An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data

Charlier, E.; Melenberg, B.; van Soest, A.H.O.

Publication date:
1997

An Analysis of Housing Expenditure

Using Semiparametric Models and Panel Data

by E. Charlier, B. Melenberg and A. van Soest[1]

Tilburg University
Department of Econometrics
P.O. Box 90153
5000 LE, Tilburg
The Netherlands
E-mail first author: E.Charlier@kub.nl

February 1997

Abstract

In this paper we model expenditure on housing for owners and renters by means of endogenous switching regression models for panel data. We explain the share of housing in total expenditure from a household specific effect, family characteristics and total expenditure, where the latter is allowed to be endogenous. We consider both random and fixed effects panel data models. We compare estimates for the random effects model with estimates for the linear panel data model in which selection only enters through the fixed effects and with estimates allowing for fixed effects and a more general type of selectivity. Differences appear to be substantial. The results imply that the random effects model as well as the linear panel data model are too restrictive.

Keywords: sample selection, Engel curves, semiparametric models, panel data

JEL classification: C14, C33, R21

[1] We thank Marno Verbeek and seminar participants at Rice University for helpful comments. Research of the third author was possible due to a fellowship of the Netherlands Royal Academy of Arts and Sciences (KNAW). We are grateful to Statistics Netherlands (CBS) for providing the data.

1. Introduction

In most industrialized countries housing is one of the main categories of household expenditure. Its understanding is therefore crucial for analyzing household consumption. The decision how much to spend on housing is strongly related to the choice between renting and owning. The standard reference is Lee and Trost (1978), who explain annual family expenditure on housing taking the decision to own or to rent explicitly into account. They use cross-section data and apply a switching regression model with endogenous switching and normally distributed error terms, which is also referred to as Tobit V by Amemiya (1984).

Several authors have focused on different aspects of the demand for housing. Ioannides and Rosental (1994) analyze the choice between renting and owning in relation to consumption and investment demand for housing. Zorn (1993) models the fact that some households cannot obtain a mortgage due to mortgage constraints, which results in a kinked budget set. Haurin (1991) investigates the same issue as Zorn (mortgage constraints) and analyzes how the intertemporal variation in income affects tenure choice.

In this paper we focus on housing expenditure and thus not on housing assets, housing equity or mortgage constraints. We will combine the model by Lee and Trost (1978), henceforth referred to as LT model, with the consumer demand literature on expenditure on goods. We extend the LT model in two ways. First, we use panel data, and can therefore allow for time constant unobserved household specific effects which can be correlated with the regressors. In other words, we will allow for fixed effects, which would be impossible in the cross-section context. The usual cross-section model imposes independence between individual effects and regressors or instruments, which, in a panel data context, leads to the more restrictive random effects model. There are two types of fixed effects models that we consider: a linear model in which selectivity only enters through the fixed effects, and a model similar to that of Kyriazidou (1995), which incorporates more general selectivity effects than the linear model. We will compare results for these two fixed effects models with those of the random effects model.

Secondly, when modelling the budget share spent on housing as a function of total expenditure, account has to be taken of the possibility of endogeneity of total expenditure. We test for this and present estimates allowing for it.

Our main findings are that the random effects model, the model in which selectivity enters through the fixed effects only, and the model which assumes that total expenditure is exogenous, are all rejected against the more general fixed effects model. Moreover, the models lead to different conclusions about aggregate elasticities of housing expenditure.

The remainder of this paper is organized as follows. In section 2 we describe the data, drawn

from the Dutch Socio Economic Panel, 1987-1989. In section 3 we discuss various parametric and semiparametric panel data models and estimates explaining housing. Section 4 concludes.

2. Data

We will use data from the waves 1987-1989 of the Dutch Socio-Economic Panel (SEP). Although this panel exists since 1984, information concerning housing is only present since 1986 and wealth data are available as of 1987. We will use a cleaned subsample for each year with information on family characteristics (including marital status, number of children living with the family, age of the head of household, education level and region of residence), and labour market characteristics (including hours of work, gross and net earnings). The labour market characteristics are used to construct household income which consists of labour earnings, other family income (mainly from letting rooms or child allowances), benefits and pensions. Personal income of children is excluded. Asset income and capital gains are also excluded, because this type of income is strongly related to the home ownership decision. Wealth data[2] are used to construct savings.[3] For issues on cleaning the savings data we refer to Camphuis (1993). Income and savings are used to construct total expenditure. Expenditure and income are reported in *Dutch guilders per month*.

The budget share spent on housing is defined as the fraction of total expenditure spent on housing. Housing expenditure for renters is the amount of money spent on rent by the family (i.e., excluding gas/water/electricity/heating as well as rental subsidy). For owners expenditure on housing consists of the following components: net interest costs on the mortgage,[4] net rent paid if the land is not owned, taxes on owned housing,[5] costs of insuring the house, opportunity costs of housing equity, maintenance costs, and minus the increase of the value of the house. The latter three costs components are not observed in the data. The opportunity cost of the foregone interest on housing equity is set equal to 4% of the value of the house minus the mortgage value. Maintenance costs and the increase of the value of the house are set equal to 2% and 1% of the value of the house, respectively. In Appendix A, we shall investigate the sensitivity of the results

---

[2] Net wealth is constructed using checking accounts, savings and deposits accounts, saving certificates, certificates of deposits, bonds and mortgage bonds, shares, options and other securities, antiques, jewels, coins etc., real estate other than one's own residence, one's own car, claims against private persons, other assets, life-insurance with saving elements, personal loan or revolving credit, hire-purchase and other loans.

[3] We also corrected for donations, bequests, and capital gains.

[4] Mortgage interest payments are tax deductible. See Data Appendix for computation of the marginal tax rate.

[5] This refers to a direct tax on housing property and to extra income tax due to adding the imputed rental value of the house to household income.

with respect to these choices. It appears that our main results are hardly affected.

The Data Appendix contains some further details on the construction of the sample and the variables of interest, and a comparison with macro data on housing expenditure. Given this sample, we excluded households with a missing observation for expenditure, and a few households with housing budget share larger than 3.[6] For the 1987 wave this reduces the dataset from 3006 to 2357 observations. Variable definitions and summary statistics for the three resulting panel waves are presented in table 1. The average budget share of housing is approximately 0.24 for renters and about 0.22 for owners. For both groups, this share decreased slightly over time. From 1987 to 1989, average total monthly expenditure increased from 2304 to 2477 for renters and from 3233 to 3606 for owners. For both owners and renters, the average age of the head of the household and the average values of the three region dummies do not change much over time. For renters the fraction of married household heads as well as the number of children living with the family decreased slightly.

We present several graphs for 1987, the year we will use to obtain estimates in the random effects model. In figure 1, nonparametric density estimates for the budget shares BS0 for renters and BS1 for owners are reported, as well as nonparametric regressions of these budget shares on log(total expenditure). Both budget share distributions are skewed to the right. Some budget shares larger than one are observed (see footnote 6). The regression estimates suggest that the housing budget share is nonlinear in log(total expenditure), but can be approximated reasonably well by a quadratic function. This is similar to what Banks et al. (1994) find for many commodity groups.

In figure 2, the result of a nonparametric regression of the probability of owning a house as a function of log(total income) is presented together with the frequency distribution of log(total income). Families with higher total income tend to have a higher probability of owning a house for the main part of the income range.

3. Models

The panel data models we consider allow for household specific effects which either are assumed to be independent of the explanatory variables (random effects) or allowed to be correlated with the explanatory variables (fixed effects). Starting point is the following system of equations.

$$d_{it} = 1(\pi' x_{it} + \eta_i - u_{it} \geq 0).$$
$$y_{0it} = \beta_0' x_{it} + \alpha_{0i} + \varepsilon_{0it} \quad \text{if } d_{it}=0$$

[6] Some budget shares are larger than one, possibly due to the fact that total expenditure is constructed from income minus savings, which might lead to substantial measurement errors for some households.

$$y_{1it} = \beta_1' x_{it} + \alpha_{1i} + \varepsilon_{1it} \quad \text{if } d_{it}=1$$

Here the indices i and t refer to household i in period t (t=1,...,T). d_{it} is a sector selection dummy variable which is 1 for owners and 0 for renters, x_{it} is a vector of explanatory variables (log total expenditure and its square, and taste shifters), y_{0it} and y_{1it} are the budget shares spent on housing for renters and owners, respectively. α_{0i}, α_{1i}, and η_i are unobserved household specific time-invariant effects, ε_{0it}, ε_{1it}, and u_{it} are error terms, varying across households as well as time. β_1, β_0 and π are vectors of unknown parameters.

3.1. Random effects

In a random effects model where α_{0i}, α_{1i}, η_i, ε_{0it}, ε_{1it}, and u_{it} are normally distributed and independent of x_{it}, we could apply the estimation procedure proposed by Vella and Verbeek (1994). However, their estimation procedure relies strongly on the normality assumptions. An alternative approach to estimate the slope parameters in the random effects panel data model is to focus on only one wave of data (i.e. a cross-section), drop the t-subscript, include the random effects in the error terms which then become $v_i=(\alpha_{0i}+\varepsilon_{0i},\alpha_{1i}+\varepsilon_{1i},\eta_i-u_i)$, and use existing estimation techniques for a cross-section endogenous switching regression model. By using a semi-parametric cross-section model estimator, consistent estimates for the slope parameters in the three equations can be obtained without imposing normality of the errors.

Even if the error terms in the cross-section endogenous switching regression model are independent of the regressors, without further distributional assumptions, identification of the parameters of this model requires that at least one component of both β_1 and β_0 is equal to zero (possibly the same), while the corresponding components of π are not equal to zero. Such exclusion restrictions are not required if normality of the errors is imposed, but are needed in a semi-parametric framework. We will therefore impose them throughout. Our main exclusion restriction is that the head of household's education level is not included in the budget share equations. Education level may affect the family's information set and interest in financial matters, and may therefore influence the family's portfolio choice, of which the choice between owning and renting is an important component. It is not obvious, however, why education should have a direct impact on housing consumption, given the ownership decision. Another variable which we exclude from the share equations is the number of children. Although there is no *a priori* reason for this, the number of children was always insignificant in the share equations at any conventional level.

As mentioned above, x_i will include the log of total expenditure and its square, which might be

endogenous. For example, in the two-stage budgeting literature[7] a household first decides how much to spend in total in each period and, given this decision, it decides how much of this to spend on food, clothing, housing, etc. Thus, total expenditure per period is a decision variable and could be endogenous. In the standard model where error terms arise due to future uncertainty only, total expenditure is exogenous to the share equations. However, introducing random preferences in a life-cycle consistent way will lead to a model in which the resulting error term is correlated with total expenditure and hence total expenditure is endogenous.

To the best of our knowledge, practically feasible semiparametric estimators of the model allowing for endogenous regressors in the binary choice selection equation are not available yet. We shall therefore assume that the log of total expenditure and its square are not present in the selection equation. Instead, this equation includes the log of household income and its square, which can be seen as instruments for the total expenditure variables.

We decompose x_i into x_{ai}, containing log total expenditure and its square, x_{bi}, containing log household income and its square, and x_{di}, containing the taste shifters. x_{ci} is a subvector of x_{di}. x_{bi} and the part of x_{di} that is not in x_{ci} are excluded from the budget share equations, while x_{ai} is excluded from the selection equation. The random effects assumption implies that we assume that the error terms $\alpha_{0i}+\varepsilon_{0i}$, $\alpha_{1i}+\varepsilon_{1i}$ and η_i-u_i are independent of $(x_{bi}',x_{di}')'$.

