

CBM  
R

8414  
1990 —  
990-64  
64



\* C I N 0 0 9 5 2 \*



No. 9064

**CAN COHORT DATA BE TREATED AS GENUINE  
PANEL DATA?**

by Marno Verbeek  
and Theo Nijman

R46  
518.9

November 1990

ISSN 0924-7815

# Can cohort data be treated as genuine panel data? \*

by Marno Verbeek and Theo Nijman

Tilburg University  
Department of Econometrics  
P.O.Box 90153  
5000 LE Tilburg  
The Netherlands

November 1990

---

\*The authors thank Bertrand Melenberg, Robert Moffitt, Guglielmo Weber, seminar participants at Texas A&M University, Rice University and the Conference on "The Econometrics of Panels and Pseudo Panels" (Venice, October 1990) and two anonymous referees for helpful comments. Rob Alessie and Pim Adang kindly provided the data. Financial support by the Royal Netherlands Academy of Arts and Sciences (K.N.A.W.) and the Netherlands Organization for Scientific Research (N.W.O.) is gratefully acknowledged.

## Abstract

If repeated observations on the same individuals are not available it is not possible to capture unobserved individual characteristics in a linear model by using the fixed effects estimator in the standard way. If large numbers of observations are available in each period one can use cohorts of individuals with common characteristics to achieve the same goal, as shown by Deaton [1985]. It is tempting to analyze the observations on cohort averages as if they are observations on individuals which are observed in consecutive time periods. In this paper we analyze under which conditions this is a valid approach. Moreover, we consider the impact of the construction of the cohorts on the bias in the standard fixed effects estimator. Our results show that the effects of ignoring the fact that only a synthetic panel is available will be small if the cohort sizes are sufficiently large and if the true means within each cohort exhibit sufficient time variation. In applications the latter condition seems hard to fulfill, which implies that fairly large cohort sizes (100, 200 individuals) are needed to validly ignore the cohort nature of the data.

## 1 Introduction

In recent years much attention is paid to the comparison of panel data with a single cross section or a series of independent cross sections (cf. Hsiao [1985]). In the context of a random effects model, for example, Nijman and Verbeek [1990] show that more efficient estimators of several functions of the parameters can be obtained from a series of cross sections than from a panel (with the same number of observations). On the other hand several authors have stressed the fact that panel data are not indispensable for the identification of many commonly estimated models and that the parameters of interest can be identified (with or without some additional assumptions) from a single cross section or a series of independent cross sections (see, for example, Heckman and Robb [1985], Deaton [1985] and Moffitt [1990]). In this paper we pay attention to a regression model with individual effects that are correlated with the explanatory variables (“the fixed effects model”), and analyze the properties of the within estimator based on aggregated data on cohorts constructed from a series of independent cross sections. In this approach “similar” individuals are grouped in cohorts, after which the averages within these cohorts are treated as observations in a synthetic panel (cf. Deaton [1985]).

Models in which the individual effects are correlated with the explanatory variables often arise naturally from economic theory, for example in life cycle models where the individual effects represent marginal utility of wealth (see, e.g., Heckman and MaCurdy [1980], MaCurdy [1981] and Browning, Deaton and Irish [1985]). Because in many countries no panel data on household consumption or labor supply are available but repeated cross sectional information is, the latter data is typically used to estimate life cycle models. Deaton [1985] has shown that the slope parameters in such models can usually be identified from a series of independent cross sections. In his approach, cohorts are defined as groups of individuals sharing common observed characteristics, such as age or sex. Because the observed cohort aggregates are error-ridden measurements of the true cohort population values, an errors-in-variables estimator is proposed which yields consistent estimators under fairly weak assumptions.

However, if the numbers of observations per cohort is large, it is tempting to ignore the errors-in-variables problem and to use standard software to handle the synthetic panel as if it were a genuine panel. This is what is

usually done in empirical studies, see, e.g., Browning, Deaton and Irish [1985] and Blundell, Browning and Meghir [1989]. In this paper we analyze to what extent this is a valid approach.

First, in Section 2, we present a general introduction to the estimation of a regression model from cohort data if the individual effects are treated as fixed unknown parameters. Moreover, we derive conditions for the consistency of the standard within estimator on the synthetic panel which ignores the measurement errors problem. In Section 3 we derive expressions for the bias of this estimator if the conditions for consistency are not met (under additional assumptions on the data generating process on an individual level and the way in which the cohorts are constructed). Furthermore, we analyze the effects of the choice of the cohort sizes on this bias. In Section 4 attention is paid to the (true) variance of the within estimator compared with the estimated variance from standard routines and the influence from the choice of the cohort sizes on these variances. In section 5 we consider the implications of our results for the estimation of Engel curves for food expenditures from Dutch monthly data. The results suggest that fairly large numbers of observations are required in each cohort to validly ignore the fact that the model is estimated from cohort data. Finally, Section 6 concludes.

