$ estimator. In principle, ξ and h^* should be chosen as to minimize the MSE. In practice however, this will be highly computer intensive. Jennen-Stenmetz and Gasser (1988) suggest that a good compromise would be a value of ξ between 0.2 and 0.5. As one can se in figure A2.2 (panel a to c) a high value of ξ makes the ends flatter and a low value of ξ makes the middle part flatter. We chose to estimate not only the curves but also the confidence interval with ξ set to 0.2, 0.25 and 0.5. It turns out that the middle part is still jumpy with low ξ, but if $\xi = 0.5$ the ends flattens out. We interpret this as $\xi = 0.5$ being the most preferable. The line with $\xi = 0.5$ is the thick line in the middle. We will however still make testing with all three different values of ξ.

Figure A2.2 (panel a): Semiparametric fit of $g(\rho)$ for various values of ξ

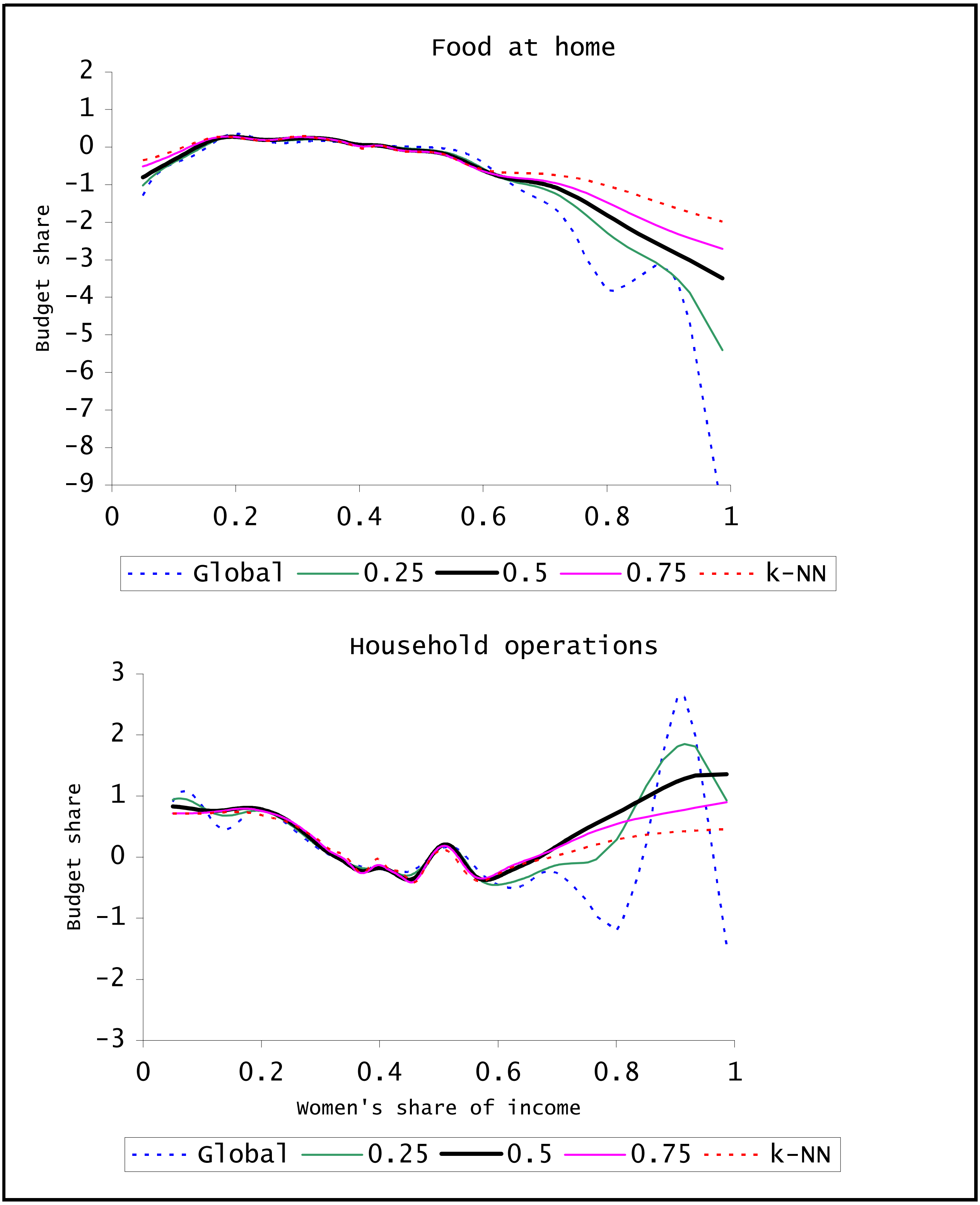


Figure A2.2 (panel b): Semiparametric fit of $g(\rho)$ for various values of ξ

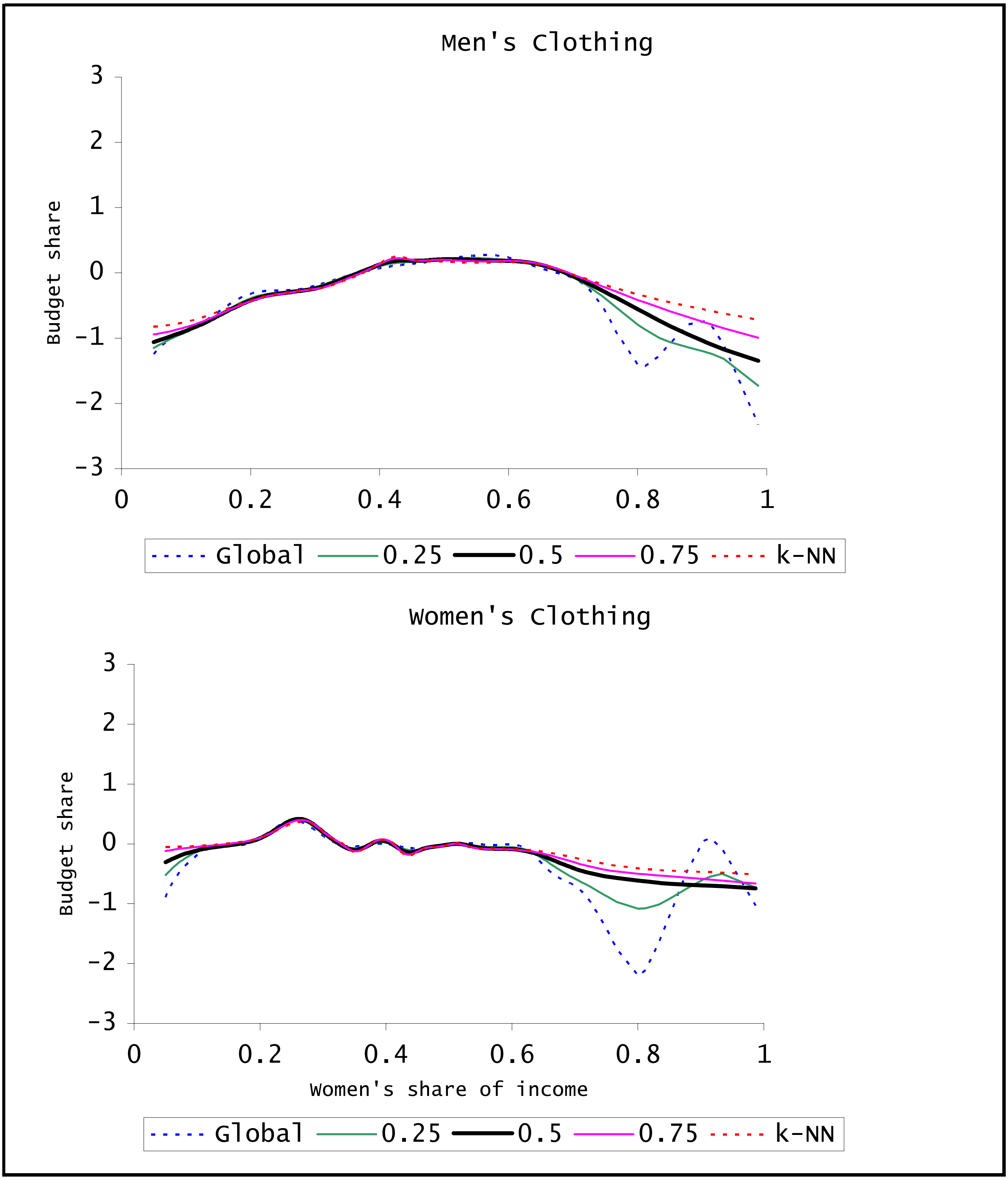
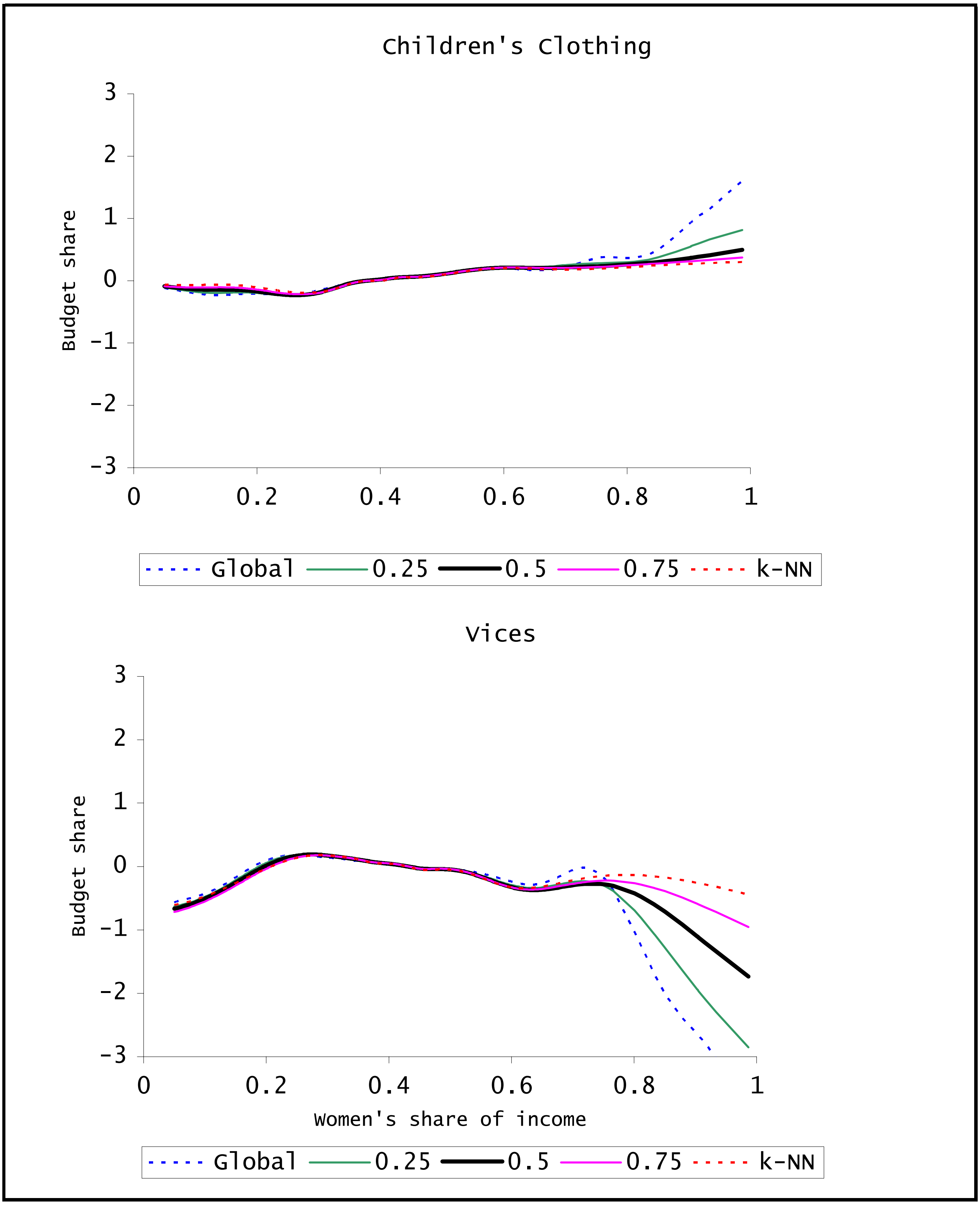


Figure A2.2 (panel c): Semiparametric fit of $g(\rho)$ for various values of ξ



6.3 Pointwise confidence intervals

6.3.1 Asymptotic confidence intervals

In order to do some inference, we need a measure for the precision of the semiparametric as well as for the nonparametric regression curves. In the nonparametric case the asymptotic properties is well founded (see Härdle (1990)). From the limiting distribution of the Nadaraya-Watson estimator confidence bands can easily be derived. If we insert the estimator of the variance, the confidence bands can be estimated as

$$C_{low}, C_{up} = \hat{m}_j(\rho) \pm c_\alpha \sqrt{c_K} \cdot \hat{\sigma}(x) / \sqrt{nh\hat{f}_h(x)}$$

where $\hat{\sigma}^2(x) = \frac{1}{n}\sum_{i=1}^{n} W_{hi}(x)\left(m_j(\rho_i) - \hat{m}_j(x)\right)^2$, $W_{hi} = K_h(x-\rho_i)/\hat{f}_h(x)$ and where c_α denotes the (100-$\alpha/2$) fractile in the normal distribution.

Now turn to the semiparametric case. The fact that β is estimated does not affect the limiting distribution of $\hat{g}_j(\rho)$, really simplifies the problem of deriving the variance. The reason is that $\hat{\beta}_j^{GMM}$ is $\sqrt{n}$-consistent, while $\hat{m}_h^{\omega_j}(\rho)$, $\hat{m}_h^X(\rho)$ and $\hat{m}_h^Z(\rho)$ are converging at a rate slower that $\sqrt{n}$. Therefore the nonparametric estimators will always be the leading terms in the sequence $\hat{g}_j(\rho_n)$. and the asymptotic distribution of $\hat{g}_j(\rho)$ is unaffected by the fact that β_j is estimated. Despite this fact, an analytical expression for the variance of $g_j(\rho)$ is not easily calculated, since the semiparametric estimator of $g_j(\rho)$ involves a complex structure of random variables. For example the Nadaraya-Watson estimator, is a ratio of random variables; thus standard central limit theorems cannot be applied directly. This is complicated further by the potential correlation between the nonparametric estimators, that enters in the resulting expression for $\hat{g}_j(\rho)$ [15].

6.3.2 Bootstrap confidence intervals

Even if we could derive an analytical expression for the asymptotic variance and confidence intervals, we have reason to believe that the asymptotic theory provides a poor guide to the precision of the estimator. First of all, we have already seen indications of small sample problems in the previous appendix, where asymptotic equivalent estimators of the restricted models gave rather different estimates of the restricted $\beta's$. The magnitude of this problem is surely not smaller in the semiparametric case, since the rate of covergence is smaller than $\sqrt{n}$. Secondly,

[15] None of the papers we have read, derived an analytical expression for the variance of the nonlinear part of the model. Perhaps it is one of those subjects that is considered too trivial to publish a paper on - or perhabs, the lack of papers on the subject is due to its non-triviality. We don't know. What we know, is that we couldn't do it.

the various nonparametric estimators $\hat{m}_h^{\omega_j}(\rho)$, $\hat{m}_h^{X}(\rho)$ and $\hat{m}_h^{Z}(\rho)$ need not converge at the same rate, since the rate of convergence depends on the optimal bandwidth, and thereby on the curvature of the relationship between the variable of interest and ρ. Therefore, we regard the bootstrap technique as an desirable alternative, since it is tractable and may provide a better finite sample approximation.

Several bootstrap algorithms are available in the literature. We chose the naive bootstrap, where bootstrap samples are drawn (with replacement) from the original sample. One advantage of this method is that it is robust to heteroscedasticity. That is, we do not have to assume that errors are iid., which is the case for the wild bootstrap, where the bootstrap samples are based on the residuals alone. This is particular relevant in our case, since we have to allow for dependence of the errors across the equations in the demand system.