A detailed analysis of various cross-section models is given in Charlier et al. (1996). Since in that paper, the normality assumption on $(\alpha_{0i}+\varepsilon_{0i},\alpha_{1i}+\varepsilon_{1i},\eta_i-u_i)'$ is strongly rejected, we here only report the results based on the approach of Newey (1988). This yields consistent estimators under weaker distributional assumptions than normality, and also has the advantage of computational convenience. Newey's approach consists of two steps. The first step is to estimate the binary choice selection equation. In our search for a flexible enough specification for this, we have experimented with several generalizations of a probit model, and found that the following one-parameter extension of the probit model performs well:

$$P\{d_i=1 \mid x_{bi},x_{di}\} = \Phi(\pi_b'x_{bi}+\pi_d'x_{di}+\tau[\pi_b'x_{bi}+\pi_d'x_{di}]^2).$$

This binary choice single index model is estimated by ML.

The second step is to estimate the budget share equations, taking account of selectivity bias and potential endogeneity of expenditure variables. Selection is accounted for by adding an additional regressor which can be seen as a correction term. This correction term is an unknown function of the single index $\pi_b'x_{bi}+\pi_d'x_{di}$ in the selection equation. The unknown function is replaced by a

[7] See Blundell and Walker (1986), for example.

polynomial with coefficients to be estimated, and the parameters π_b and π_d are replaced by their first round estimates. Newey shows that, for the case of exogenous regressors, OLS on the respective subsamples with the terms of the polynomials added as additional regressors, leads to consistent estimates if the order of the polynomial tends to infinity with the number of observations. He also derives the asymptotic covariance matrix of the estimator and a consistent estimate for it.

Potential endogeneity of x_{ai} can be accounted for using IV (with x_{ci} and log family income and its square as instruments) instead of OLS in the second step. This is all described extensively in Charlier et al. (1996). To make the current paper self-contained, we have also included the details of this estimator and its implementation (choice of smoothing parameters, etc.) in Appendix B.

Results for the wave of 1987 are presented in Table 2. In the lower panel are the ML estimates of the selection equation as specified above. τ, the coefficient of $(\pi_b'x_{bi}+\pi_d'x_{di})^2$, is significantly negative, but the probability $P\{d_i=1 \mid x_{bi},x_{di}\}$ increases with the index $\pi_b'x_{bi}+\pi_d'x_{di}$ over the sample range.[8] The income pattern is U-shaped, and the probability of ownership increases with income over most of the income range. The education effect is also positive, and much stronger and significant than the income effect. The age pattern is inversely U-shaped with a maximum probability of ownership at about 47 years. Being married increases the probability of ownership, the number of children is insignificant. The regional dummies imply that ownership is higher in other regions than in the west of the country, where house prices are higher than elsewhere.

The semiparametric estimates based upon Newey (1988) for the case that LEXP and L2EXP are assumed to be exogenous, are presented in the second column of (the upper part of) table 2. In the series approximation of the correction term, six terms were used for owners and four for renters. These choices resulted from estimating models with up to nine terms; the estimates did not change much after including more than six and four terms, respectively.

The estimated standard errors, which take into account the first stage estimation error in the parameters of the selection equation, appear to differ substantially from the standard OLS standard error estimates, but are similar to the Eicker-White standard errors. This indicates that the first stage errors hardly affect the standard errors of the second stage estimates.

We present Newey instrumental variables (IV) estimates, allowing for endogeneity of LEXP and

[8] Using LM tests similar to those in Chesher and Irish (1987), normality in this extended probit model could not be rejected. Homoskedasticity, however, is still rejected, suggesting that the single index specification might be inadequate. Due to the lack of feasible alternatives, however, we have to retain this assumption (see also previous footnote).

L2EXP in the budget share equation, in the fourth column of table 2.[9] We used series approximations of six terms for owners and five terms for renters. Using IV instead of OLS mainly affects the parameter estimates related to LEXP and L2EXP. A Hausman type test on exogeneity of LEXP and L2EXP is based on the difference between the share equation estimates in table 2. The realization of the test statistic is 1.2 for owners and 12.9 for renters. Both are below the critical value of a χ^2_8 distribution for any conventional significance level, so that exogeneity of LEXP and L2EXP cannot be rejected.

The age terms are insignificant in both equations for both estimators. The regional dummies suggest that housing costs are lower in the north than in the rest of the country. Marital status is insignificant for owners. The only substantial difference between IV and OLS estimates is that marital status is significantly negative for renters according to the former, and virtually zero according to the latter.

The estimated shares spent on housing as a function of LEXP are presented as dotted curves in graphs 1 and 3 in figure 3. The other explanatory variables are set to their sample means. The constant terms are not estimated; they are chosen such that the means of the predicted budget shares equal the means of the observed budget shares. Therefore, only the shapes of the curves can be compared, and not their level. For both owners and renters, we find that allowing for endogeneity of total expenditure makes a big difference for high levels of total expenditure.

For each panel wave implied elasticities of housing expenditure with respect to total expenditure are presented in table 5. We present means of these elasticities for owners and renters separately, weighted with total household expenditure. These can be interpreted as aggregate elasticities (cf. Banks et al. (1994)). We present the means and their standard errors, and the fraction of households for which the elasticity estimate is larger than zero.[10] In all cases, the elasticities are much smaller than one, suggesting that housing is a necessity. The standard errors are often quite large, so that the means are insignificantly different from zero.

To see whether the negative sign for the elasticity in the Newey IV model is caused by an inappropriate choice of the instruments, we also replaced the instruments by the lagged values of log(household income) and its square. This, however, led to similar parameter estimates as before and the elasticities for renters increased only slightly.

3.2. Fixed effects

[9] Results in Appendix A show that the results of the Newey (1988) estimates are not sensitive with respect to the definition of the expenditure measure for owners.

[10] The median elasticities (not reported), were very close to zero in all cases.

Using more than one wave for estimation requires that we explicitly include the time period in the notation. As in the previous model, we decompose x_{it} into x_{ait}, containing log total expenditure and its square, x_{bit}, containing log household income and its square, and x_{dit}, containing the taste shifters. x_{cit} again is a subvector of x_{dit}. x_{bit} as well as the part of x_{dit} that is not in x_{cit} are excluded from the budget share equations, while x_{ait} is excluded from the selection equation. We allow for correlation between the household specific effects and $(x'_{ait},x'_{cit})'$. Throughout, we assume strict exogeneity of x_{bit} and x_{dit}, i.e., $\{(\varepsilon_{0it},\varepsilon_{1it},u_{it}),\ t=1,...,T\}$ is independent of $\{(x_{bit},x_{dit}),t=1,...,T\}$. Estimation can be based on taking differences between periods t and τ, $t\neq\tau$. This yields, for households with $d_{it}=d_{i\tau}$

$$y_{pit}-y_{pi\tau} = \beta'_{pa}(x_{ait}-x_{ai\tau}) + \beta'_{pc}(x_{cit}-x_{ci\tau})+ (\varepsilon_{pit}-\varepsilon_{pi\tau}) \quad \text{if } d_{it}=d_{i\tau}=p,\ p=0,1,$$

with

$$d_{is} = 1(\pi'_b x_{bis} + \pi'_d x_{dis} + \eta_i - u_{is} \geq 0),\ s=t,\tau.$$

Thus, if $d_{it}=d_{i\tau}=p$, p=0,1, we can write

$$y_{pit}-y_{pi\tau} = \beta'_{pa}(x_{ait}-x_{ai\tau}) + \beta'_{pc}(x_{cit}-x_{ci\tau}) + g_{pt\tau}(x_{bit},x_{bi\tau},x_{dit},x_{di\tau}) + \tilde{\varepsilon}_{pit\tau}$$

where the functions $g_{pt\tau}$, p=0,1, are given by

$$g_{pt\tau}(x_{bit},x_{bi\tau},x_{dit},x_{di\tau})=E\{\varepsilon_{pit}-\varepsilon_{pi\tau}|x_{bit},\ x_{bi\tau},\ x_{dit},\ x_{di\tau},\ d_{it}=d_{i\tau}=p\}.$$

and where $\tilde{\varepsilon}_{pit\tau}$ satisfies

$$E\{\tilde{\varepsilon}_{pit\tau}|x_{bit},\ x_{bi\tau},\ x_{dit},\ x_{di\tau},\ d_{it}=d_{i\tau}=p\} = 0,\ p=0,1.$$

The assumptions with respect to the error terms $(\varepsilon_{0it},\varepsilon_{1it},u_{it})$ determine the functions $g_{0t\tau}$ and $g_{1t\tau}$ and the way to estimate the parameters. We discuss the two that will be applied.

(i) Linear Panel Data Model

If we assume that no selection bias is present after differencing, i.e., $g_{pt\tau}\equiv 0$, p=0,1, standard panel data estimation procedures can be used. In this case there is no reason to estimate the auxiliary selection equation. Only the budget share equations need to be estimated. This corresponds to the assumption that η_i-u_{it} is independent of ε_{0it} and ε_{1it}, for all t, implying that possible selection

effects on the budget share equations only enter through correlation between α_i and $(\eta_i,u_{i1},..,u_{it})$. This assumption is often used in applications, for example, by Pedersen et al. (1990) in a model for wage differentials between public and private sector.

Estimation results for the linear panel data model estimator are presented in table 3, both under the assumption that LEXP and L2EXP are exogenous (OLS), and allowing for their endogeneity (IV). A Hausman type test comparing these two leads to rejecting exogeneity for renters but not for owners. The only significant variables are LEXP and its square. Graphs of the budget share as a function of LEXP are presented as solid curves in figure 3. Not only the other observed characteristics are fixed, but also the unobserved household specific effects. We chose them in such a way that the average shares for owners and renters equal the observed sample means. For owners the difference between the curves for exogenous and endogenous LEXP and L2EXP are substantial. For renters, the two curves are more similar to each other.

Elasticities of housing expenditure with respect to total expenditure can be calculated in the same way as in the random effects panel data model. These elasticities are now not only conditional upon the exogenous variables and the choice between renting and owning, but also on the household specific fixed effects. We calculated the aggregate elasticities (weighted with total expenditure) for each panel wave. The results are presented in table 5. For owners the aggregate elasticity is significantly positive when LEXP and L2EXP are treated as exogenous, but insignificant if endogeneity is allowed for. The latter conclusion also holds for renters. In general, the elasticities are close to zero. Comparing the results with the ones for the random effects model, the fairly large standard errors do not allow strong conclusions concerning differences in sign or magnitude. The main difference occurs when LEXP and L2EXP are assumed to be exogenous: the results for owners in the linear panel data model are significant but insignificant in the random effects model.

(ii) Semiparametric model

For a panel with two time periods Kyriazidou (1995) proposes an estimator requiring weaker assumptions than those in the model discussed above. The main assumption in her paper is the exchangeability of the error terms. For the share equation of owners, this means that, conditional on the household specific effects, $(\varepsilon_{1it},\varepsilon_{1i\tau},u_{it},u_{i\tau})$ and $(\varepsilon_{1i\tau},\varepsilon_{1it},u_{i\tau},u_{it})$ are identically distributed. It implies that for households for which $d_{it}=d_{i\tau}$ and $\pi_b'x_{bit}+\pi_d'x_{dit}=\pi_b'x_{bi\tau}+\pi_d'x_{di\tau}$, the effect of selection on the budget share equation (i.e., the g-functions) is the same in periods t and τ. For such observations, differencing will not only eliminate the fixed effect, but also the selection effect. Note the difference with the linear model introduced above, where we could use all the

observations, since the assumptions implied that correction terms were zero. Now, we only use that the correction terms are the same for certain observations. The subsample consisting of these observations is used for estimation.