## 2 Estimation from cohort data

Consider the following linear model

$$y_{it} = x_{it}\beta + \theta_i + \varepsilon_{it}, \quad t = 1, \dots, T \quad (1)$$

where  $i$  indexes individuals and  $t$  indexes time periods and suppose  $\beta$  is the parameter of interest. Throughout the paper we assume that  $E\{\varepsilon_{it} \mid x_{js}\} = 0$  for all  $s, t = 1, \dots, T$  and all  $i, j$ . In each period, observations on  $N$  individuals are available. Throughout we assume that the data set is a series of independent cross sections. This assumption does not rule out the possibility that some individuals are observed more than once. It is sufficient that each cross section is a random sample of the population such that all covariances between individuals observed in different periods are zero.

In many applications the individual effects  $\theta_i$  are likely to be correlated with the explanatory variables in  $x_{it}$  so that estimation procedures treating

the  $\theta_i$  as random drawings from some distribution lead to inconsistent estimators, unless the correlation is explicitly taken into account. When panel data are available this problem can be solved by treating the  $\theta_i$  as fixed unknown parameters. Usually the fixed effects are eliminated before estimation, for example by a within or first difference transformation. Obviously, this strategy no longer applies if no repeated observations on the same individuals are available.

Deaton [1985] suggests the use of cohorts to obtain consistent estimators for  $\beta$  in (1) if repeated cross sections are available, even if the individual effects are correlated with the explanatory variables. Let us define  $C$  cohorts, which are groups of individuals sharing some common characteristics. These cohorts are defined in such a way that each individual is a member of exactly one cohort which is the same for all periods. For example, a particular cohort may consist of all individuals born in 1945-1949, or of all males having a university degree on January 1, 1990. Aggregation of all observations to cohort level results in

$$\bar{y}_{ct} = x_{ct}\beta + \bar{\theta}_{ct} + \bar{\varepsilon}_{ct}, \quad c = 1, \dots, C; t = 1, \dots, T \quad (2)$$

where  $\bar{y}_{ct}$  and  $\bar{x}_{ct}$  are the averages of all observed  $y_{it}$ 's and  $x_{it}$ 's in cohort  $c$  at time  $t$ . The resulting data set is a synthetic (or pseudo) panel with repeated observations on  $C$  cohorts over  $T$  periods. The main problem with the estimation of this model is that  $\bar{\theta}_{ct}$  in (2) depends on  $t$ , is unobserved and is likely to be correlated with  $\bar{x}_{ct}$ . Therefore, treating the  $\bar{\theta}_{ct}$  as random (and uncorrelated with the explanatory variables) is likely to lead to inconsistent estimators and treating them as fixed will result in an identification problem unless the variation of  $\bar{\theta}_{ct}$  over  $t$  can be neglected. Intuitively, the latter will be the case if the number of observations within each cohort is large. In the remainder of this section we analyze in more detail the conditions under which the within estimator on the synthetic panel will be consistent.

An alternative way to approach the problem is adopted by Deaton [1985], who considers the cohort population version of (1),

$$y_{ct}^* = x_{ct}^*\beta + \theta_c^* + \varepsilon_{ct}^*, \quad c = 1, \dots, C; t = 1, \dots, T \quad (3)$$

where the asterisks denote (unobservable) population cohort means and where  $\theta_c^*$  is the cohort fixed effect, which is constant over time because population cohorts contain the same individuals in all periods (ignoring birth and death

of individuals). If the population cohort means would be observable, equation (3) could be used to estimate  $\beta$  using standard procedures for a panel consisting of  $C$  cohorts observed in  $T$  periods. However, we can regard the observed cohort means  $\bar{y}_{ct}$  and  $\bar{x}_{ct}$  as error-ridden measurements of the true population cohort means  $y_{ct}^*$  and  $x_{ct}^*$ . Deaton [1985] assumes that the measurement errors in  $\bar{y}_{ct}$  and  $\bar{x}_{ct}$  are normally distributed with zero mean and independent of the true values  $y_{ct}^*$  and  $x_{ct}^*$ , in particular<sup>1</sup>.

$$\begin{pmatrix} \bar{y}_{ct} \\ \bar{x}_{ct} \end{pmatrix} \sim N \left( \begin{pmatrix} y_{ct}^* \\ x_{ct}^* \end{pmatrix}; \begin{pmatrix} \sigma_{00} & \sigma' \\ \sigma & \Sigma \end{pmatrix} \right). \quad (4)$$