1. *Sample with replacement* $\{\omega_{1i}, ...\omega_{Mi}, X_i, \rho_i, Z_i\}_{i=1}^{n}$ *from original sample.*

2. *Compute* $\hat{g}_j(\rho)$

3. *Repeat 1 and 2 B times (B=200 in our experiment)*

4. *Calculate confidence intervals as the* $1-\alpha/2$ *and* $\alpha/2$ *fractiles*

7 Appendix 3: Spline vs semiparametric estimation

In appendix 1 we made the following points:

- Testing a collective model with caring men and women against a collective model where only women exhibit caring we find that: the probability that one finds men to exhibit caring on a given interval of ρ increases when doubling the number of splines of the unrestricted model form 10 to 20. This, we argue, is because the variance becomes too high, making it more difficult to reject the hypothesis of flatness. If this point holds in general, then the robustness of the BL results is weakened.

- It is problematic to rely on asymptotic properties of the parametric model when the sample is relatively small. The difference in $\hat{\beta}_{BL}^{R}$ and $\hat{\beta}_{GMM}^{R}$ indicate that there are small sample problems and point to the need to study implications further.

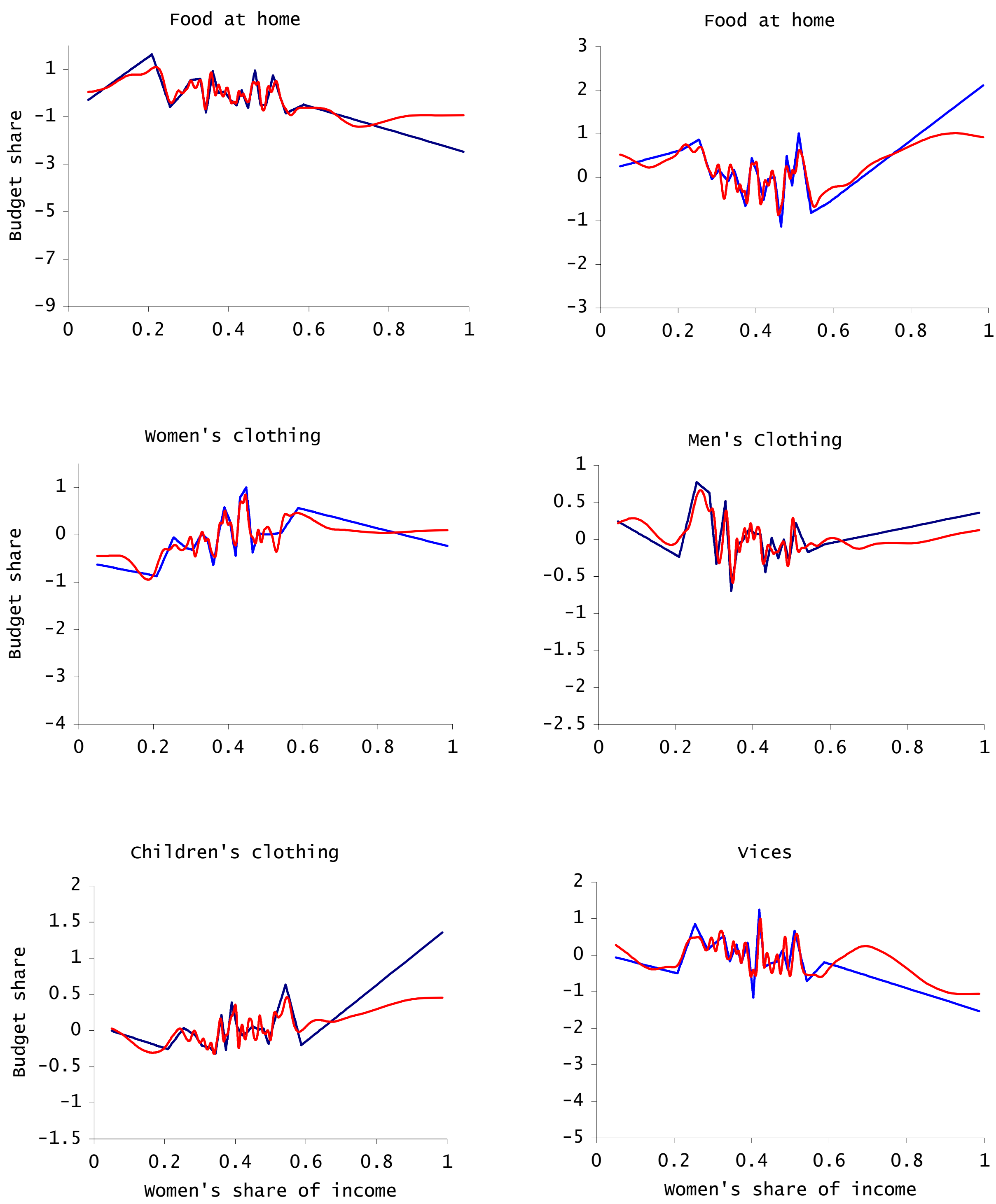
In Appendix 1 we did not reach any decisive conclusions about how many splines to use in the unrestricted model or which model to prefer. In the technical Appendix 2 we studied alternative approaches to estimating the effect of ρ on ω. In the present appendix we will try to use what we learned in appendix 2 to address some of the unanswered questions from appendix 1. In particular, we would like to gain certainty about these the following: i) if we have to do Minimum Chi Squared gridsearching, how do we decide the number of splines in the unrestricted model? ii) can we test restricted parametric models against an unrestricted semiparametric alternative.

7.1 Splines and Bandwidth

Is there a parallel between the choice of splines in the parametric regression and the choice of bandwidth in the semiparametric regression? It seems so, since both these choices determine the contribution of surrounding observations in the estimation of the slope at a given point. The difference is, ofcourse, that the spline technique is discontinuous and gives uniform weight to a limited number of observations, whereas kernel technique is continuous and gives non-uniform weight to surrounding observations.

But how close are the two estimation techniques? Can we get similar results using parametric and semi-parametric estimation? The closest we can get to the BL spline estimation semi-parametrically is K-nearest-neighbor estimation. With the K-N-N adaptive bandwidth adaptor ξ=1.0 and a low bandwidth, $h = 0.01$, we can produce a $g(\rho)$ curve that closely resembles the 20 spline $D_s\Gamma_j$ curve. See figure A3.1 below

Table A3.1: 20-spline unrestricted v. semiparametric fit with bandwidth $h = 0.01$ and $\xi = 1$