Since observations with $\pi_b'x_{bit}+\pi_d'x_{dit} = \pi_b'x_{bi\tau}+\pi_d'x_{di\tau}$ are scarce, all observations for which the difference between these two values is sufficiently close to zero are used. This leads to weighted IV or weighted LS estimators for $(\beta_{0a}',\beta_{0c}')'$ and $(\beta_{1a}',\beta_{1c}')'$. We present the IV estimation procedure for the owners' share equation; the procedure applied to the other cases is very similar.

Denote the regressors in the budget share equations by $\tilde{x}_{it}=(x_{ait}',x_{cit}')'$, and the corresponding instruments by $w_{it}=(x_{bit}',x_{cit}')'$ (of the same dimension as $\tilde{x}_{it}$). Let

$$\hat{S}_{wx} = \sum_{i=1}^{n} \hat{\omega}_i (w_{i\tau} - w_{it}) (\tilde{x}_{i\tau} - \tilde{x}_{it})' d_{it} d_{i\tau}$$

$$\hat{S}_{wy1} = \sum_{i=1}^{n} \hat{\omega}_i (w_{i\tau} - w_{it}) (y_{1i\tau} - y_{1it}) d_{it} d_{i\tau}$$

$$\hat{\omega}_i = \frac{1}{s_{1n}} K\left(\frac{\hat{\pi}_b'(x_{bi\tau} - x_{bit}) + \hat{\pi}_d'(x_{di\tau} - x_{dit})}{s_{1n}} \right)$$

where $\hat{\pi}_b$ and $\hat{\pi}_d$ are estimates of π_b and π_d (to be discussed below), and K is a kernel with bandwidth satisfying $s_{1n}\to 0$ as $n\to\infty$. Then the IV estimator for $(\beta_{1a}',\beta_{1c}')'$ is $\hat{S}_{wx}^{-1}\hat{S}_{wy1}$. The estimator is asymptotically normal with an asymptotic bias and an asymptotic covariance matrix that can be estimated consistently. The rate of convergence is $(ns_{1n})^{½}$.

We use the standard normal density function for the kernel. In choosing the bandwidth, we used the plug-in procedure as described by Horowitz (1992): first, some initial value for the bandwidth is chosen and the parameter estimates, the estimate of the asymptotic bias and the estimate of the covariance matrix are computed. These estimates are used to compute the MSE minimizing bandwidth and then the bias and the covariance matrix are re-estimated.[11]

The approach for two time periods can easily be generalized to the case of more than two time periods. Given some estimates for the selection equation, the budget share equations can be estimated using the IV approach for each combination of panel waves (t,τ). Minimum Distance, preferably with the optimal weighting matrix, can then be applied to combine these estimates. Details can be found in Appendix C. To estimate the optimal weighting matrix, an estimate for the covariance matrix of the estimators for the different time periods is required. These covariances converge to zero due to the fact that the bandwidth tends to zero. The proof is similar to that in Charlier (1994) and is included in Appendix C. The Minimum Distance estimator is therefore a

[11] Details on this procedure are available upon request from the authors.

weighted average of the estimators for each pair (t,τ), $t\neq\tau$, with weights given by the inverse of the corresponding covariance matrix estimate.

The above estimator requires a first stage estimator $(\hat{\pi}_b',\hat{\pi}_d')'$ for $(\pi_b',\pi_d')'$. This can be, for instance, smoothed maximum score (see Charlier et al., 1995 or Kyriazidou, 1995) or conditional logit, depending on the distributional assumptions for the selection equation. Kyriazidou proposes to use smoothed maximum score. Both estimators only use transitions from owning to renting and from renting to owning. Such transitions are scarce in our data, however. Consequently, it is impossible to estimate a very flexible specification. Therefore, we will impose the stronger assumptions that the u_{it} $(t=1,...,T)$ are iid with a logistic distribution, and use the conditional logit ML estimator to estimate the selection part of the model (see Chamberlain, 1980). Since this estimator for $(\pi_b',\pi_d')'$ converges at a faster rate than those for $(\beta_{1a}',\beta_{1c}')'$ and $(\beta_{0a}',\beta_{0c}')'$, the former will not affect the limit distribution of the latter. This is similar to the result in Kyriazidou (1995).

In order to retain as many observations as possible, we extend the conditional logit estimator to the case of unbalanced panels. Let $c_i=(c_{i1},..,c_{it})$ denote a vector of zeros and ones, with $c_{it}=1$ indicating that all the variables are observed for household i in time period t. Assuming independence between y_i and c_i conditional on x_i, it is easy to show that c_i can be treated as exogenous. The conditional likelihood contribution of an observation then only depends on observed values of (y_{it},x_{it}).

We estimate the selection equation using the unbalanced panel for the years 1987 till 1989, consisting of 4089 households, with 2348 present in all three years, 943 in two years, and 798 in only one year.[12] This leads to 3065, 3276 and 3387 observations in the three waves. Important for the precision of the estimates, however, is the number of households that switch at least once between the two states renting and owning. This number is 170.

In the fixed effects logit model, only the coefficients corresponding to the time varying regressors are identified. This implies that, due to little or no time variation in these variables, the constant term and the parameters related to the education dummies, the dummy for being married, and the regional dummies cannot be estimated. Only the parameters of AGE, AGE2, LINC, L2INC and NCH remain. We supplemented the equation with time dummies for each of the three years, two of which can be estimated; the coefficient for the dummy for 1989 is normalized to zero.

The results are presented in the second column of table 4. The estimates for the time dummies show that the ownership rate increases over time, *ceteris paribus*. The age variables imply an inversely U-shaped pattern of the probability of owning similar to that in table 2. The coefficients

[12] Since total expenditure does not play a role in the selection equation, observations with missing information on total expenditure were also used.

related to LINC and L2INC are jointly insignificant. Excluding L2INC still leads to an insignificant parameter estimate for LINC. This result is different from that for the random effect panel data model, where income had a positive impact on the probability of home ownership. That finding was probably due to the positive relation between permanent income and home ownership. In the fixed effects model, permanent income is part of the fixed effect, and the interpretation of our result is that transitory income components do not affect the home ownership decision significantly. This makes sense in a life cycle context.

The other columns of table 4 contain the minimum distance estimates and their standard error estimates for owners and renters.[13] The bias in the first step Kyriazidou estimates was generally large for AGE, AGE2 and the time dummy whereas it was small for LEXP and L2EXP (±4% of the parameter estimates) using 87/88 or 88/89 in estimation. However, the bias for these parameters was a lot larger for 87/89 (±30%). The parameters related to AGE, AGE2, LEXP and L2EXP are substantially different from their random effects counterparts based on IV. For renters, the age variables are insignificant. The coefficients of LEXP and L2EXP are strongly significant. They imply that, *ceteris paribus*, the budget share spent on housing negatively responds to a change in total expenditure. For owners the main difference between the two estimates are the estimates for LEXP and L2EXP as well as the significance of the time dummy for 1987 when endogeneity of LEXP and L2EXP is taken into account. For renters the same remark applies, but also the time dummy for 1988 is significant and the age pattern changes.

To test the assumption of no selectivity bias in the linear panel data model, we perform a Hausman type test comparing the IV parameter estimates in tables 3 and 4. Because the Kyriazidou estimator converges slower than the linear panel data estimator, the limit distribution of the difference between the estimators is determined by the limit distribution of the Kyriazidou estimator only. The resulting values for the test statistics are 138.3 for owners and 2881.3 for renters. Both are larger than the critical values of the χ^2_6 at any conventional significance level. This indicates that the model that does not allow for correlation between the error terms in the share equations and the error term or fixed effect in the selection equation is misspecified.

To test the assumption of no correlation between the household specific effects and $(x'_{bi},x'_{di})'$ we perform a Hausman type test based on the difference between the Newey IV and the Kyriazidou IV estimates for those explanatory variables present in both estimates (AGE, AGE2, LEXP and L2EXP). The limit distribution of the difference between the estimators is again determined by the

[13] Results in Appendix A show that most parameters tend to change slightly but not significantly with the different definitions for housing expenditure for owners.

limit distribution of the Kyriazidou estimator only. The resulting values for the test statistics are 117.9 for owners and 28.9 for renters. For owners this is larger than the critical values of the χ^2_4 at any conventional significance level. This indicates that the random effects panel data model that does not allow for correlation between the household specific effects and the explanatory variables is misspecified. This result continues to hold when we compare the estimates for owners and renters simultaneously.

To test whether the model could be simplified to a model with one budget share equation instead of separate equations for renters and owners, we use a Wald test to check whether β_1 is equal to β_0. Because T=3, no household can both own a house for two periods or more and rent a house for two periods or more. As a consequence, the covariance between the estimates for β_0 and β_1 in table 4 is zero, which makes it straightforward to perform the Wald test. The value of the test statistic is 31.27 which exceeds the critical value of the χ^2_6 distribution at all conventional significance levels. This implies that the model cannot be simplified in this direction.

Graphs of the budget share spent on housing according to the Kyriazidou model are presented as dashed lines in figure 3. In general most curves are again decreasing except for the very high levels of total expenditure. For owners the curves based upon estimates allowing and not allowing for endogeneity of LEXP and L2EXP are very similar. They are also similar to the curve for the linear panel data model with exogenous LEXP and L2EXP. For renters, the curve allowing for endogeneity of LEXP and L2EXP in the Kyriazidou model alters the shape of the curve and makes it closer to linear. For owners we also present the curves for alternative definitions of housing expenditure (BS12, BS10, see Appendix A) in graph 2. The curves for the different definitions of housing expenditure for owners are similar.

In table 5 we present the weighted elasticity estimates for the Kyriazidou model, i.e. the aggregate elasticities of housing expenditure with respect to total expenditure. For owners the results are similar to those in the linear panel data model: elasticity estimates are significantly positive under exogeneity of LEXP and L2EXP, and insignificant when LEXP and L2EXP are endogenous. For renters the elasticity estimates change substantially compared to those in the linear panel data model. Compared to the random effects model the results change substantially both for owners and for renters. For the Kyriazidou model the estimated elasticities have the wrong sign under endogeneity. To see whether the negative sign is due to an inappropriate choice of instruments, we also replaced the instruments by the lagged values of log(household income) and its square.[14] Although the parameter estimates changed, the elasticity estimates remained negative

[14] We used the balanced panel. Due to the extra time lag in the instruments we can only compute the estimates for the 1988 and 1989 waves of the panel so no minimum distance step is required.

and they became significant. Therefore the choice of current income variables as instruments does not seem to explain the negative sign for the elasticities.

Again, the importance of the fixed effects for the interpretation of these results should be emphasized. Permanent income effects enter through the fixed effect, and can still be positive. The estimates imply that for renters transitory shocks on total expenditure are more likely to be negatively correlated to changes in housing expenditure. We have no economic explanation for this.

The final panel of table 5 contains the results for the elasticities for different measures of housing expenditure for owners (see Appendix A). The elasticities as well as the standard errors are slightly affected.

Finally, we performed a specification test on the Kyriazidou model. A natural approach here is to perform a test on overidentifying restrictions in the minimum distance step. However, as discussed in Appendix C, we have to choose smoothing parameters. When choosing the smoothing parameters as in Appendix C, the realizations of the test statistics are 26.29 for renters and 27.56 for owners, which both exceed the critical value of a χ^2_9 distribution at conventional significance levels.[15] The choice of smoothing parameters we employ, corresponds to setting weights for the first step estimates equal to one. However, these weights that depend on n, the first round smoothing parameters, and the minimum distance smoothing parameters, only have to converge to one for the sample size approaching infinity. Small changes in the smoothing parameters do not affect the conclusion of misspecification, but more substantial changes in the weights (say, weight 1.5 instead of 1) yield as conclusion that the null of no misspecification cannot be rejected.[16]

4. Conclusions

We have modelled expenditure on housing for owners and renters using endogenous switching regression models for panel data. Attention was paid to the construction of the variables needed in the econometric model, especially to the definition of housing expenditure for owners. In choosing the model assumptions we were guided by economic theory, but to a large extent also by the availability of suitable estimators and the nature of the data. We extended the standard switching regression model in several directions. First, we used (unbalanced) panel data instead of cross-

[15] There are 6 parameters to be estimated. Using one pair of waves, only the difference of the two corresponding time dummies is identified. Therefore we have 4x3+3=15 constraints in the minimum distance step. This yields 15−6=9 degrees of freedom.