One way to estimate the parameter  $\beta$  in (3) is to analyze the model in (3) and (4) as a model with measurement errors. If the row vector of cohort dummies is denoted by  $d_c$  and the column vector of corresponding parameters is denoted by  $\theta^* = (\theta_1^*, \dots, \theta_C^*)'$ , the errors-in-variables estimator (on the model in levels) proposed by Deaton [1985] is given by

$$\begin{pmatrix} \tilde{\theta} \\ \tilde{\beta} \end{pmatrix} = \left( \sum_{c=1}^C \sum_{t=1}^T \begin{pmatrix} d_c' d_c & d_c' \bar{x}_{ct} \\ \bar{x}_{ct}' d_c & \bar{x}_{ct}' \bar{x}_{ct} - \hat{\Sigma} \end{pmatrix} \right)^{-1} \left( \sum_{c=1}^C \sum_{t=1}^T \begin{pmatrix} d_c' \bar{y}_{ct} \\ \bar{x}_{ct}' \bar{y}_{ct} - \hat{\sigma} \end{pmatrix} \right). \quad (5)$$

where  $\hat{\Sigma}$  and  $\hat{\sigma}$  are estimates of  $\Sigma$  and  $\sigma$  based on all individual observations. If the following assumption holds, the estimator  $\tilde{\beta}$  is consistent for  $\beta$  if the number of observations  $CT$  tends to infinity, while  $\tilde{\theta}$  is consistent for  $\theta^*$  if the total number of observations per cohort ( $TN/C$ ) tends to infinity.

**Assumption 2.1** *The moments matrix of the population means of the explanatory variables*

$$\text{plim}_{CT \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T \begin{pmatrix} d_c' d_c & d_c' \bar{x}_{ct} \\ \bar{x}_{ct}' d_c & \bar{x}_{ct}' \bar{x}_{ct} - \hat{\Sigma} \end{pmatrix} \quad (6)$$

*is nonsingular.*

If the number of observations per cohort is not too small, it is tempting to ignore the errors-in-variables problem and to estimate (2) assuming equality

<sup>1</sup>Note that, contrary to Deaton, we do not include the cohort dummies in the vector of  $x$ 's. These dummies are of course observed without error.



of population and sample means. The resulting estimator for  $\beta$  is the within estimator on the synthetic panel,  $\hat{\beta}_W$ , given by

$$\hat{\beta}_W = \left( \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)' (\bar{x}_{ct} - \bar{x}_c) \right)^{-1} \left( \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)' (\bar{y}_{ct} - \bar{y}_c) \right), \quad (7)$$

where  $\bar{x}_c$  is the time average of  $\bar{x}_{ct}$ , i.e.  $\bar{x}_c = \frac{1}{T} \sum_{t=1}^T \bar{x}_{ct}$  and analogously for  $\bar{y}_c$ . Using (2) it is easy to show that  $\hat{\beta}_W$  is unbiased if

$$E\{\bar{\theta}_{ct} - \bar{\theta}_c \mid \bar{x}_{ct} - \bar{x}_c\} = 0 \quad (8)$$

provided the following assumption holds.

**Assumption 2.2** *The moments matrix of the observed cohort means of the explanatory variables*

$$\text{plim}_{CT \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)' (\bar{x}_{ct} - \bar{x}_c) \quad (9)$$

*is nonsingular.*

It is important to note that Assumption 2.2 is implied by Assumption 2.1 but that the converse is not true. Condition (8) will be satisfied if  $\theta_i$  is independent of  $x_{it}$  (for all  $t$ ) or if the averaged individual effects  $\theta_i$  are constant over time ( $\bar{\theta}_{ct} = \bar{\theta}_c$ ). If the number of observations per cohort,  $N/C$ , is large, one is tempted to assume that the latter condition holds. In the sequel of this paper we shall pay attention to the bias in the cohort within estimator  $\hat{\beta}_W$  given the number of observations per cohort ( $N/C$ ). Note that increasing the number of observations per cohort implies a decrease in the number of observations in the synthetic panel and thus an increase in the variance of the within estimator on the synthetic panel. Evidently, the optimal choice of the cohorts will depend on both its impact on the bias and its impact on the variance, which will be analyzed (for a simple model) in Sections 3 and 4 respectively.