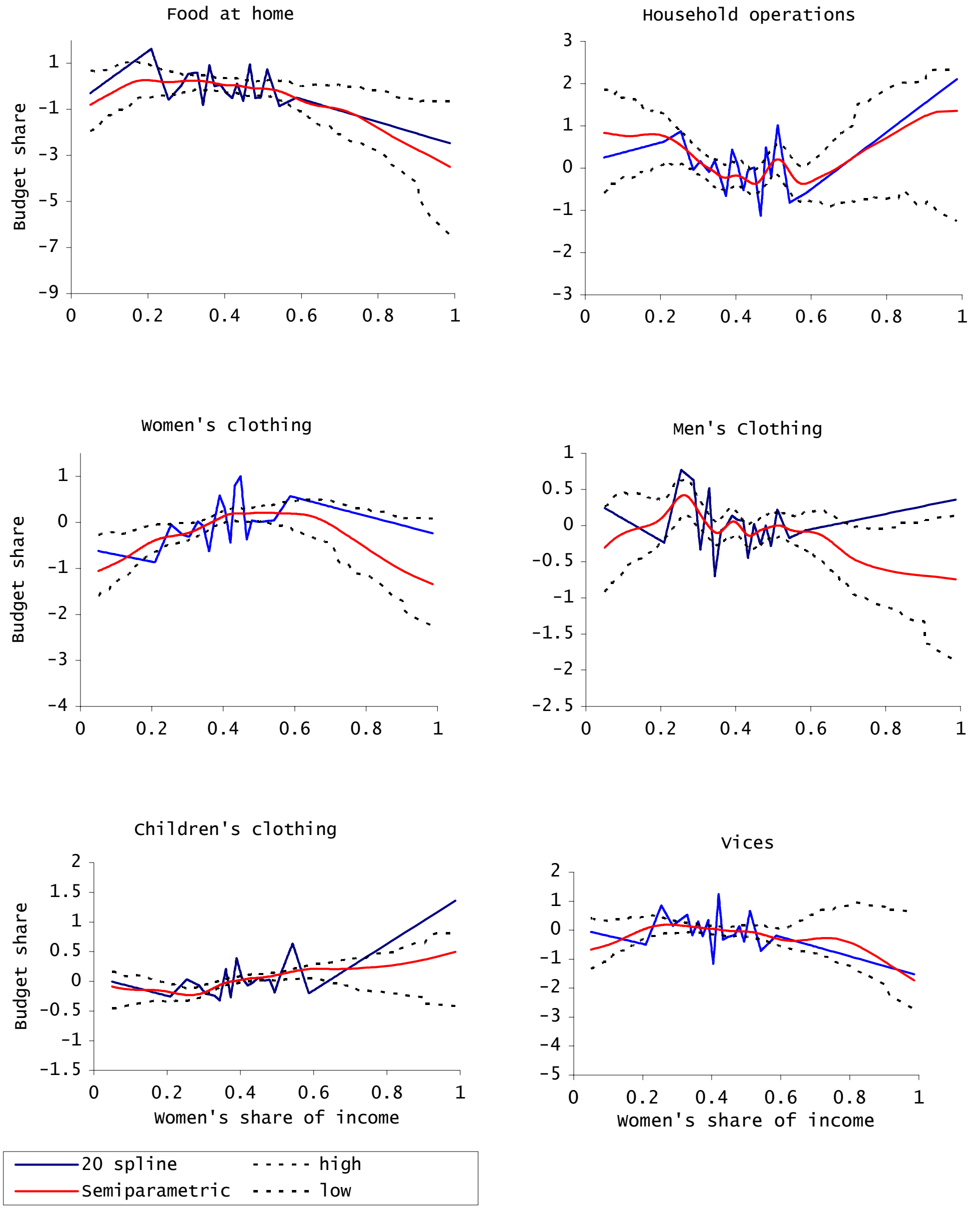
The bandwidth used for this, h=0.01, is much lower than the optimal global bandwidth, h^*, which we found to be between 0.2 and 0.4. We take this to indicate that 20 splines in a spline sense represents a too low bandwidth, and that a lower number of splines should have been chosen for the unrestricted model. This however is not a conclusive argument to drop the 20-spline model and we need to substantiate our case further.

7.2 Spline curves and confidence intervals

7.2.1 The unrestricted model v. semi-parametric confidence intervals

In appendix 2 we learned how to calculate bootstrapped confidence intervals for the semiparametric estimator. Semiparametric estimation is generally a less efficient estimation technique than GMM, so our ex ante expectation was that a semiparametric 95%confidence interval would comfortably contain the parametrically estimated curve. However, the 20-spline $D_s\Gamma_j$ curve does not fit inside the bootstrapped 95%-confidence intervals of the semi-parametric estimation. We have verified this for ξ=0.2, 0.25, and 0.5. We feel this demonstrates that choosing too many splines introduces spurious volatility. We document this for ξ=0.5 in figure A3.2 below.

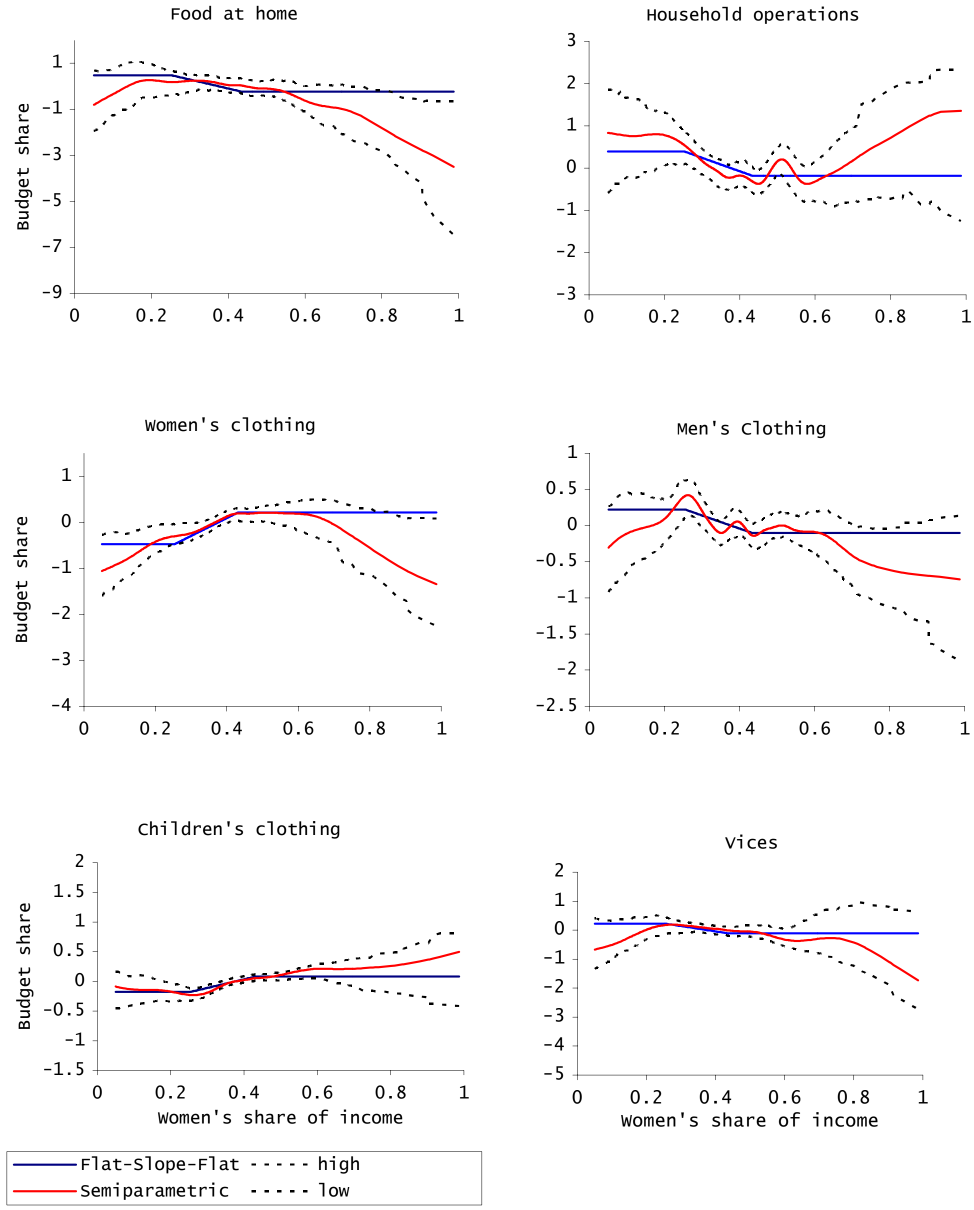
Figure A3.2: 20 unrestricted splines v. semiparametric fit with optimal bandwidth and $\xi = 0.5$



7.2.2 The collective model with caring v. the semi-parametric confidence intervals

The flat-slope-flat model generally fits inside the bootstrapped confidence intervals of the semi-parametric estimation. We have verified this for ξ=0.2, 0.25, and 0.5. We feel this lends evidence to the robustness of a collective model with caring. The curve does, however, stray outside the 95% confidence intervals for some of the less densely populated parts of rho. We document this for ξ=0.5 in figure A3.3 below.

Figure A3.3: Flat-Slope-Flat v. semiparametric fit with optimal bandwidth and $\xi = 0.5$



How do all these observations weigh up? From looking at the plots we saw indications that the 20-spline unrestricted model had a more problematic fit than the flat-slope-flat model. But we need a more systematic way of testing the fit of different model specifications..

7.3 Specification testing

How well does the linear parametric specifications, suggested by Browning and Lechene (2001), fit in comparison with the semiparametric specifications? The relevance and appropriateness of our non- and semiparametric analysis may be judged by its ability to answer this question. If the semiparametric approach shall be used for more than looking at data, it can be argued that it must be able to discriminate between different types of structural models.

On the other hand, if the semiparametric model can be used to identify some specification problems in the parametric model, we have certainly gained a lot. One issue is the number of splines in the unrestricted model. As we have seen, the choice of bandwidth is crucial for the shape of the semiparametric curve. In the unrestricted model in BL, the analogue to the bandwidth is the number of join point in the spline regression. We will argue that since the bandwidth is chosen optimally in the semiparametric regression, the number of join points in the spline regression, should be chosen such that the parametric curve fits the semiparametric curve best.