[16] The same sensitivity analysis can also be performed in the tests comparing the random effects IV model or the linear panel data IV model with the Kyriazidou IV model, and the Wald test comparing the estimates for β_1 and β_0. In all these cases, the sensitivity analysis yields similar results.

section data, and considered random effects and fixed effects models. For the random effects case, cross-section models and data can be used to obtain consistent estimates, but the fixed effects case requires different techniques. We used two of them, allowing for different types of selection effects. Where possible, we tried to avoid normality assumptions and relied on semiparametric techniques. Finally, we focused on estimation techniques which allow some of the explanatory variables in the budget share equations to be endogenous.

We estimated the slope coefficients in the random effects model using the cross-section data for 1987 on the basis of a semiparametric model. We have compared results which do and do not take account of potential endogeneity of the variables related to total expenditure. Differences between these two sets of estimates mainly concern the parameter estimates related to the total expenditure variables themselves.

For the fixed effects panel data case we estimated two models. The first one is the linear panel data model which can be estimated using standard estimation techniques. The alternative estimator based on weaker assumptions was proposed by Kyriazidou (1995). Here the parameters in the selection equation were estimated using conditional logit. The parameters in the budget share equations are estimated in a second step, making use of the conditional logit estimates. The models were compared using Hausman type tests. The results indicate that both the random effects and the linear panel data model are too restrictive. Exogeneity of total expenditure variables is not always rejected. Finally, we also applied a test on overidentifying restrictions in the Kyriazidou (1995) model, the most general model that we considered. The results suggest that an even more general model might yield better results.

References

Ahn, H. and J.L. Powell (1993), "Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism," *Journal of Econometrics*, 58 (1/2), 3-29.

Amemiya, T. (1984), "Tobit models: a survey," *Journal of Econometrics*, 24, 3-63.

Andrews, D. and M. Schafgans (1995), "Semiparametric estimation of a sample selection model," mimeo, Yale University.

Banks, J., R. Blundell and A. Lewbel (1994), "Quadratic Engel Curves, Indirect tax Reform and Welfare Measurement," University College London Discussion Paper 94-04.

Blundell, R. and I. Walker (1986), "A Life-Cycle Consistent Empirical Model of Labour Supply Using Cross-Section Data," *Review of Economic Studies*, 53, 539-558.

Budgethandboek NIBUD (1987), "*Gegevens Omtrent Inkomsten, Uitgaven en Bestedingspatronen van Particuliere Huishoudens*," NIBUD.

Camphuis, H. (1993), "Checking, Editing and Imputation of Wealth Data of the Netherlands Socio-Economic Panel for the period '87-'89," VSB-CentER Savings Project Discussion paper.

CBS (1987), "*Woningbehoeftenonderzoek* 1985/1986," CBS, The Hague.

Chamberlain, G. (1980), "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47, 225-238.

Charlier, E. (1994), "A Smoothed Maximum Score Estimator for the Binary Choice Panel Data Model with Individual Fixed Effects and Application to Labour Force Participation," CentER Discussion Paper no. 9481, Tilburg University.

Charlier, E., B. Melenberg and A.H.O. van Soest (1996),"An Analysis of Housing Expenditure using Semiparametric Cross-Section Models", mimeo, Tilburg University.

Chesher, A. and M. Irish (1987), "Residual Analysis in the Grouped and Censored Normal Linear Model," *Journal of Econometrics*, 34, 33-61.

Euwals, R. and A.H.O. van Soest (1995), "Desired and Actual Family Labour Supply in the Netherlands," CentER Discussion Paper no. 9623, Tilburg University.

Haurin, D.R. (1991), "Income Variability, Homeownership, and Housing Demand," *Journal of Housing Economics*, 1, 60-74.

Horowitz, J. (1992), "A Smoothed Maximum Score Estimator for the Binary Choice Response Model," *Econometrica*, 60, 505−531.

Ioannides, Y.M. and S.S. Rosental (1994), "Estimating the consumption and investment demands for housing and their effect on housing tenure status," *The Review of Economics and Statistics*, 76, 127-141.

Klein, R.W. and R.S. Spady, (1993), "An Efficient Semiparametric Estimator of the Binary Response Model," *Econometrica*, 61, 387-423.

Kyriazidou, E. (1995), "Estimation of a Panel Data Sample Selection Model," mimeo, Yale University.

Lee, L.F. and R.P. Trost (1978), "Estimation of Some Limited Dependent Variable Models with Application to Housing Demand," *Journal of Econometrics*, 8, 357-382.

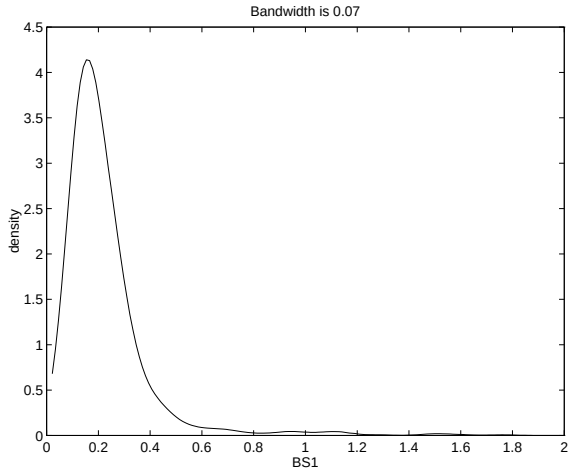
Newey, W.K. (1988), "Two Step Series Estimation of Sample Selection Models," mimeo, MIT (revised version October 1991).

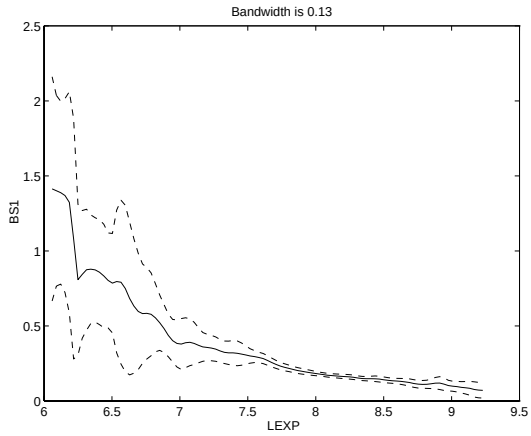
Pedersen, P.J., J.B. Schmidt-Sørensen, N. Smith and N. Westergård-Nielsen (1990), "Wage Differentials Between the Public and Private Sectors," *Journal of Public Economics*, 41, 125-145.

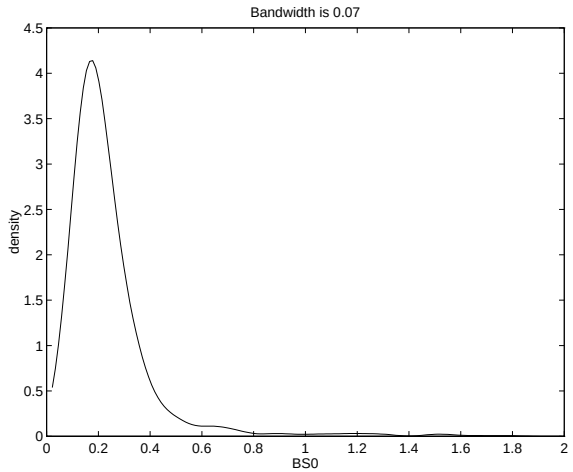
Vella, F. and M. Verbeek (1994), "Two-Step Estimation of Simultaneous Equation Panel Data Models with Censored Endogenous Variables," CentER Discussion Paper No. 9455.

Zorn, P.M. (1993), "The Impact of Mortgage Qualification Criteria on Households' Housing Decisions: An Empirical Analysis Using Microeconomic Data," *Journal of Housing Economics*, 3, 51-75.

Figure 1: *Nonparametric density estimates for BS1 and BS0 and nonparametric regression estimates of the same variables on log total expenditure (LEXP), together with 95% uniform confidence bands*







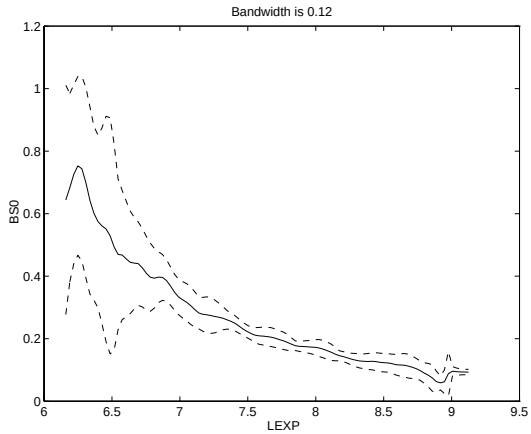
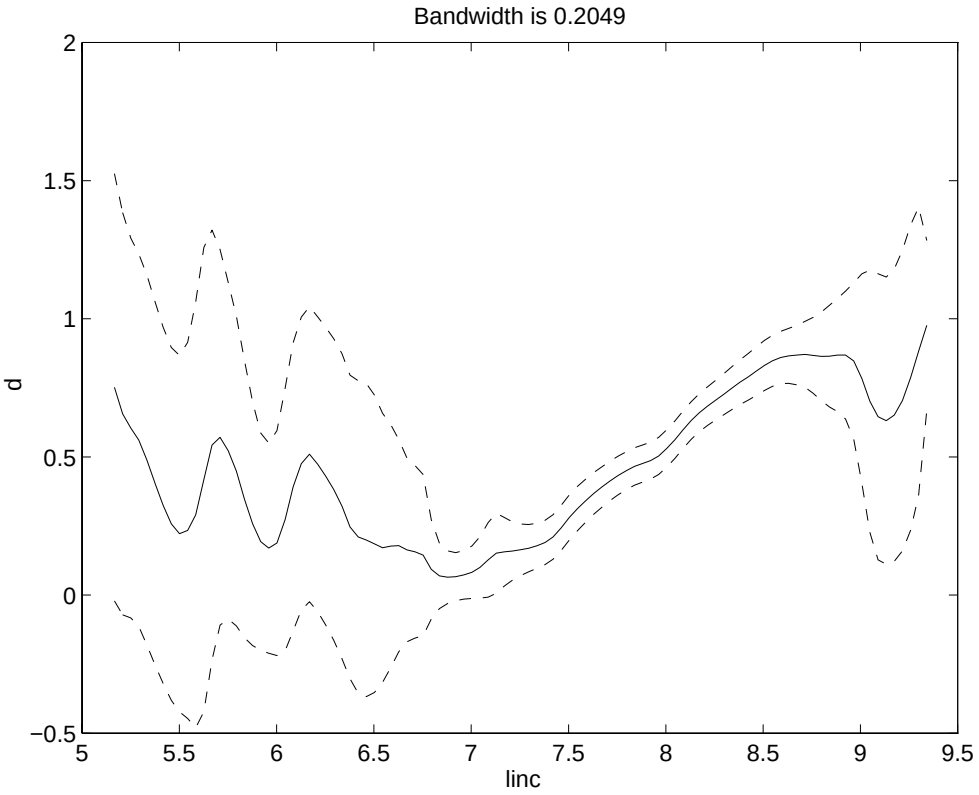