A striking point from our results is that it is possible that Deaton [1985]'s estimator has a nonexisting probability limit (for  $CT \rightarrow \infty$ ), while  $\hat{\beta}_W$  has a well-defined probability limit which may even equal the true value  $\beta$ . This will happen when Assumption 2.2 is satisfied but Assumption 2.1 is not. We will return to this point in the next section.

### 3 The effects of the choice of cohorts on the bias

Our basic interest lies in the validity of the argument that “the number of observations per cohort is large enough to ignore the errors-in-variables problem” (cf., e.g., Browning, Deaton and Irish [1985]). We therefore concentrate on the case where the number of observations per cohort  $N/C$  is fixed. To simplify the analytical results we approximate the finite sample bias by the asymptotic bias for large  $C$  and large  $N$ . Numerical checks reveal that this approximation is accurate if  $C$  is not too small (10–20). Under Assumption 2.1, this is one of the situations in which Deaton’s estimator is consistent (for  $C \rightarrow \infty$ ) and the standard within estimator need not be consistent.

In this section we will derive the asymptotic bias in  $\hat{\beta}_W$  for the special case of a linear model with only one explanatory variable,

$$y_{it} = \beta x_{it} + \theta_i + \varepsilon_{it} \quad (10)$$

where  $x_{it}$  is a scalar variable. Following Chamberlain [1984], we assume that the dependence of  $x_{it}$  and  $\theta_i$  can be characterized as follows.

**Assumption 3.1** *The individual effects  $\theta_i$  are correlated with the  $x$ ’s in the following way*

$$\theta_i = \lambda \bar{x}_i + \xi_i \quad (11)$$

where  $E\{\xi_i | x_{it}\} = 0$  for all  $t = 1, \dots, T$  and  $V\{\xi_i\} = \sigma_\xi^2$ .

Then, under Assumptions 2.2 and 3.1,  $\lambda = 0$  is a sufficient condition for consistency of  $\hat{\beta}_W$  as in that case the cohort effects  $\bar{\theta}_{ct}$  in (2) are uncorrelated with the regressors. Cohorts are assumed to be constructed in the following way.

**Assumption 3.2** *Cohorts are defined on the basis of an absolute continuous distributed variable  $z$  which is distributed independently across individuals with variance normalized to unity. Moreover, the cohorts are chosen such that the (unconditional) probability of being in a particular cohort is the same for all cohorts.*

According to this assumption the support of the density of  $z$  is split into  $C$  intervals with equal probability mass, implying that all cohorts have approximately the same number of members in the sample. In practice, the variable

$z$  may be based on more than one underlying variable. It should be noted that the choice of  $z$  (or the underlying variables) is restricted. First,  $z_i$  should be constant over time for each individual  $i$  because individuals cannot move from one cohort to another. Second,  $z_i$  should be observed for all individuals in the sample. The latter requirement rules out variables like “wage earnings in 1988” or “family size at January, 1st, 1990”, because these variables are typically not observed for all individuals in the sample. In applications variables like date of birth or sex will be chosen to define the cohorts.

For Assumptions 2.1 and 2.2 to be satisfied it is required that the true cohort means vary over cohorts and/or over time. To model this, we assume that the correlation between  $x_{it}$  and  $z_i$  (on an individual level) is of the following form.

**Assumption 3.3** *The regressor variable  $x_{it}$  is correlated with  $z_i$  in the following fashion*

$$x_{it} = \mu_t + \gamma_t z_i + v_{it} \quad (12)$$

where  $v_{it}$  is uncorrelated with  $z_i$ , has expectation zero and (for the sake of simplicity) a constant variance  $\sigma_v^2$ . Its correlation over time is characterized by  $E\{v_{it}v_{is}\} = \rho\sigma_v^2$  if  $s \neq t$ . The  $\mu_t$  are fixed (unknown) constants (fixed time effects).

This assumption implies that  $v_{it}$  has the commonly assumed error components structure with an individual specific effect. The results can easily be generalized to, for example, the case where  $E\{v_{it}v_{is}\} = \rho_{|t-s|}\sigma_v^2$  ( $s \neq t$ ). In this case  $1 + (T-1)\rho$  in the expressions below should be replaced by  $1 + 2\sum_{k=1}^T \frac{T-k}{T}\rho_k$ .