In order to determine the number of join points in the spline regression, we suggest a procedure where the model that minimizes the distance to the semiparametric fit, is chosen. In order to evaluate the distance, we suggest a convenient goodness of fit test. Specifically, we evaluate the Weighted Average Squared Error between the semiparametric estimate $\hat{g}_{jh}(\rho)$ and the predicted values based on the spline regression $\tilde{g}_{js}(\rho)$ (the subscript s denotes the number of splines). The test statistic is expressed as

$$Q_j = \frac{1}{n}\sum_{i=1}^{n}\frac{\left(\hat{g}_{jh}(\rho_i) - \tilde{g}_{js}(\rho_i)\right)^2}{Var\left(\hat{g}_{jh}(\rho_i)\right)} \tag{14}$$

where $Var\left(\hat{g}_{jh}(\rho_i)\right)$ is the variance of the semiparametric fit. The reason why we choose the variance of the semiparametric estimator is, that this is the unrestricted model against which we test the parametric fit.

Under the null $\tilde{g}_{js}(\rho_i)$ lies within the bootstrapped confidence band of $\hat{g}_{jh}(\rho_i)$. Clearly the distribution of this test statistic is non-standard under the null, therefore we simulated a distribution of the test statistic, under the assumption that semiparametric fit based on the original data was equal to the semiparametric fit based on the bootstrap samples.

If the test statistic lies within the simulated distribution, the parametric model does not deviate more from the semiparametric fit than the semiparametric fit based on the bootstrap sample does, or, equivalently, the parametric model could lie within the confidence intervals of the semiparametric fit. Thereby we can rank how the parametric models with different number of splines fit the semiparametric curves.

Below we present the simulated distribution of the estimator for various adaptors.

Table A3.1: Table of Critical Values for the Bootstrapped Goodness of Fit Statistic

Adaptor=0.2

Bootstrap Critcal values for the Godness of fit-test statistic			
	95%	90%	80%
Food at home	**2.696**	2.094	1.530
Household operations	**2.344**	2.078	1.421
Women's clothing	**2.669**	2.210	1.850
Men's clothing	**1.819**	1.673	1.449
Children's clothing	**3.417**	2.726	2.080
Vices	**3.253**	2.447	1.910
Joint test	**1.778**	1.603	1.396

Adaptor=0.25

Bootstrap Critcal values for the Godness of fit-test statistic			
	95%	90%	80%
Food at home	**2.421**	2.080	1.432
Household operations	**2.175**	1.820	1.473
Women's clothing	**2.702**	2.116	1.770
Men's clothing	**1.995**	1.789	1.353
Children's clothing	**3.079**	2.592	1.987
Vices	**3.214**	2.845	1.869
Joint test	**1.857**	1.543	1.364

Adaptor = 0.5

Bootstrap Critcal values for the Godness of fit-test statistic			
	95%	90%	80%
Food at home	**2.490**	1.972	1.500
Household operations	**2.104**	1.727	1.357
Women's clothing	**2.573**	2.002	1.676
Men's clothing	**1.974**	1.715	1.365
Children's clothing	**3.000**	2.501	2.012
Vices	**2.744**	2.317	1.781
Joint test	1.643	1.457	1.314

It may be argued, that to justify the method, too little evidence has been presented in terms formal studies of the asymptotic properties and simulation results of the performance of the test. However, we find substantial evidence for consistency between the plots and the goodness of fit test.

We can put this test to use in two distinct ways: i) Testing the fit of different theory models directly against the semiparametric fit. ii) Using the semiparametric fit as a *meta*-unrestricted model. That is, it can guide us in choosing an appropriate number of splines for the unrestricted *parametric* model.

7.3.1 Goodness of fit for relevant theory models

We are interested to compare various theoretical models with the semiparametric fit of the data. The way we do this is the following: for every model we take the join point combination that fared best in our Minimum Chi Squared grid search with 20 splines and compute the Goodness of Fit statistic. Then we compare the result of each model to the critical values and to the other models' goodness of fit statistics. This is a suboptimal procedure as we have no certainty that the minimum chi-squared grid search found the optimal join points for each model. This could be the subject for further study. We present the results in table A3.2 below.

Table A3.2: Test Results for restricted models v. semiparametric

Adaptor 0.2

Test statitics: For five restricted models vs semiparametric						
Test statistics:	Flat-slope-flat	Slope-flat	Flat-slope	Slope	Unitary	95%
Food at home	1.595	1.499	0.850	0.847	0.942	2.696
Household operations	1.568	1.521	2.198	2.143	2.458	2.344
Women's clothing	0.923	0.415	1.661	1.449	6.355	2.669
Men's clothing	0.794	0.780	0.891	0.912	1.317	1.819
Children's clothing	1.297	1.733	0.809	0.967	9.822	3.417
Vices	1.017	0.880	0.461	0.490	1.328	3.253
Joint test	**1.199**	**1.138**	**1.145**	**1.135**	**3.704**	**1.778**

Adaptor 0.25

Test statitics: For five restricted models vs semiparametric						
Test statistics:	Flat-slope-flat	Slope-flat	Flat-slope	Slope	Unitary	95%
Food at home	1.217	1.143	0.645	0.643	0.812	2.421
Household operations	1.263	1.252	1.887	1.848	2.281	2.175
Women's clothing	0.714	0.331	1.720	1.523	5.648	2.702
Men's clothing	0.795	0.804	0.905	0.926	1.310	1.995
Children's clothing	1.248	1.768	0.922	1.090	9.477	3.079
Vices	1.001	0.868	0.428	0.466	1.293	3.214
Joint test	**1.040**	**1.028**	**1.084**	**1.083**	**3.470**	**1.857**

Adaptor 0.5

Test statitics: For five restricted models vs semiparametric						
Test statistics:	Flat-slope-flat	Slope-flat	Flat-slope	Slope	Unitary	95%
Food at home	0.952	0.931	0.473	0.485	1.014	2.490
Household operations	1.198	1.208	1.830	1.799	2.317	2.104
Women's clothing	0.378	0.228	1.467	1.332	4.802	2.573
Men's clothing	0.870	0.929	1.054	1.077	1.499	1.974
Children's clothing	1.103	1.600	0.889	1.048	7.820	3.000
Vices	0.856	0.775	0.395	0.437	1.115	2.744
Joint test	**0.893**	**0.945**	**1.018**	**1.030**	**3.094**	**1.643**

We see that the unitary model is rejected and that the other all pass the joint test for goodness of fit with test statistics which are very close. For an adaptor of 0.2 the collective-no-caring models fares best, for an adaptor of 0.25, slope-flat fares best, and for an adaptor of 0.5 flat-slope-flat fares best. It appears that the primary strength of this test is, that we can have great certainty in discarding models that fail the test (because of the wide confidence bands of the semiparametric model as compared to those of parametric models). But the test does not give us much guidance in choosing between the models that pass it.

As seen flat-slope-flat fares better as the adaptor increases. The reason is that a higher ξ increases the bandwidth in areas with sparce data, which effectively flattens the curve in these regions. Therefore, when the model is estimated with high values of ξ, it favors the hypothesis of Becker regions. Since the parametric model in BL2001 has an equal number of observations for each spline, the length of the splines increases as we move to the tails of ρ. That resembles the $k-NN$ estimator, which is equivalent to a kernel regression with $\xi = 1$. Hence, the parametric specification chosen by BL2001, implicitly favors the hypothesis of existence of Becker regions.

7.3.2 Goodness of fit for unrestricted spline models

We have argued that the data can not sustain an unrestricted model of 20 splines. But the argument has not been conclusive. We will now proceed to exclude unrestricted models with a number of splines that is incompatible with the semiparametric fit. We saw above that the unitary model was rejected so we conclude that 0 splines are to few for an unrestricted model! We also saw that the collective model passed the goodness of fit test, so we have found the lower limit of the number of splines to be one. In table A3.5 below we try to find the upper limit to number of splines that may be included in the unrestricted model.