Figure 2: *Nonparametric estimates of the probability of owning a house as a function of log household income (LINC), and distribution of LINC*



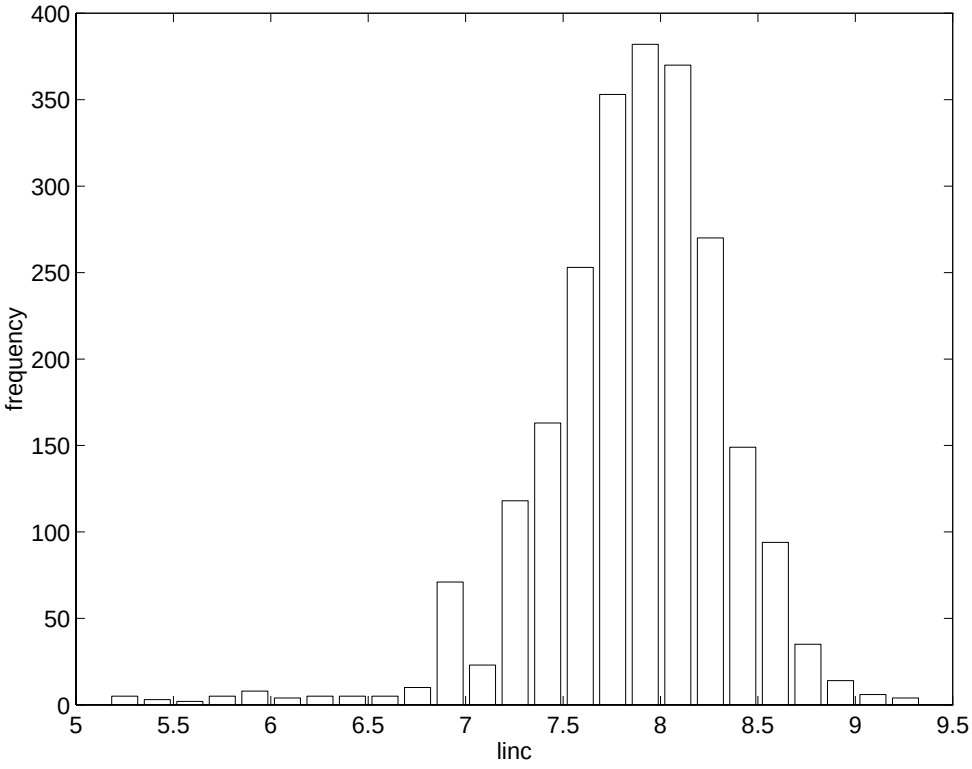
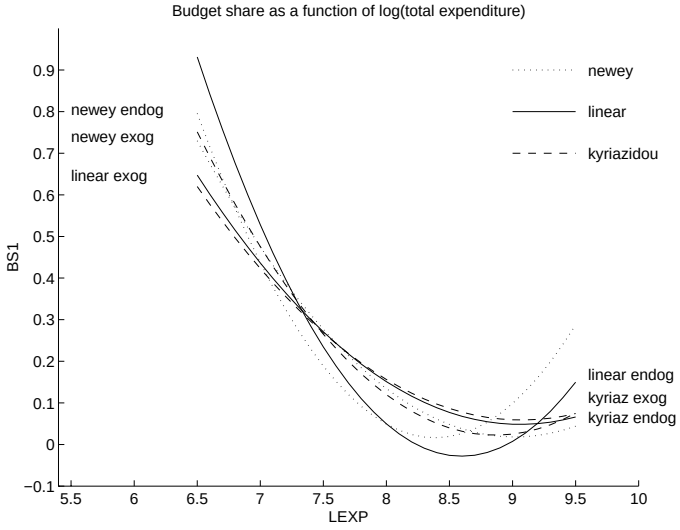
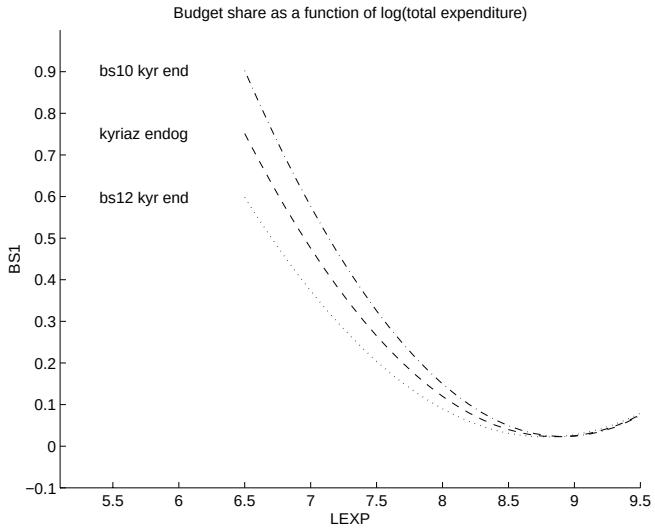


Figure 3: Budget share spent on housing as a function of LEXP for the panel data models

Owners:





Renters:

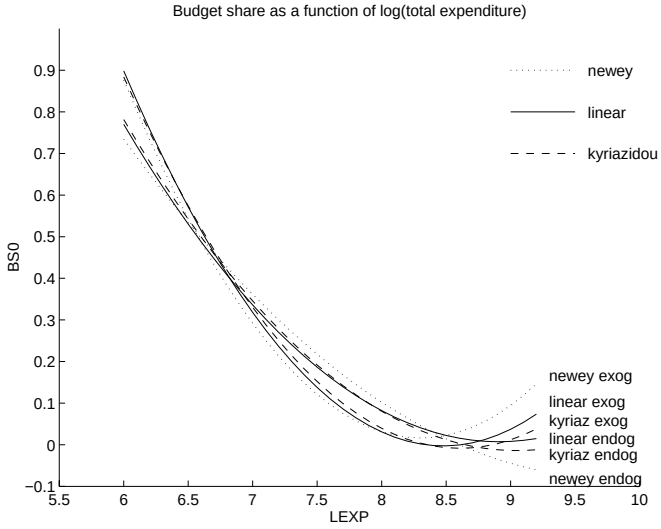


Table 1. Overview of variables and summary statistics for 1987, 1988 and 1989 (standard errors in parentheses)

Variable	Description	Mean	Renters		Mean	Owners	
	year	1987	1988	1989	1987	1988	1989
	number of obs.	1190	1235	1191	1167	1235	1278
BS0, BS1	Budget share (i.e. monthly expenditure on housing divided by monthly total expenditure)	0.24 (0.23)	0.23 (0.17)	0.23 (0.18)	0.22 (0.18)	0.22 (0.19)	0.21 (0.19)
DOP2	dummies for education level	0.24	0.26	0.26	0.15	0.16	0.17
DOP3		0.37	0.36	0.38	0.48	0.44	0.47
DOP4		0.09	0.08	0.11	0.19	0.18	0.20
DOP5		0.03	0.03	0.03	0.07	0.06	0.07
AGE	age of the head of the household in decennia	4.03 (1.21)	3.94 (1.22)	3.97 (1.23)	4.10 (0.98)	4.10 (0.96)	4.11 (0.97)
AGE2	and its square	17.64	16.98	17.24	17.78	17.70	17.83
LINC	logarithm of monthly family income and	7.69 (0.47)	7.70 (0.48)	7.73 (0.48)	8.04 (0.44)	8.05 (0.47)	8.13 (0.42)
L2INC	its square (in guilders)	59.34	59.45	59.96	64.90	65.09	66.24
EXP	monthly total family expenditure	2304 (1117)	2422 (1139)	2477 (1223)	3233 (1483)	3440 (1702)	3606 (1737)
LEXP	logarithm of monthly total family expenditure	7.62 (0.54)	7.67 (0.53)	7.68 (0.56)	7.97 (0.51)	8.02 (0.54)	8.08 (0.50)
L2EXP	and its square	58.56	59.11	59.31	63.73	64.55	65.49
DMAR	dummy for married	0.73	0.70	0.68	0.94	0.93	0.93
NCH	number of children living with the family	0.83	0.80	0.77	1.22	1.22	1.15
DREG1	region dummies for north, east and south respectively	0.11	0.11	0.12	0.11	0.10	0.11
DREG2		0.20	0.20	0.20	0.22	0.24	0.23
DREG3		0.23	0.24	0.24	0.28	0.27	0.27

Table 2. Estimation results for the random effects panel data model using BS1 for owners and BS0 for renters (standard errors in parentheses)[a]

Variable	Newey[b]		Newey IV[c]	
BS owners				
CONSTANT	9.234[d]		15.454[d]	
AGE	−0.027	(0.033)	0.023	(0.066)
AGE2	0.002	(0.004)	−0.005	(0.008)
LEXP	−2.024**	(0.336)	−3.670	(1.939)
L2EXP	0.112**	(0.021)	0.212	(0.121)
DMAR	0.024	(0.022)	0.041	(0.024)
DREG1	−0.050**	(0.011)	−0.037*	(0.016)
DREG2	−0.011	(0.011)	−0.003	(0.012)
DREG3	−0.020	(0.010)	−0.009	(0.014)
BS renters				
CONSTANT	11.290[e]		5.589[e]	
AGE	0.017	(0.036)	−0.068	(0.042)
AGE2	−0.002	(0.004)	0.008	(0.004)
LEXP	−2.690**	(0.324)	−1.107*	(0.469)
L2EXP	0.162**	(0.021)	0.056	(0.032)
DMAR	−0.002	(0.016)	−0.049*	(0.018)
DREG1	−0.028*	(0.011)	−0.038*	(0.016)
DREG2	−0.015	(0.012)	−0.024	(0.015)
DREG3	−0.003	(0.011)	−0.003	(0.013)
Selection				
CONSTANT	23.113**	(4.440)	«	
DOP2	0.207*	(0.083)	«	
DOP3	0.510**	(0.075)	«	
DOP4	0.599**	(0.121)	«	
DOP5	0.570**	(0.208)	«	
AGE	1.252**	(0.223)	«	
AGE2	−0.134**	(0.026)	«	
LINC	−7.991**	(1.193)	«	
L2INC	0.581**	(0.080)	«	
DMAR	0.481**	(0.088)	«	
NCH	0.051	(0.033)	«	
DREG1	0.303**	(0.094)	«	
DREG2	0.240**	(0.078)	«	
DREG3	0.302**	(0.073)	«	
τ_2	−0.172**	(0.036)	«	

[a] * means significant at the 5% level, ** means significant at the 1% level.
[b] series approximation using single index ML probit in estimating the selection equation.
[c] IV using AGE, AGE2, LINC, L2INC, DMAR, DREG1, DREG2, DREG3 as instruments
[d] estimates include the estimate for the constant term in the series approximation

Table 3: Estimation results based on the linear panel data model using the unbalanced panel, standard errors in parentheses

Equation	Variable	OLS Estimates	IV[a] Estimates
BS1 owners	AGE	0.041 (0.073)	−0.082 (0.122)
	AGE2	0.005 (0.008)	0.026 (0.013)
	LEXP	−1.655** (0.051)	−3.750** (0.454)
	L2EXP	0.091** (0.003)	0.218** (0.028)
	Dummy87	0.009 (0.006)	0.017 (0.011)
	Dummy88	0.003 (0.003)	0.009 (0.005)
BS0 renters	AGE	0.038 (0.063)	0.093 (0.072)
	AGE2	0.0000 (0.007)	−0.005 (0.008)
	LEXP	−2.487** (0.054)	−1.604** (0.192)
	L2EXP	0.147** (0.004)	0.090** (0.013)
	Dummy87	−0.0001 (0.007)	0.002 (0.007)
	Dummy88	−0.005 (0.004)	−0.005 (0.004)

[a] In IV estimation AGE, AGE2, LINC, L2INC, Dummy87 and Dummy88 are used as instruments.