It can be shown (see Appendix) that under Assumptions 2.2, 3.1, 3.2 and 3.3 the asymptotic bias of the within estimator  $\hat{\beta}_W$  is given by

$$\text{plim}_{C \rightarrow \infty}(\hat{\beta}_W - \beta) = \lambda \left[ \frac{1 + (T-1)\rho}{T} \right] \frac{\tau\omega_2}{\omega_1 + \tau\omega_2} = \delta, \quad (13)$$

where  $\tau = (T-1)/T$ ,  $\omega_1$  is the true within cohort variance defined as<sup>2</sup>

$$\omega_1 = \lim_{C \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (x_{ct}^* - \bar{x}_c^*)^2. \quad (14)$$

<sup>2</sup>The true cohort means are treated here as fixed but unknown constants.

with  $\bar{x}_c^* = \frac{1}{T} \sum_{t=1}^T x_{ct}^*$ , and where  $\omega_2$  is the measurement error variance in  $\bar{x}_{ct}$ , i.e.

$$\omega_2 = \text{plim}_{C \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - x_{ct}^*)^2 = n_c^{-1} \sigma_v^2, \quad (15)$$

where  $n_c$  is the number of individuals in each cohort  $(N/C)^3$ . Using (12) we can write

$$\omega_1 = \frac{1}{T} \sum_{t=1}^T \left( \mu_t - \frac{1}{T} \sum_{s=1}^T \mu_s \right)^2 + \frac{1}{T} \sum_{t=1}^T \left( \gamma_t - \frac{1}{T} \sum_{s=1}^T \gamma_s \right)^2. \quad (16)$$

Under Assumption 3.3 it can be easily checked that Assumption 2.1 implies that  $\omega_1 > 0$ , while Assumption 2.2 implies that  $\omega_1 + \tau\omega_2 > 0$ . The choice of the cohort identifying variable  $z_i$  determines the values of  $\mu_t$ ,  $\gamma_t$ ,  $\sigma_v^2$  and  $\rho$ . Thus, the choice of  $z_i$  determines  $\omega_1$ . Note that  $\omega_1 > 0$  requires that  $\mu_t$  or  $\gamma_t$  vary with  $t$ . If this is not the case the probability limit of Deaton's errors-in-variables estimator does not exist, while the bias in the within estimator is maximal, i.e.

$$\text{plim}_{C \rightarrow \infty} (\hat{\beta}_W - \beta) = \lambda \left[ \frac{1 + (T-1)\rho}{T} \right] = \delta_{max}, \quad (17)$$

which is independent of the cohort sizes. The choice of larger cohorts (decreasing  $\omega_2$ ) will reduce the bias if  $\omega_1 > 0$  only. Because  $\omega_2$  is a decreasing function of  $n_c$  the bias in the within estimator is smallest if the number of observations in each cohort is as large as possible.

In Table 1 we present some numerical results on the bias in  $\hat{\beta}_W$  as a fraction of the maximum bias  $\delta_{max}$  for several values of  $n_c$  and  $\omega_1/\sigma_v^2$ . We can see from the table that for the chosen values of  $\omega_1/\sigma_v^2$  the number of observations per cohort ( $n_c$ ) should be fairly large for the bias of the within estimator to be a small fraction of the maximum bias  $\lambda(1 + (T-1)\rho)/T$ . For example, if the cohorts have 100 or more observed members each, the bias is 1-16 % of the maximum bias. If the chosen values of  $\omega_1/\sigma_v^2$  are relevant for practical situations, this finding more or less justifies the fact that in most empirical studies (see, e.g., Browning, Deaton and Irish [1985] or Blundell, Browning and Meghir [1989]) the measurement errors are ignored and the

---

<sup>3</sup>If cohort sizes are unequal the observations should be reweighted first by the square root of the cohort size, as in Deaton [1985].

$n_c$	$T = 4$				$T = 12$			
	$\omega_1/\sigma_v^2 =$				$\omega_1/\sigma_v^2 =$			
	0.05	0.10	0.20	0.50	0.05	0.10	0.20	0.50
2	0.882	0.789	0.652	0.429	0.902	0.821	0.696	0.478
5	0.750	0.600	0.429	0.231	0.786	0.647	0.478	0.268
10	0.600	0.429	0.273	0.130	0.647	0.478	0.314	0.155
25	0.375	0.231	0.130	0.057	0.423	0.268	0.155	0.068
50	0.231	0.130	0.070	0.029	0.268	0.155	0.084	0.035
75	0.167	0.091	0.048	0.020	0.196	0.109	0.058	0.024
100	0.130	0.070	0.036	0.015	0.155	0.084	0.044	0.018
150	0.091	0.048	0.024	0.010	0.109	0.058	0.030	0.012
200	0.070	0.076	0.018	0.007	0.084	0.044	0.022	0.009

Table 1: Bias in the standard within estimator  $\hat{\beta}_W$  as a fraction of the maximum bias  $\delta_{max}$

standard within estimator is used. It is important to note that cohort sizes may be chosen smaller if the cohort identifying variable is chosen in such a way that the true within cohorts variance is large relative to  $\sigma_v^2$ .