Table A3.3 (panel a): Results for Unrestricted Spline Models vs. Semiparametric

Adaptor = 0.2								
		Test statistics: For 1-20 Splines vs. Semiparametric						
		Food at home	Household operations	Women's clothing	Men's clothing	Children's clothing	Vices	**Joint test**
95% critical		**2.696**	**2.344**	**2.669**	**1.819**	**3.417**	**3.253**	**1.778**
Splines	1	**0.847**	**2.143**	**1.449**	**0.912**	**0.967**	**0.490**	**1.135**
	2	**0.719**	**0.740**	**0.371**	**0.755**	**1.784**	**0.464**	**0.806**
	3	**0.523**	**0.822**	**0.388**	**0.892**	**2.048**	**0.756**	**0.905**
	4	**0.982**	**0.827**	**2.174**	**1.033**	**2.658**	**2.088**	**1.627**
	5	**0.846**	**0.628**	5.065	**1.026**	**0.992**	**1.558**	**1.686**
	6	**1.346**	**0.999**	4.154	**1.725**	5.385	**1.338**	2.491
	7	**1.292**	**0.712**	4.147	**1.196**	10.228	**2.292**	3.311
	8	**1.240**	**1.021**	9.347	**0.986**	8.042	**1.908**	3.757
	9	**1.787**	**1.615**	9.330	**1.164**	8.980	6.006	4.814
	10	**2.546**	2.366	7.422	**1.311**	9.538	3.789	4.495
	11	**1.863**	3.048	25.085	**1.402**	12.467	5.850	8.286
	12	**2.259**	**1.947**	11.150	2.309	10.142	8.848	6.109
	13	4.254	4.687	23.542	2.242	16.048	7.125	9.65
	14	5.414	2.347	24.682	1.956	17.558	15.238	11.199
	15	3.883	4.134	22.891	2.960	22.892	6.188	10.491
	16	3.807	3.986	25.720	3.838	21.461	17.818	12.771
	17	7.112	5.156	26.627	3.190	27.162	9.590	13.139
	18	6.242	4.053	27.054	3.184	27.919	23.971	15.404
	19	5.804	7.017	26.777	7.536	27.701	15.416	15.042
	20	7.983	5.499	29.453	5.331	27.676	24.597	16.756

Table A3.3 (panel b): Results for Unrestricted Spline Models vs. Semiparametric

Adaptor = 0.25								
		Test statistics: For 1-20 Splines vs. Semiparametric						
		Food at home	Household operations	Women's clothing	Men's clothing	Children's clothing	Vices	**Joint test**
95% critical		**2.421**	**2.175**	**2.702**	**1.995**	**3.079**	**3.214**	**1.857**
Splines	1	**0.643**	**1.848**	**1.523**	**0.926**	**1.090**	**0.466**	**1.083**
	2	**0.543**	**0.665**	**0.285**	**0.808**	**1.919**	**0.432**	**0.775**
	3	**0.406**	**0.727**	**0.311**	**0.905**	**2.197**	**0.656**	**0.867**
	4	**0.773**	**0.731**	**1.931**	**1.038**	**2.453**	**1.965**	**1.482**
	5	**0.651**	**0.576**	4.344	**1.044**	**0.923**	**1.426**	**1.494**
	6	**1.081**	**0.843**	3.716	**1.548**	4.938	**1.267**	2.232
	7	**1.021**	**0.645**	3.695	**1.094**	9.864	**2.102**	3.070
	8	**0.999**	**0.870**	7.834	**0.900**	7.218	**1.871**	3.282
	9	**1.422**	**1.353**	7.839	**1.079**	8.658	5.548	4.316
	10	**2.001**	**1.951**	6.400	**1.203**	8.849	3.612	4.002
	11	**1.506**	2.389	21.031	**1.266**	12.192	5.716	7.350
	12	**1.851**	**1.635**	9.341	2.164	9.170	8.625	5.464
	13	3.379	3.767	20.463	2.072	15.730	6.667	8.679
	14	4.276	**1.874**	19.774	**1.848**	15.981	15.282	9.839
	15	3.231	3.345	20.102	2.851	22.164	5.951	9.607
	16	3.033	3.307	21.531	3.670	20.135	18.197	11.646
	17	5.505	4.071	22.172	2.905	26.539	9.202	11.732
	18	4.987	3.207	23.328	2.961	26.271	24.154	14.151
	19	4.745	5.913	22.251	7.267	26.374	15.535	13.681
	20	6.138	4.477	24.795	5.035	27.272	25.106	15.471

Table A3.3 (panel c): Results for Unrestricted Spline Models vs. Semiparametric

Adaptor=0.5								
		Test statistics: For 1-20 Splines vs. Semiparametric						
		Food at home	Household operations	Women's clothing	Men's clothing	Children's clothing	Vices	**Joint test**
95% critical		**2.490**	**2.104**	**2.573**	**1.974**	**3.000**	**2.744**	**1.643**
Splines	1	**0.485**	**1.799**	**1.332**	**1.077**	**1.048**	**0.437**	**1.030**
	2	**0.403**	**0.738**	**0.218**	**0.944**	**1.695**	**0.341**	**0.723**
	3	**0.364**	**0.763**	**0.229**	**1.027**	**1.873**	**0.431**	**0.781**
	4	**0.610**	**0.787**	**0.974**	**1.156**	**1.753**	**1.431**	**1.119**
	5	**0.524**	**0.622**	**2.392**	**1.040**	**0.736**	**1.045**	**1.060**
	6	**0.815**	**0.654**	**1.927**	**1.112**	3.836	**0.968**	**1.552**
	7	**0.838**	**0.632**	**1.895**	**0.828**	7.411	**1.541**	2.191
	8	**0.743**	**0.682**	4.817	**0.618**	5.367	**1.437**	2.277
	9	**1.098**	**0.941**	5.207	**0.732**	6.320	3.983	3.047
	10	**1.673**	**1.331**	3.783	**0.835**	6.675	2.747	2.841
	11	**1.251**	**1.723**	13.749	**0.805**	8.996	4.155	5.113
	12	**1.577**	**1.113**	6.333	**1.449**	6.963	6.456	3.982
	13	3.142	2.601	12.736	**1.469**	11.407	4.895	6.042
	14	4.162	**1.396**	13.485	**1.268**	11.806	11.443	7.260
	15	3.049	2.291	12.751	2.135	16.270	4.431	6.821
	16	2.746	2.297	14.007	2.994	15.001	13.647	8.449
	17	5.212	2.843	14.706	2.255	19.710	6.811	8.590
	18	4.859	2.195	15.001	2.279	20.013	18.058	10.401
	19	4.449	4.143	14.676	6.103	20.086	11.559	10.169
	20	5.914	3.176	16.107	4.226	20.543	18.884	11.475

The three tables in A3.3 gives us a background for choosing the optimal number of splines in an unrestricted model. We see that 20 splines is rejected by a huge margin for all adaptors shown and that no more than 6 splines can be accepted for $\xi = 0.5$ when testing jointly. For $\xi = 0.2$ or 0.25 no more than 5 splines can be accepted when testing jointly. We expect that for $\xi = 1$(the K-N-N case) a higher number of splines could be accepted. However, the tests substantiates our suspicion, that the BL unrestricted model had too many splines causing spurious fluctuations.

Among the models that pass the test, we find 2 splines to have the best fit for all tested values of ξ. We were surprised at how many models were rejected. Looking at plots for each market of each model we found, however, that plots that looked problematic represented models that were consistently rejected by the Goodness-of-fit test. Finally, looking at the plots we find that already when going from 3 to 4 splines we start to see jumpy and theoretically implausible movements in the center of the spline curve. (Inter-active Excel graps showing this are available from the authors and will be available at www.metrics.dk).

7.4 Redoing the grid search with fewer splines

What we found above indicates that the number of splines in the unrestricted model should be substantially reduced. We redid the grid searches with unrestricted models of 5-12 splines. Testing on a 5% level, the results can be summarized as follows:

- For 5 and 6 splines in the unrestricted model, all less restricted models can be reduced to the slope model - that is, the collective model and no caring preferences.
- For 7 and 8 splines in the unrestricted model, all less restricted models can be reduced to the slope-flat model, and the slope–flat model cannot be reduced further. This lends support to a collective model with caring women.
- For 9-12 and 20 splines in the unrestricted model, all less restricted models can be reduced to the flat-slope-flat model, and the flat-slope–flat model cannot be reduced further. This lends support to a collective model with caring.