Table 4: Fixed effects logit selection equation estimates and the results for the budget share equations after performing minimum distance with the optimal weighting matrix

Variable	Fixed effects logit	'OLS' Estimates	IV[a] Estimates
		owners	
AGE	16.592* (6.587)	0.158 (0.152)	0.115 (0.106)
AGE2	−1.964* (0.755)	−0.018 (0.016)	−0.012 (0.012)
LEXP		−1.538** (0.095)	−2.312** (0.205)
L2EXP		0.084** (0.006)	0.130** (0.013)
Dummy87	−2.432** (0.735)	−0.004 (0.006)	−0.009* (0.004)
Dummy88	−1.245** (0.459)	−0.0001 (0.004)	−0.001 (0.003)
LINC	7.658 (12.894)		
L2INC	-0.534 (0.813)		
NCH	−0.545 (0.481)		
		renters	
AGE		−0.283 (0.119)	0.115* (0.055)
AGE2		0.034* (0.014)	−0.012 (0.007)
LEXP		−2.262** (0.171)	−1.545** (0.123)
L2EXP		0.131** (0.011)	0.085** (0.008)
Dummy87		0.007 (0.006)	−0.013** (0.002)
Dummy88		−0.002 (0.003)	−0.007** (0.003)

[a] In IV estimation AGE, AGE2, LINC, L2INC, Dummy87 and Dummy88 are used as instruments.

Choices for initial bandwidth (s_1 and s_0) and resulting optimal bandwidths (s_1^* and s_0^*):

	exogenous LEXP and L2EXP				IV			
Year	s_1	s_1^*	s_0	s_0^*	s_1	s_1^*	s_0	s_0^*
87/88	0.5	0.54	0.5	0.44	0.6	0.60	0.4	0.41
87/89	0.3	0.38	0.3	0.33	0.7	0.91	0.5	0.50
88/89	0.5	0.50	0.6	0.60	0.5	0.50	0.6	0.61

Table 5: Budget elasticities for the panel data models (standard errors in parentheses)[a]

		owners		fr > 0	renters		fr > 0
Random Effects,	1987	−0.037	(0.101)	0.37	0.242**	(0.092)	0.55
exogenous LEXP	1988	0.010	(0.119)	0.38	0.292**	(0.103)	0.59
and L2EXP	1989	0.047	(0.128)	0.41	0.342**	(0.109)	0.62
Random Effects,	1987	0.507	(0.590)	0.62	−0.138	(0.226)	0.32
endogenous LEXP	1988	0.670	(0.718)	0.65	−0.159	(0.250)	0.31
and L2EXP	1989	0.766	(0.783)	0.68	−0.149	(0.261)	0.32
Linear, exogenous	1987	0.113**	(0.028)	0.56	0.049	(0.031)	0.43
LEXP and L2EXP	1988	0.149**	(0.031)	0.56	0.082*	(0.034)	0.46
	1989	0.179**	(0.032)	0.57	0.125**	(0.035)	0.51
Linear,	1987	0.002	(0.109)	0.35	0.014	(0.087)	0.46
endogenous LEXP	1988	0.140	(0.127)	0.36	0.016	(0.095)	0.48
and L2EXP	1989	0.228	(0.138)	0.43	0.039	(0.100)	0.50
Kyriazidou,	1987	0.166**	(0.038)	0.62	−0.024	(0.059)	0.38
exogenous LEXP	1988	0.199**	(0.042)	0.61	−0.002	(0.065)	0.42
and L2EXP	1989	0.226**	(0.044)	0.64	0.035	(0.068)	0.45
Kyriazidou,	1987	0.019	(0.057)	0.41	−0.061	(0.057)	0.36
endogenous LEXP	1988	0.083	(0.064)	0.42	−0.064	(0.062)	0.40
and L2EXP	1989	0.130*	(0.068)	0.46	−0.043	(0.064)	0.41
Kyriazidou IV							
BS12	1987	−0.018	(0.053)	0.38			
BS10	1987	0.078	(0.063)	0.45			

[a] * means significant at the 5% level, ** means significant at the 1% level.

DATA APPENDIX

In this appendix we give some details on the construction of the variables for 1986 till 1989, used in the application. Although we only use wealth from the 1986 data (to subtract from the wealth in 1987 to get savings for 1987) we include them here because the features discussed below are representative for the other years as well (unless indicated) and because we can compare the 1986 data to macro data that are not available for 1987 and 1988.

Housing

Initial dataset: 3850, 3613, 3818, 3896 households for ’86, ’87, ’88 and ’89 respectively.

Dropped from the analysis are:

- families that live for free (±0.8 % in 1986);
- families with a total income below Dfl. 1,- per month (±200 obs);
- families that receive a so called *huurgewenningsbijdrage* (i.e., a governmental allowance for people who experienced a large rent increase because of renovation of their dwelling or who had to search for a different dwelling after pull down of their previously rented dwelling). The reason for this latter drop is that the amount is a substantial part of the housing expenditure (16% on average) and it is not clear from the data whether this amount is included in the answers on rent payments or not (±1.2% of the renters in 1986).

Housing consumption for owners:

(1-tax)**erfpacht* + tax**huurwaardeforfait* + (1-tax)*interest payment + foregone interest − increase in the value of the house + maintenance costs + *eigenaarsgedeelte onroerend goedbelasting* + *opstalverzekering*.

Here *erfpacht* is the amount of money you have to pay if you do not own the land on which your dwelling is built (which is partly deductible), tax is the marginal tax rate of the most earning adult in the household, *huurwaardeforfait* is tax levied on the value of the house of owners, *eigenaarsgedeelte onroerend goedbelasting* is municipal tax for house owners and *opstalverzekering* is a house insurance for fire, broken windows etc. Expenditure on gas/water/electricity/heating is excluded.

Computation of the variables in expenditure for owners

Approximately 140 house owning families dropped because the value of the house is not known, which is necessary to correct for, among other things, *huurwaardeforfait*. In the data we have either the amount spent on interest payments on the mortgage or the interest rate on the mortgage.

If we only have the interest rate on the mortgage we computed the interest payments by multiplying this percentage with the mortgage value. If the mortgage value is not reported we used 149000, 155000 and 163000 (the average value of a house for 1987, 1988 and 1989). Foregone interest is set equal to 0.04 times the difference in the value of the house and the mortgage value. Maintenance costs are defined as 2 percent of the value of the house. In the main text we investigate the sensitivity of the results with respect to the percentage increase in the value of a house and the percentage used in the maintenance costs. Because the *eigenaarsgedeelte onroerend goedbelasting* can differ per municipal it is calculated as follows: we have data over 1986-1989 on Tilburg and we will consider Tilburg to be representative for its province. Per province we have the amount of tax that was payed to the local government per inhabitant of the municipality (CBS, Statistiek der gemeentebegroting). The *eigenaarsgedeelte onroerend goedbelasting* per province is calculated as the figure for Tilburg times the relative tax per inhabitant of the province. The relative tax for the provinces is approximately constant over time. The *opstalverzekering* is simply 12.95 times the value of the house divided by 100000 (Budgethandboek NIBUD, 1987).

Computation of marginal tax rate

In the SEP we only observe net income like net wages, net unemployment benefits, net pensions etc. To calculate the marginal tax rate we need gross income of the spouse that earns most because he/she will have to report the tax related issues of owning a house (like e.g. *huurwaardeforfait*). From the net income we could try to invert the tax system and infer gross income. However, this is a very cumbersome approach. Therefore we will follow Euwals and Van Soest (1995). Gross income is already available for individuals with a payed job. We now estimate a net wage equation using the households in which at least one individual has a paid job. An important variable to be included is the tax free allowance (TFA). Constructing this for married couples involves the gross income of the other spouse. All the households for whom we could determine the TFA were included in estimation. The equation estimated is the same as in Euwals and Van Soest (1995), i.e. without a constant term. Without making differences between men and women we got an R^2 of .9955 and the parameter estimates are fairly similar. Given the net income we can now estimate gross income by inverting the relationship. By taking derivatives of net income with respect to gross income we can estimate the marginal tax rate.

General remarks concerning the data

The following data cleaning operations have been applied.

- People who got married or divorced are left out in the analysis to avoid dependence between

households in the sample (±140 households per year);

- households that spend more than 1.5 times their monthly income on housing (±70 households per year) are also left out.

In general we loose approximately 600 households per year. In 1986 we loose 100 more because we do not have good data on the value of the house. If we use only the observations with income budget shares smaller than 1.5 we end up with 3122, 3006, 3224 and 3321 observations.

Comparing the data with macro data

We will compare the 1989 and the 1986 data with the figures in *Woningbehoeftenonderzoek 1989/1990* and the *Woningbehoeftenonderzoek 1985/1986* reported by Statistics Netherlands (CBS). Their definitions for rent and income are the same as the ones we use. For renters the CBS tabulates rent, net annual income and budget shares. The definition of expenditure on housing for owners differs from our measure. The CBS measure of housing expenditure includes expenditure on the mortgage, *erfpacht*, *opstalverz.*, *eigenaarsgedeelte onroerend goed belasting*, *rijksbijdrage eigen woning bezit* and tax issues like interest, *erfpacht*, *huurwaardeforfait* en *rijksbijdrage eigen woning bezit*. We constructed this measure without *opstalverz.*, *eigenaarsgedeelte onroerend goed belasting*, *rijksbijdrage eigen woning bezit* and related tax issues. For owners the CBS tabulates net yearly income and budget shares.

Comparing the 1989 data with the statistics in the *Woningbehoeftenonderzoek 1989/1990* we conclude that:

- house owners are overrepresented in our sample. We see two reasons for this: the group of one-person households is underrepresented and 75% of this group rents, and the owners are overrepresented in the more-than-one-person households;
- for renters our rent data follow the results of the CBS, but low income households (<26000 net per year) are underrepresented and the higher budget shares are overrepresented yielding an average budget share of 0.22 instead of 0.18. The median is 0.19 but this is not reported in the CBS figures;
- the data for owners with a mortgage are similar to the CBS results in the sense that the distribution of net annual income is similar, the distribution of budget shares is similar and the average budget share is 0.133 instead of 0.137;
- households owning a house without a mortgage are underrepresented.

Comparing the 1986 data with the statistics in the *Woningbehoeftenonderzoek 1985/1986* we conclude that:

- again house owners are overrepresented in our sample, probably for same two reasons discussed above.
- the distribution of the rent per month is similar to the CBS data but for renters the high net yearly incomes (>38000) are underrepresented and, related to this, higher budget shares are overrepresented (on average it is 0.21 whereas it should be 0.17 according to the CBS). Especially the budget shares larger than 0.32 are overrepresented. The median is 0.18 but this figure is not reported in the CBS figures;
- for owners the low net yearly incomes (<25000) are underrepresented but the budget shares (for the ones with a mortgage) are conform the CBS data except that again the budget shares larger than 0.32 seem to be a bit overrepresented the mean is 0.17 whereas it should be approximately 0.15). The median is 0.13 but this figure is not reported by the CBS;
- households owning a house without a mortgage are underrepresented.

In general the data have the following features:

- the density of income shifts a little bit to the right over time;
- the density of the budget shares for renters remains approximately the same over time;
- the density of the budget shares for owners is slightly shifted to the right when compared to the 1989 macro data. The data for 1986 contain too many people with budget shares over 0.32;
- the density of the interest payments on mortgages looks the same for all years. However, the average value for 1986 is still a bit high but the average is increasing (slightly) over 1987-1989.