## 4 The effects of the choice of cohorts on the variance

In the previous section we have shown that the bias in the within estimator from the synthetic panel may be small if the number of observations per cohort is sufficiently large. However, an increase in the number of observations per cohort implies a decrease in the number of observations in the synthetic panel ( $CT$ ) and – consequently – an increase in the variance of  $\hat{\beta}_W$ . In this section we will analyze the impact of the choice on the number of cohorts on this variance in more detail. Moreover, we show that the difference between the true variance of  $\hat{\beta}_W$  and the probability limit of its routinely estimated variance is a function of the bias only.

The asymptotic variance of  $\hat{\beta}_W$  can be written as

$$V\{\hat{\beta}_W\} = \frac{1}{CT}(\omega_1 + \tau\omega_2)^{-2}V^* \quad (18)$$

where

$$V^* = \lim_{C \rightarrow \infty} V \left\{ \frac{1}{\sqrt{CT}} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)(\bar{\theta}_{ct} - \bar{\theta}_c + \bar{\varepsilon}_{ct} - \bar{\varepsilon}_c) \right\}. \quad (19)$$

It should be noted that the expression within curved braces in (19) does not have expectation zero, because of the inconsistency of the estimator (if  $\lambda \neq 0$ ). Moreover, the summations over  $c$  and  $t$  are neither summations over independently nor identically distributed variables. This complicates elaboration of the expression in (19). In the Appendix it is shown that under the additional assumption that  $\bar{x}_{ct}$ ,  $\bar{\theta}_{ct}$  and  $\bar{\varepsilon}_{ct}$  are normally distributed, the variance of  $\hat{\beta}_W$  is given by

$$V\{\hat{\beta}_W\} = \frac{1}{CT} [(\sigma_\theta^2 + \sigma_\varepsilon^2 n_c^{-1})(\omega_1 + \tau\omega_2)^{-1} + \delta^2] \quad (20)$$

where  $\delta$  is the asymptotic bias of the within estimator defined in (13), and

$$\sigma_\theta^2 = \sigma_\xi^2 n_c^{-1} + \lambda^2 \left[ \frac{1 + (T-1)\rho}{T} \right] \omega_2, \quad (21)$$

which is the variance of  $\bar{\theta}_{ct} - \theta_c^*$ .

An increase in the cohort sizes  $n_c$  influences the variance of the within estimator  $\hat{\beta}_W$  in two ways. First, the measurement error variance  $\omega_2$  and the equation error variance  $\sigma_\theta^2 + \sigma_\varepsilon^2 n_c^{-1}$  are reduced. Second, the total number of observations  $CT$  is decreased. The latter effect is dominant, so an increase in  $n_c$  will cause a decrease in the variance of the within estimator on the synthetic panel. We will present some numerical results in the next section.

If standard software is used to compute  $\hat{\beta}_W$ , the routinely computed estimator of the variance,

$$\hat{V}\{\hat{\beta}_W\} = \hat{\sigma}^2 \left[ \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)^2 \right]^{-1} \quad (22)$$

will not be consistent for  $V\{\hat{\beta}_W\}$  in (20). In general, it converges to

$$\bar{V}\{\hat{\beta}_W\} = \bar{\sigma}^2 \frac{1}{CT} (\omega_1 + \tau\omega_2)^{-1} \quad (23)$$

where  $\bar{\sigma}^2 = \text{plim}_{C \rightarrow \infty} \hat{\sigma}^2$ . One can see from this expression that the true variance of the within estimator is underestimated by the routinely computed variance in two ways. First, the dependence of the errors and the explanatory variables invalidates the standard formula for the variance. This is reflected by the second term in (20). Second, the estimator for the error variance,

$$\hat{\sigma}^2 = \frac{1}{C(T-1)} \sum_{c=1}^C \sum_{t=1}^T [(\bar{y}_{ct} - \bar{y}_c) - \hat{\beta}_W(\bar{x}_{ct} - \bar{x}_c)]^2. \quad (24)$$

will underestimate the true error variance  $\sigma^2 = (\sigma_\xi^2 + \sigma_\varepsilon^2)n_c^{-1}$ , in particular

$$\text{plim}_{C \rightarrow \infty} \hat{\sigma}^2 = \bar{\sigma}^2 = \sigma_\varepsilon^2 n_c^{-1} + \sigma_\theta^2 - \delta^2 (\omega_1 + \tau \omega_2)^{-1}. \quad (25)$$

Note that both aspects work in the same direction. Using (25) the probability limit of the estimated variance of  $\hat{\beta}_W$  can be written as

$$\bar{V}\{\hat{\beta}_W\} = \frac{1}{CT} [(\sigma_\theta^2 + \sigma_\varepsilon^2 n_c^{-1})(\omega_1 + \tau \omega_2)^{-1} - \delta^2]. \quad (26)$$

As will be clear from the formulae above, the difference between the true variance and the probability limit of the estimated variance equals  $2\delta^2/CT$  so it will be small if the bias  $\delta$  is small.