The protagonist for models with caring might argue, that obviously fewer splines means that the first and last spline might be 'too long' over ρ to uncover a real and existent caring phenomenon. We agree and suggest that a larger and better data set is procured or that his conclusions be based on theoretical arguments and semi-parametric analysis. Till then we are not satisfied that anyone cares - especially not men.

5 Spline Framework **GMM** Estimation		flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
						Unrestricted model								
Restricted model	flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
	slope				1.62	**7.41**	4.11	3.28	11.08	78.56	0.27	14.63	5.62	58.13
	Flat-slope					**0.05**	0.15	0.13	0.67	13.49	0.01	1.52	0.52	20.71
	Slope-flat						18.27	13.98	38.12		1.24	39.32	16.93	82.79
	Flat-slope-flat						**9.35**	**7.28**	**24.32**	**99.99**	**0.56**	**28.15**	**11.11**	**76.00**
	Slope-slope							14.43			0.66	67.35	26.04	95.82
	Slope-flat-slope										0.64	83.86	33.43	98.18
	fl-sl-fl-sl										0.04	39.50	11.78	90.54
	sl-fl-sl-fl										0.00	1.57	0.49	37.57
	fl-sl-fl-sl-fl													
	Slope-slope-slope												4.20	95.72
	sl-fl-sl-fl-sl													100.00
	sl-sl-sl-sl-sl													

5 Spline Framework **BL** Estimation		flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
						Unrestricted model								
Restricted model	flat		0.00	0.01	0.00	**0.00**	0.00	0.00	0.00	0.06	0.00	0.02	0.01	1.02
	slope				2.47	**10.20**	8.51	12.76	20.16	83.91	1.21	28.88	15.44	79.66
	Flat-slope					**0.16**	0.66	1.23	2.47	24.42	0.09	5.90	2.85	46.25
	Slope-flat						27.99	37.61	51.84		3.98	57.97	34.19	94.00
	Flat-slope-flat						**15.95**	**23.80**	**36.49**	**99.98**	**2.05**	**45.80**	**25.10**	**90.61**
	Slope-slope							80.51			1.73	79.06	42.29	98.79
	Slope-flat-slope										0.64	69.76	33.43	98.18
	fl-sl-fl-sl										0.12	52.08	22.18	96.52
	sl-fl-sl-fl										0.00	4.28	1.93	60.89
	fl-sl-fl-sl-fl													
	Slope-slope-slope												7.24	98.20
	sl-fl-sl-fl-sl													100.00
	sl-sl-sl-sl-sl													

Probablitity that the difference in Minimum Chi Squared Criteria between models is zero

		Unrestricted model												
	6 Spline Framework **GMM** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
	flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
	slope				5.58	**8.96**	5.66	2.90	9.40	53.68	49.02	8.62	48.70	31.75
	Flat-slope					**0.17**	0.46	0.24	1.07	11.21	10.67	1.38	15.23	11.51
	Slope-flat					**28.02**	12.25	6.19	18.88	82.75	76.08	15.81	68.95	44.60
Restricted model	**Flat-slope-flat**						**11.37**	**5.60**	**18.48**	**87.08**	**79.69**	**15.65**	**70.98**	**45.40**
	Slope-slope							6.54	55.40			33.85	99.41	69.41
	Slope-flat-slope											60.41		83.63
	fl-sl-fl-sl											23.97	99.88	65.07
	sl-fl-sl-fl										22.48	1.79	37.44	23.67
	fl-sl-fl-sl-fl											1.60	41.52	25.30
	Slope-slope-slope													80.83
	sl-fl-sl-fl-sl													21.67
	sl-sl-sl-sl-sl													

		Unrestricted model												
	6 Spline Framework **BL** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
	flat		0.00	0.01	0.00	**0.00**	0.00	0.00	0.00	0.02	0.03	0.01	0.13	0.32
	slope				6.33	**11.73**	10.07	16.03	17.28	57.76	56.87	19.66	66.07	54.66
	Flat-slope					**0.38**	1.34	2.62	3.16	16.74	17.68	5.08	30.45	28.96
	Slope-flat					**36.00**	20.07	30.25	31.51	84.56	81.93	32.14	83.72	68.32
Restricted model	**Flat-slope-flat**						**17.30**	**27.59**	**29.10**	**85.72**	**82.60**	**30.45**	**83.88**	**67.80**
	Slope-slope								67.53			51.30	99.89	84.78
	Slope-flat-slope								32.24			37.72	99.47	79.44
	fl-sl-fl-sl											36.38	99.98	80.08
	sl-fl-sl-fl										30.68	5.87	58.82	46.28
	fl-sl-fl-sl-fl											4.76	60.31	46.47
	Slope-slope-slope													87.56
	sl-fl-sl-fl-sl													34.64
	sl-sl-sl-sl-sl													

Probablitity that the difference in Minimum Chi Squared Criteria between models is zero

		Unrestricted model												
	7 Spline Framework **GMM** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
	flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	slope				3.52	**3.07**	3.83	1.43	4.06	32.64	25.87	3.52	26.47	16.42
	Flat-slope					**0.10**	0.48	0.17	0.61	7.80	6.18	0.70	8.58	6.28
	Slope-flat					**11.13**	10.91	4.06	10.74	66.17	53.68	8.34	47.74	28.10
	Flat-slope-flat						**16.44**	**5.92**	**15.73**	**85.23**	**70.99**	**11.68**	**59.77**	**34.74**
Restricted model	Slope-slope							3.87	25.42			18.25	89.41	49.79
	Slope-flat-slope											44.18		71.05
	fl-sl-fl-sl											19.46	98.19	54.72
	sl-fl-sl-fl										14.70	1.21	26.95	16.42
	fl-sl-fl-sl-fl											1.38	35.25	19.89
	Slope-slope-slope													74.28
	sl-fl-sl-fl-sl													18.59
	sl-sl-sl-sl-sl													

		Unrestricted model												
	7 Spline Framework **BL** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
	flat		0.00	0.01	0.00	**0.00**	0.00	0.00	0.00	0.02	0.02	0.01	0.07	0.16
	slope				4.73	**5.99**	8.21	11.69	10.99	40.46	38.50	11.06	46.74	37.00
	Flat-slope					**0.33**	1.58	2.62	2.63	13.47	13.41	3.41	22.19	20.09
	Slope-flat					**19.32**	19.26	25.85	23.61	71.13	66.63	21.20	68.97	52.29
	Flat-slope-flat						**22.19**	**30.04**	**27.20**	**81.00**	**75.47**	**23.95**	**75.00**	**56.28**
Restricted model	Slope-slope							62.20	41.59			32.80	95.66	69.60
	Slope-flat-slope								21.89			25.25	93.81	65.27
	fl-sl-fl-sl											27.84	98.47	69.08
	sl-fl-sl-fl										25.53	4.57	49.02	36.52
	fl-sl-fl-sl-fl											4.03	52.61	37.94
	Slope-slope-slope													81.83
	sl-fl-sl-fl-sl													29.66
	sl-sl-sl-sl-sl													

Probablitity that the difference in Minimum Chi Squared Criteria between models is zero

						Unrestricted model							
	8 Spline Framework **GMM** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
Restricted model	flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	slope				1.37	**0.85**	2.87	0.87	2.69	19.54	12.92	3.18	16.89	17.27
	Flat-slope					**0.05**	0.53	0.16	0.58	5.76	3.72	0.89	6.50	8.35
	Slope-flat					**6.32**	14.51	4.47	12.16	61.97	43.76	11.68	43.91	36.97
	Flat-slope-flat						**29.71**	**9.05**	**23.36**	**90.34**	**69.81**	**20.27**	**62.87**	**49.82**
	Slope-slope							2.78	20.40			20.86	77.67	56.76
	Slope-flat-slope											56.67	99.93	80.78
	fl-sl-fl-sl											26.21	93.11	65.10
	sl-fl-sl-fl										9.72	2.37	25.80	25.66
	fl-sl-fl-sl-fl											3.70	40.38	34.28
	Slope-slope-slope													78.99
	sl-fl-sl-fl-sl													32.37
	sl-sl-sl-sl-sl													