APPENDIX A

In this appendix we will investigate the sensitivity of the cross-section Newey IV results and the panel data Kyriazidou IV results with respect to the maintenance costs and the mortgage costs in housing consumption for owners. Let BS1ab denote the Budget Share spent on housing for owners with a% increase of the value of a house (a=0,1,2,3,4) and b% of the value of the house as the maintenance costs (b=1,2). In the main text a equals 1 and b equals 2. From the definition of housing costs for owners it follows that BS1ab=BS1a+1,b+1 so eg. BS121=BS132. Because the averages for BS142, BS132 (and henceBS131 and BS121) are very low compared to the average for renters we only consider BS122, BS112 and BS102. The last digit is then dropped because it is fixed at 2. Hence we consider BS1a with the maintenance costs fixed at 2 % of the value of the house. BS11 is used throughout the main text. The means for BS12, BS11 and BS10 are respectively 0.18, 0.22 and 0.27 with standard errors of 0.15, 0.18 and 0.22.

In the next table we indicate the sensitivity of the parameter estimates of the Newey IV estimates with respect to the measure for housing expenditure for owners. The coefficients related to LEXP, L2EXP, DMAR and DREG1 tend to change somewhat, but the main conclusions remain the same. The standard errors remain rather large such that we do not find significant differences in the parameter estimates when varying housing expenditure for owners.

Sensitivity of the estimation results with respect to the measure for housing expenditure of owners, cross-section[a]

Variable	BS12 Newey IV[b,c]		BS11 Newey IV[b,c]		BS10 Newey IV[b,c]	
CONSTANT	7.289[d]		15.454[d]		18.108[d]	
AGE	0.028	(0.053)	0.028	(0.066)	0.035	(0.077)
AGE2	−0.004	(0.007)	−0.005	(0.008)	−0.005	(0.009)
LEXP	−3.040	(1.634)	−3.670	(1.939)	−4.300	(2.241)
L2EXP	0.181	(0.105)	0.219	(0.121)	0.256	(0.144)
DMAR	0.032	(0.020)	0.041	(0.024)	0.050	(0.028)
DREG1	−0.029[*]	(0.013)	−0.037[*]	(0.016)	−0.045[**]	(0.018)
DREG2	−0.005	(0.010)	−0.003	(0.012)	−0.001	(0.014)
DREG3	−0.010	(0.011)	−0.009	(0.014)	−0.008	(0.017)

[a] [*] means significant at the 5% level, [**] means significant at the 1% level. The results for renters and for the selection equation are the ones presented in the second and third column of table 3

[b] series approximation using single index ML probit in estimating the selection equation

[c] IV using AGE, AGE2, LINC, L2INC, DMAR, DREG1, DREG2 and DREG3 as instruments

[d] estimates include the estimate for the constant term in the series approximation

In the next table we indicate the sensitivity of the parameter estimates of the Kyriazidou IV panel estimates with respect to the measure for housing expenditure for owners. Most coefficients change somewhat but the main conclusions remain the same.

Sensitivity of the estimation results with respect to the measure for housing expenditure of owners, panel[a]

Variable	Kyriaz. IV BS12		Kyriaz. IV BS11		Kyriaz. IV BS10	
AGE	0.051	(0.095)	0.115	(0.106)	0.1605	0.117
AGE2	−0.006	(0.010)	-0.012	(0.012)	−0.0160	0.013
LEXP	−1.942**	(0.162)	-2.312**	(0.205)	−2.6841	0.249
L2EXP	0.110**	(0.010)	0.130**	(0.013)	0.1505	0.015
Dummy87	−0.004	(0.004)	-0.009*	(0.004)	−0.0143	0.005
Dummy88	−0.001	(0.002)	-0.001	(0.003)	−0.0017	0.004

[a] * means significant at the 5% level, ** means significant at the 1% level. The results for the selection equation are the ones presented in the second and third column of table 6 and the results for renters are the ones in the sixth and seventh column of table 6. Because the smoothing parameters are related to the index of the first step estimates only, the smoothing parameters are the ones reported in table 6.

Appendix B

In this appendix we discuss some details of implementing the Newey (1988) estimator discussed in section 3.1. Starting point is the random effects model which, for one cross-section, can be written as

$$d_i = 1(\pi' x_i - v_{si} \geq 0).$$
$$y_{0i} = \beta_0' x_i + v_{0i} \quad \text{if } d_i=0$$
$$y_{1i} = \beta_1' x_i + v_{1i} \quad \text{if } d_i=1$$

Compared to the notation in section 3.1, the time index t is omitted and the random effects are incorporated in the error terms $v_i=(v_{si},v_{0i},v_{1i})$, which is independent of x_i. Newey uses the fact that the independence assumption implies that the distribution of v_i depends on (x_{bi},x_{di}) only through the index $\pi' x_i=\pi_b' x_{bi}+\pi_d' x_{di}$. This implies that

$$y_{pi} = \beta_{pa}' x_{ai} + \beta_{pc}' x_{ci} + g_p(x_{bi},x_{di}) + \tilde{\varepsilon}_{pi}, \quad \text{with}$$
$$g_p(x_{bi},x_{di})=E\{\varepsilon_{pi} | x_{bi}, x_{di}, d_i=p\} \text{ and } E\{\tilde{\varepsilon}_{pi} | x_{bi}, x_{di}, d_i=p\} = 0, \quad p=0,1.$$

and where the functions g_0 and g_1 can then be written as

$$g_p(x_{bi},x_{di}) = \tilde{g}_p(\pi_b' x_{bi}+\pi_d' x_{di}), \quad p=0,1.$$

To estimate the budget equations, $\tilde{g}_0$ and $\tilde{g}_1$ are approximated by $\Sigma_{k=0}^{K}\alpha_{pk}(\pi_b' x_{bi}+\pi_d' x_{di})^k$, p=0,1, with K=K(p,n) (p=0,1, n is the number of observations). The following regression equations can now be used for the subsamples of renters and owners separately

$$y_{pi} = \beta_{pa}' x_{ai} + \beta_{pc}' x_{ci} + \Sigma_{k=0}^{K}\alpha_{pk}(\hat{\pi}_b' x_{bi}+\hat{\pi}_d' x_{di})^k + \hat{\varepsilon}_{pi}, \qquad (1)$$

where $\hat{\pi}_b$ and $\hat{\pi}_d$ denote estimates of π_b and π_d, respectively (to be discussed below). If x_{ai} is exogenous, consistent and asymptotically normal estimates for $(\beta_{0a}',\beta_{0c}')$ and $(\beta_{1a}',\beta_{1c}')$ can be obtained by applying OLS to equation (1) for each subsample. This was shown by Newey (1988), who also derives a consistent estimator for the asymptotic covariance matrices of the estimators.

We apply Newey's procedure to the case that x_{ai} is allowed to be endogenous by replacing OLS with IV. Denote the regressors in equation (1) corresponding for the case p=1 by $\hat{x}_i^s$, i.e. $\hat{x}_i^s=(x_{ai}',x_{ci}',1,(\hat{\pi}_b' x_{bi}+\hat{\pi}_d' x_{di})^1,..,(\hat{\pi}_b' x_{bi}+\hat{\pi}_d' x_{di})^K)'$ (with now K=K(1,n)) and let $\hat{X}^s=(\hat{x}_1^s,..,\hat{x}_{n1}^s)'$ where n1 is the number of observations with $d_i=1$. Furthermore, let $\hat{w}_i^s$ be the vector of instruments, i.e. $\hat{x}_i^s$ with

x_{ai} replaced by x_{bi} (hence $\hat{w}_i^s$ is of the same dimension as $\hat{x}_i^s$), and let $\hat{W}^s=[\hat{w}_1^s,..,\hat{w}_{n1}^s]'$. The parameters β_{1a}, β_{1c}, and α_{11} to α_{1K} can now be estimated by applying IV to equation (1). Under appropriate regularity conditions[17] the IV-estimates for β_{1a} and β_{1b} will be consistent and asymptotically normal: $\sqrt{n}((\hat{\beta}_{1a}',\hat{\beta}_{1c}')'-(\beta_{1a}',\beta_{1c}')')\rightarrow^{d}N(0,V)$. Notice, however, that the constant term in the regression equation cannot be estimated separately, since the series approximation also includes a constant term.[18] The asymptotic covariance matrix V can be estimated consistently by

$$[I,0](\hat{W}^{s\prime}\hat{X}^s)^{-1}\left\{\sum_{i=1}^{n1}\hat{w}_i^s\hat{w}_i^{s\prime}\tilde{e}_i^2 + \hat{H}_w\hat{V}(\hat{\pi}_b,\hat{\pi}_d)\hat{H}_w'\right\}(\hat{X}^{s\prime}\hat{W}^s)^{-1}[I,0]'$$

where $\tilde{e}_i$ is the IV residual and

$$\hat{H}_w = \sum_{i=1}^{n1}\left\{\hat{w}_i^s(x_{bi}',x_{di}')\left(\sum_{k=1}^{K}k\hat{\alpha}_{1k}(\hat{\pi}_b'x_{bi}+\hat{\pi}_d'x_{di})^{k-1}\right)\right\}$$

where $\hat{\alpha}_{1k}$, k=1,..,K are the IV estimates of the α_{1k}. The expressions in Newey (1988) are a special case with $\hat{W}^s$ replaced by $\hat{X}^s$, $\tilde{e}_i$ by the OLS residuals and $\hat{\alpha}_{1k}$, k=1,..,K by the OLS estimates. The parameters in the other equation (p=0) can be estimated analogously.

The smoothing parameter in the estimation procedure is the number of terms in the series approximation, which is chosen such that adding more terms in the series approximation no longer affects the parameter estimates for the regression coefficients. In practice, often only a few terms in the series approximation turn out to be required.

The Newey approach for estimating β_p, p=0,1, requires estimation of a single index binary choice model[19] to obtain estimates for (π_b',π_d'). Klein and Spady (1993) have proposed an estimator which is semiparametrically efficient under weak regularity assumptions. This estimator, however, is difficult to compute. Instead, we started with the probit ML estimates for (π_b',π_d'). We

---

[17] Appropriate regularity conditions should include conditions guaranteeing consistency of the IV estimates of β_{1a} and β_{1c} and conditions that allow one to derive the presented limit distribution. The former conditions will be different from Newey's, since identification should now be based on moment restrictions. Given identification (and consistency) the latter conditions will be comparable to Newey's conditions.

[18] Andrews and Schafgans (1995) show how the constant term can be estimated if observations with selection probability close to one are available. Since, however, we do not have many observations with probability of ownership close to zero or one, this approach is practically infeasible for both renters and owners.

[19] Ahn and Powell (1993) allow for a more general model, in which the probability of ownership is estimated completely nonparametrically. Due to the large number of explanatory variables in the selection equation, such an approach is practically infeasible for our purposes.

tested for normality and heteroskedasticity of exponential form using tests described in Chesher and Irish (1987). Both normality and homoskedasticity were rejected. Therefore, we experimented with the following specification, in which the single index assumption is retained:

$$P\{d_i=1 \mid x_{bi},x_{di}\} = \Phi(m(\tau,\pi_b'x_{bi}+\pi_d'x_{di})/\exp\{\sigma(\gamma,\pi_b'x_{bi}+\pi_d'x_{di})\})$$

Here m and σ are power series in $\pi_b'x_{bi}+\pi_d'x_{di}$ with coefficients τ and γ, respectively. This can be seen as a series approximation to an arbitrary single index model. Let τ_j and γ_j denote the coefficients related to $(\pi_b'x_{bi}+\pi_d'x_{di})^j$. The normalizations imposed are $\tau_0=0$, $\tau_1=1$ and $\gamma_0=0$. We estimated this model for several lengths of the two power series, and found one significant term in m: $(\pi_b'x_{bi}+\pi_d'x_{di})^2$.