## 5 An empirical illustration

In this section we consider the implications of the results in the previous sections on the estimation of Engel curves for food expenditures of Dutch households. We use a monthly panel data set to analyze what the properties of the within estimator on a synthetic panel would have been if one would analyze a series of repeated cross sections instead of a panel. The data used are the 367 complete monthly observations for 1986 in the so-called Expenditure Index Panel conducted by INTOMART, a marketing research agency in the Netherlands.

The model which is analyzed is the Engel curve for consumer expenditures on food,

$$w_{it} = \beta \log x_{it} + \theta_i + \varepsilon_{it}, \quad t = 1, \dots, 12, \quad (27)$$

where  $w_{it}$  is the budget share of food (in total expenditures on non-durables) and  $\log x_{it}$  is the natural logarithm of total expenditures on non-durables. The individual effects  $\theta_i$  reflect the influence of household specific characteristics (age, education, family size, etcetera) that are constant over the sample period (12 months). Obviously, these variables are likely to be correlated with total expenditures on non-durables and a fixed effects treatment of the  $\theta_i$  is desired. As in the previous sections we shall impose Assumption 3.1,

$$\theta_i = \overline{\lambda \log x_i} + \xi_i. \quad (28)$$

The construction of the cohorts will be based on the date of birth of the head of the household, as in many applied studies (see, e.g., Browning, Deaton and Irish [1985]). Because the relationship between age and total expenditures is likely to be nonlinear we choose the cohort identifying variable  $z_i$  as a quadratic function of the deviation of individual  $i$ 's date of birth from the average date of birth in the sample (in years and months). The variance of  $z_i$  is normalized to one. Under Assumption 3.3 it holds that

$$\log x_{it} = \mu_t + \gamma_t z_i + v_{it}. \quad (29)$$

Using the 367 household observations of the balanced sub-panel, we easily obtain consistent estimates of the model parameters using ordinary least squares, which are given in Table 2. All estimated  $\gamma_t$ 's are negative implying that (in a given period) total expenditures on non-durables are maximal at the average age of 49.2. Although all  $\gamma_t$ 's and  $\mu_t$ 's differ significantly from zero, the variation in the  $\gamma_t$ 's and  $\mu_t$ 's (reflected in  $\omega_1 = 0.00681$ ) is relatively small in comparison with the estimated variance of  $v_{it}$ . Although the dependence of age and total expenditures is significantly large, there does not seem to be much time variation in this dependence. Particularly for Deaton's errors-in-variables estimator this is something to worry about because its variance is inversely related with  $\omega_1$ . Of course, the small variation in the  $\gamma_t$ 's and  $\mu_t$ 's may be caused by the fact that we are using monthly data.

Before we discuss the consequences of these parameter values, we present some specification tests. First, we shall test the functional form of (27) by testing whether  $x_{it}$  (total expenditures on non-durables) should be included in (27). Subsequently we do the same for the triple  $x_{it}$ ,  $x_{it}^2$  and  $\sqrt{x_{it}}$ . This results in values for the Lagrange Multiplier test statistics of 2.75 and 7.83,



$\beta$	-0.188 (0.006)	$\mu_1$	12.235 (0.041)	$\gamma_1$	-0.147 (0.028)
$\lambda$	0.110 (0.007)	$\mu_2$	12.085 (0.041)	$\gamma_2$	-0.132 (0.028)
$\sigma_\xi$	0.105	$\mu_3$	12.202 (0.037)	$\gamma_3$	-0.164 (0.026)
$\sigma_\varepsilon$	0.072	$\mu_4$	12.238 (0.041)	$\gamma_4$	-0.150 (0.028)
$\sigma_v^2$	0.305	$\mu_5$	12.270 (0.043)	$\gamma_5$	-0.170 (0.030)
$\rho$	0.634	$\mu_6$	12.165 (0.041)	$\gamma_6$	-0.156 (0.028)
		$\mu_7$	12.161 (0.046)	$\gamma_7$	-0.156 (0.022)
$\omega_1$	0.00681	$\mu_8$	12.152 (0.042)	$\gamma_8$	-0.139 (0.029)
		$\mu_9$	12.180 (0.039)	$\gamma_9$	-0.154 (0.027)
		$\mu_{10}$	12.328 (0.042)	$\gamma_{10}$	-0.162 (0.029)
		$\mu_{11}$	12.224 (0.043)	$\gamma_{11}$	-0.181 (0.030)
		$\mu_{12}$	12.385 (0.048)	$\gamma_{12}$	-0.233 (0.033)