						Unrestricted model							
	8 Spline Framework **BL** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
Restricted model	flat		0.00	0.01	0.00	**0.00**	0.00	0.00	0.00	0.02	0.02	0.01	0.08	0.32
	slope				2.96	**2.45**	7.61	11.51	8.64	33.64	26.60	11.76	37.31	41.46
	Flat-slope					**0.21**	2.15	3.72	2.83	14.34	11.22	4.92	20.49	27.27
	Slope-flat					**10.12**	23.03	31.75	23.56	70.93	57.91	26.68	64.37	61.78
	Flat-slope-flat						**36.66**	**48.30**	**35.50**	**90.65**	**77.11**	**36.94**	**77.84**	**71.36**
	Slope-slope							79.79	31.19			37.18	88.66	76.30
	Slope-flat-slope								13.24			27.65	83.02	71.34
	fl-sl-fl-sl											38.70	95.85	79.59
	sl-fl-sl-fl										14.58	6.88	42.69	47.09
	fl-sl-fl-sl-fl											8.72	55.62	54.61
	Slope-slope-slope													85.90
	sl-fl-sl-fl-sl													46.49
	sl-sl-sl-sl-sl													

Probablitity that the difference in Minimum Chi Squared Criteria between models is zero

Unrestricted model

9 Spline Framework **GMM** Estimation (Restricted model)	flat	slope	Flat-slope	Slope-flat	**Flat-slope -flat**	Slope -slope	Slope-flat -slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope -slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
slope				0.79	**0.39**	2.38	0.40	1.58	10.32	4.92	2.20	12.10	12.51
Flat-slope					**0.02**	0.46	0.07	0.34	2.84	1.29	0.63	4.58	5.94
Slope-flat					**4.38**	16.81	3.05	10.14	48.05	25.49	10.84	40.31	32.89
Flat-slope-flat						**41.16**	**7.67**	**23.49**	**83.87**	**51.01**	**21.84**	**63.65**	**48.46**
Slope-slope							1.16	12.13		52.95	16.69	67.99	48.51
Slope-flat-slope											67.19	100.00	81.97
fl-sl-fl-sl											28.88	93.65	63.33
sl-fl-sl-fl										5.19	3.52	31.90	27.72
fl-sl-fl-sl-fl											8.49	62.59	43.20
Slope-slope-slope													74.93
sl-fl-sl-fl-sl													30.15
sl-sl-sl-sl-sl													

Unrestricted model

9 Spline Framework **BL** Estimation (Restricted model)	flat	slope	Flat-slope	Slope-flat	**Flat-slope -flat**	Slope -slope	Slope-flat -slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope -slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
flat		0.00	0.02	0.00	**0.00**	0.01	0.01	0.01	0.03	0.02	0.02	0.11	0.42
slope				2.83	**2.05**	8.22	11.66	7.57	25.34	17.68	10.62	32.37	37.74
Flat-slope					**0.21**	2.61	4.17	2.69	10.99	7.49	4.74	18.28	25.40
Slope-flat					**8.51**	25.28	32.75	21.37	58.83	42.72	24.81	58.72	58.02
Flat-slope-flat						**43.59**	**53.21**	**34.80**	**82.86**	**63.11**	**36.52**	**74.51**	**69.24**
Slope-slope							61.48	22.43		72.79	31.40	79.89	70.54
Slope-flat-slope								9.81		59.10	24.09	74.27	66.33
fl-sl-fl-sl											38.90	93.43	77.76
sl-fl-sl-fl										10.81	9.10	46.43	49.79
fl-sl-fl-sl-fl											14.08	66.50	60.73
Slope-slope-slope													83.87
sl-fl-sl-fl-sl													46.99
sl-sl-sl-sl-sl													

Probablitity that the difference in Minimum Chi Squared Criteria between models is zero

	Unrestricted model												
10 Spline Framework **GMM** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
flat		0.00	0.00	0.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
slope				0.16	**0.08**	0.69	0.06	0.53	3.31	1.63	0.93	4.54	4.43
Flat-slope					**0.00**	0.14	0.01	0.12	0.88	0.43	0.27	1.66	2.00
Slope-flat					**3.86**	15.11	1.52	9.80	41.60	22.68	11.18	34.00	23.59
Restricted model — **Flat-slope-flat**						**39.82**	**4.18**	**24.21**	**79.08**	**48.49**	**23.58**	**57.61**	**37.74**
Slope-slope							0.49	13.52		50.45	19.10	61.33	37.27
Slope-flat-slope											91.75		80.21
fl-sl-fl-sl											31.06	86.33	49.12
sl-fl-sl-fl										5.83	4.76	29.28	20.73
fl-sl-fl-sl-fl											10.67	55.84	32.55
Slope-slope-slope													56.50
sl-fl-sl-fl-sl													22.84
sl-sl-sl-sl-sl													

	Unrestricted model												
10 Spline Framework **BL** Estimation	flat	slope	Flat-slope	Slope-flat	**Flat-slope-flat**	Slope-slope	Slope-flat-slope	fl-sl-fl-sl	sl-fl-sl-fl	fl-sl-fl-sl-fl	Slope-slope-slope	sl-fl-sl-fl-sl	sl-sl-sl-sl-sl
flat		0.00	0.02	0.00	**0.00**	0.00	0.01	0.00	0.02	0.01	0.01	0.06	0.18
slope				1.06	**0.74**	4.14	6.19	3.94	14.23	9.52	6.39	19.45	21.57
Flat-slope					**0.08**	1.41	2.34	1.49	6.29	4.15	3.01	10.95	14.15
Slope-flat					**7.03**	23.29	30.45	19.46	54.29	38.14	23.59	52.04	46.22
Restricted model — **Flat-slope-flat**						**44.14**	**53.71**	**34.60**	**81.74**	**60.79**	**37.10**	**70.24**	**59.01**
Slope-slope							61.16	21.70		66.92	31.72	74.03	58.78
Slope-flat-slope								9.44		52.23	24.40	67.65	54.14
fl-sl-fl-sl											40.12	89.05	66.38
sl-fl-sl-fl										10.12	9.74	41.59	38.93
fl-sl-fl-sl-fl											15.73	61.74	49.57
Slope-slope-slope													70.32
sl-fl-sl-fl-sl													37.18
sl-sl-sl-sl-sl													

Probablitity that the difference in Minimum Chi Squared Criteria between models is zero

References

[1] Blundell, R. M. Browning and I. Crawford (2003), "Non-parametric Engel Curves and Revealed Preference", *Econometrica, 71(1), 205-240.*

[2] Blundell, R. and A. Duncan (1998): "Kernel regression in empirical microeconomics", *Journal of Human Resources,* **33**, 62-87.

[3] Blundell, R., A. Duncan and Krishna Pendakur (1998): "Semiparametric Estimation and Consumer Demand", *Journal of Applied Econometrics,* **13**, 435-461.

[4] Browning, M. and P. Chiappori (1998), "Efficient Intra-household Allocation: A General Characterization and Empirical Tests", *Econometrica, 66(6), 1241-1278.*

[5] Browning, M. and V. Lechene (2001): "Caring and Sharing: Tests Between Alternative Models of Intra-household Allocation", *Institute of Economics, University of Copenhagen, Discussion Paper 01-07*

[6] Browning, M., F. Bourguignon, P. Chiappori and V. Lechene (1994): "Incomes and Outcomes: A Structural Model of Intrahousehold Allocation", *Journal of Political Economy,* **102**, 1067-1096.

[7] Greene, H. W.(2002): "Econometric Analysis", *Prentice Hall*, Fifth Edition.

[8] Härdle, W. (1990): "Applied Nonparametric regression", *Econometric Society Monographs No. 19.*

[9] Jennen-Steinmetz, C and T. Gasser(1988): "A Unifying Approach to Nonparametric Regression Estimation", *Journal of the American Statistical Association,* ***83****, 1302-1309.*

[10] Johnston, J. and J. DiNardo (1997): "Econometric Methods", *McGraw-Hill,* Fourth Edition

[11] Robinson, P.M.(1988): "Root-N Consistent Semiparametric Regression", *Econometrica, 56(4), 931-954.*