APPENDIX C

In this appendix we derive the limit distribution of the minimum distance estimator for the Kyriazidou panel data model with more than two time periods. The estimators used in the first step are the Kyriazidou estimators based on two time periods. They play a major role in determining the limit distribution of the minimum distance estimator. Particularly, we will show that the asymptotic covariance between the Kyriazidou estimators based on a different combination of two different time periods is asymptotically zero. For notational convenience we will show the results comparing the estimator based on time periods one and two with the one based on the periods two and three. The result can be easily extended including more estimators in the first step.

Let $\tilde{\beta}_{1,ts}$ denote the estimator for $(\beta'_{1a},\beta'_{1c})'$ based on time periods s and t. It is easy to show that for the second step minimum distance estimator, b_1, say, we can write

$$\sqrt{ns_{3n}}(b_1 - \beta_1) = A_n\sqrt{ns_{3n}}\begin{bmatrix}\tilde{\beta}_{1,21} - \beta_1 \\ \tilde{\beta}_{1,32} - \beta_1\end{bmatrix} \qquad \textbf{(C.1)}$$

for some matrix A_n converging in probability to A, say, when $n\to\infty$, and some smoothing parameter s_{3n}. Hence the limit distribution of the minimum distance estimator is determined by the limit distribution of

$$\sqrt{ns_{3n}}\begin{bmatrix}\tilde{\beta}_{1,21} - \beta_1 \\ \tilde{\beta}_{1,32} - \beta_1\end{bmatrix} \qquad \textbf{(C.2)}$$

From Kyriazidou (1995) we have

$\sqrt{(ns_{1n})}(\tilde{\beta}_{1,21}-\beta_1) \to^d N(AB_1,V_1)$, and
$\sqrt{(ns_{2n})}(\tilde{\beta}_{1,32}-\beta_1) \to^d N(AB_2,V_2)$,

with AB_1, AB_2 the asymptotic bias, and V_1, V_2 the asymptotic covariance matrices.

Using the optimal estimators (i.e minimizing asymptotic MSE) in the first round it follows that $s_{1n}=O(n^{-\alpha})$ and $s_{2n}=O(n^{-\alpha})$ for some $0<\alpha<½$. Therefore also $s_{3n}=O(n^{-\alpha})$.
Now define

$$\lim_{n\to\infty}\frac{ns_{3n}}{ns_{1n}} = c_{31}, \text{ and } \lim_{n\to\infty}\frac{ns_{3n}}{ns_{2n}} = c_{32} \qquad \textbf{(C.3)}$$

with $0<c_{31}<\infty$ and $0<c_{32}<\infty$.

Then

$$\sqrt{ns_{3n}}\begin{bmatrix}\tilde{\beta}_{1,21}-\beta_1\\ \tilde{\beta}_{1,32}-\beta_1\end{bmatrix} = \begin{bmatrix}\sqrt{ns_{1n}}(\tilde{\beta}_{1,21}-\beta_1)\dfrac{\sqrt{ns_{3n}}}{\sqrt{ns_{1n}}}\\ \sqrt{ns_{2n}}(\tilde{\beta}_{1,32}-\beta_1)\dfrac{\sqrt{ns_{3n}}}{\sqrt{ns_{2n}}}\end{bmatrix} \rightarrow^{d} N\left(\begin{bmatrix}AB_1\sqrt{c_{31}}\\ AB_2\sqrt{c_{32}}\end{bmatrix},\begin{bmatrix}c_{31}V_1 & cov\\ cov & c_{32}V_2\end{bmatrix}\right) \qquad \textbf{(C.4)}$$

We will now show that cov=cov($\sqrt{(ns_{3n})}(\tilde{\beta}_{1,21}-\beta_1),\sqrt{(ns_{3n})}(\tilde{\beta}_{1,32}-\beta_1)$) tends to zero as n tends to infinity. Because the first round estimator for π converges at a faster rate, the limit distribution of the estimators can be analyzed assuming we know the true value for π (analogously to Kyriazidou, 1995). Then $\tilde{\beta}_{1,21}$ and $\tilde{\beta}_{1,32}$ are (following the notation in the main text)

$$\begin{aligned}\tilde{\beta}_{1,21} &= \left[\sum_{i=1}^{n}\frac{1}{s_{1n}}K\left(\frac{\pi_b'(x_{bi2}-x_{bi1})+\pi_d'(x_{di2}-x_{di1})}{s_{1n}}\right)(w_{i2}-w_{i1})(\tilde{x}_{i2}-\tilde{x}_{i1})'d_{i1}d_{i2}\right]^{-1}\\ &\quad \sum_{i=1}^{n}\frac{1}{s_{1n}}K\left(\frac{\pi_b'(x_{bi2}-x_{bi1})+\pi_d'(x_{di2}-x_{di1})}{s_{1n}}\right)(w_{i2}-w_{i1})(y_{1i2}-y_{1i1})d_{i1}d_{i2}\\ &= \beta_1 + [S_{w21,x21}]^{-1}S_{w21,\epsilon 21}\end{aligned} \qquad \textbf{(C.5)}$$

$$\begin{aligned}\tilde{\beta}_{1,32} &= \left[\sum_{i=1}^{n}\frac{1}{s_{1n}}K\left(\frac{\pi_b'(x_{bi3}-x_{bi2})+\pi_d'(x_{di3}-x_{di2})}{s_{1n}}\right)(w_{i3}-w_{i2})(\tilde{x}_{i3}-\tilde{x}_{aci2})'d_{i2}d_{i3}\right]^{-1}\\ &\quad \sum_{i=1}^{n}\frac{1}{s_{1n}}K\left(\frac{\pi_b'(x_{bi3}-x_{bi2})+\pi_d'(x_{di3}-x_{di2})}{s_{1n}}\right)(w_{i3}-w_{i2})(y_{1i3}-y_{1i2})d_{i2}d_{i3}\\ &= \beta_1 + [S_{w32,x32}]^{-1}S_{w32,\epsilon 32}\end{aligned} \qquad \textbf{(C.6)}$$

Because the inverted matrices in (C.5) and (C.6) converge in probability they will be ignored in the remainder.

Analogous to Kyriazidou (1995, proof of lemma 1) one can show that

$$\begin{aligned}&\sqrt{ns_{3n}}S_{w21,\epsilon 21} = \frac{\sqrt{ns_{3n}}}{\sqrt{ns_{1n}}}\sqrt{ns_{1n}}S_{w21,\epsilon 21} = \frac{\sqrt{ns_{3n}}}{\sqrt{ns_{1n}}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{i21n},\\ &\text{where } \xi_{i21n} = \frac{1}{\sqrt{s_{1n}}}K\left(\frac{\pi_b'(x_{bi2}-x_{bi1})+\pi_d'(x_{di2}-x_{di1})}{s_{1n}}\right)(w_{i2}-w_{i1})\Delta v_{i21}d_{i1}d_{i2}\end{aligned} \qquad \textbf{(C.7)}$$

$$\text{and } \Delta v_{i21}=v_{i2}-v_{i1}\ ,\ v_{it}=\epsilon_{it}-E\{\epsilon_{it}|d_{i1}=d_{i2}=1,x_{bi1},x_{di1},x_{bi2},x_{di2},\alpha_i,\eta_i\}$$

and a similar expression holds for $S_{w32,\epsilon 32}$. Now drop the subscript i and define $\Delta w_{21}=w_2-w_1$,

$\Delta w_{32}=w_3-w_2$, $G_{21}=\pi_b'(x_{bi2}-x_{bi1})+\pi_d'(x_{di2}-x_{di1})$ and $G_{32}=\pi_b'(x_{bi3}-x_{bi2})+\pi_d'(x_{di3}-x_{di2})$. Using that ξ_{i21n} and ξ_{i32n} have expectation zero it follows that (suppressing i subscripts)

$$
\begin{aligned}
&\operatorname{cov}\left(\sqrt{ns_{1n}}S_{w21,\epsilon 21},\sqrt{ns_{2n}}S_{w32,\epsilon 32}\right)\\
&= \mathbb{E}\left\{\xi_{21n}\xi_{32n}'\right\}\\
&= \int\int \mathbb{E}\left\{\Delta w_{21}\Delta w_{32}'\Delta v_{21}\Delta v_{32}\middle| G_{21},G_{32}\right\}\frac{1}{\sqrt{s_{1n}s_{2n}}}\mathbb{K}\left(\frac{G_{21}}{s_{1n}}\right)\mathbb{K}\left(\frac{G_{32}}{s_{2n}}\right)f_{G_{21},G_{32}}(G_{21},G_{32})\,dG_{21}\,dG_{32}\\
&= \sqrt{s_{1n}s_{2n}}\int\int \mathbb{E}\left\{\Delta w_{21}\Delta w_{32}'\Delta v_{21}\Delta v_{32}\middle| G_{21}=v_{21}s_{1n},G_{32}=v_{32}s_{1n}\right\}\mathbb{K}(v_{21})\mathbb{K}(v_{32})f_{G_{21},G_{32}}(v_{21},v_{32})\,dv_{21}\,dv_{32}\\
&\to 0*f_{G_{12},G_{32}}(0,0)\,\mathbb{E}\left\{\Delta w_{21}\Delta w_{32}'\Delta v_{21}\Delta v_{32}\middle| G_{21}=G_{32}=0\right\}\left[\int \mathbb{K}(v)\,dv\right]^2\\
&= 0 \quad (n\to\infty)
\end{aligned}
\tag{C.8}
$$

Using (C.3) and (C.8) it now follows that

$$
\sqrt{ns_{3n}}\begin{bmatrix}\tilde{\beta}_{1,21}-\beta_1\\ \tilde{\beta}_{1,32}-\beta_1\end{bmatrix}\to^d N\left(\begin{bmatrix}AB_1\sqrt{c_{31}}\\ AB_2\sqrt{c_{32}}\end{bmatrix},\begin{bmatrix}c_{31}V_1 & 0\\ 0 & c_{32}V_2\end{bmatrix}\right) \tag{C.9}
$$

In practice we need to estimate

$$
\begin{bmatrix}\dfrac{AB_1}{\sqrt{ns_{1n}}}\dfrac{\sqrt{c_{31}}\sqrt{ns_{1n}}}{\sqrt{ns_{3n}}}\\ \dfrac{AB_2}{\sqrt{ns_{2n}}}\dfrac{\sqrt{c_{32}}\sqrt{ns_{2n}}}{\sqrt{ns_3n}}\end{bmatrix},\ \text{and}\ \begin{bmatrix}\dfrac{V_1}{ns_{1n}}\dfrac{c_{31}ns_{1n}}{ns_{3n}} & 0\\ 0 & \dfrac{V_2}{ns_{2n}}\dfrac{c_{32}ns_{2n}}{ns_{3n}}\end{bmatrix} \tag{C.10}
$$

The quantities $AB_1/\sqrt{(ns_{1n})}$, $AB_2/\sqrt{(ns_{2n})}$, $V_1/(ns_{1n})$ and $V_2/(ns_{2n})$ are what we estimate in the first step of the estimation procedure, so the question is how to estimate the other quantities. A possible way to do this is to assume that $s_{jn}=c_jn^{-\alpha}$ for some c_j, j=1,2,3. Then it follows that all the remaining quantities in (C.10) are equal to 1 and hence we only need the bias and variance estimates from the first step. We use this choice in the main text. For s_{jn}, j=1,2, Kyriazidou (1995) assumes the structure mentioned before. However, the assumption that $s_{3n}=c_3n^{-\alpha}$, although natural, can be restrictive in small samples. Therefore, we also investigated the sensitivity of the results when the remaining quantities in (C.10) are slightly different from 1.