Table 2: Parameter estimates based on 367 observations from the balanced sub-panel (standard errors – if computed – in parentheses)

respectively. Comparing these numbers with the critical values of a Chi-square distribution with one and three degrees of freedom, respectively, we do not take them as evidence against the null. Furthermore, we test Assumption 3.3, in particular the structure of the variance covariance matrix of  $v_{it}$ . We perform the (pseudo) LM test against first order autocorrelation, as discussed in Nijman and Verbeek [1990, Appendix], which yields a value of 0.057, clearly implying that we cannot reject our null hypothesis. Apparently, the error components structure imposed on  $v_{it}$  fits the data very well. In summary, we may conclude that our model, though far simpler than many related models discussed in the literature, is not evidently in conflict with the data.

From (17) we immediately obtain that the maximum bias in the within estimator based on cohort data over 12 periods equals 0.0731, which is 39% of the (estimated) true value. Given our choice of the cohort identifying variables it is possible to eliminate some of this bias by choosing large cohorts. This is illustrated in Table 3, where the theoretical biases in the within estimator are given for several cohort sizes. Note that the bias decreases slowly with the cohort size. In the table also the probability limit of the estimated standard error is given (based on (23)) and the true standard error (based on

$n_c$	bias (in %)		plim est. st. error/ $\sqrt{N}$	true st. error/ $\sqrt{N}$
2	0.0695	37.0	0.099	0.124
5	0.0650	34.6	0.152	0.171
10	0.0586	31.2	0.205	0.220
25	0.0453	24.1	0.287	0.298
50	0.0329	17.5	0.348	0.356
75	0.0258	13.7	0.379	0.386
100	0.0212	11.3	0.398	0.404
150	0.0157	8.3	0.420	0.424
200	0.0124	6.6	0.433	0.436

Table 3: Bias in the standard within estimator  $\hat{\beta}_W$ , plim of estimated standard error and true standard error

(20)). Both are based on the asymptotic distribution. Although the bias is substantial the differences in these two standard errors are fairly small. Note that both standard errors increase if the cohort sizes are increased, which is caused by the fact that the number of (cohort) observations decreases if the cohort sizes are increased. Although there is the counteracting effect that the observations are more precise (contain less measurement error) if the cohort sizes are large, this effect is almost negligible.

Our empirical illustration in this section draws attention to the fact that in practice it may be the case that the common choice for the cohort identifying variable (date of birth) does not lead to a sufficiently large  $\omega_1/\sigma_v^2$ . If this is the case the bias in the standard within estimator is substantial, even if the number of individuals in each cohort is very large. For exactly the same reason, Deaton's errors-in-variables estimator is not a good alternative, because its variance is inversely related with  $\omega_1$ . In practice one would therefore hope to find a better cohort identifying variable  $z_i$ , or that the correlation between individual effects and explanatory variables ( $\lambda$ ) is small.

## 6 Concluding remarks

In this paper we analyzed the validity of treating cohort data as genuine panel data. Because the observed cohort averages are error-ridden measurements of the true cohort means, in general errors-in-variables estimators are required to obtain consistent estimators. If the individual effects and the explanatory variables in the model are correlated, a bias will occur in the standard fixed effects estimator, which will only be small if the number of observations in each cohort is large and if the time variation in the true cohort means is sufficiently large. To illustrate this we used genuine panel data on consumer expenditures to calibrate the possible magnitude of bias from using the synthetic panel data. The results show that in practice fairly large cohort sizes (100, 200 individuals) are needed to validly ignore the cohort nature of the data.

In summary, one can say that there is no guarantee that the bias from ignoring the measurement errors and treating the synthetic panel as if it were a true panel is small and negligible if the number of observations in each cohort is quite large. Only if there is enough time variation in the true cohort means of the explanatory variables, the bias may be negligible for reasonable cohort sizes (100 individuals). If there is no time variation at all ( $\omega_1 = 0$ ), the bias in the within estimator is bounded, while the errors-in-variables estimator proposed by Deaton [1985] has a non-existing probability limit.

## Appendix. Some technical details

In this appendix we sketch the derivation of (13) and (20). Using (12) we can write for the observed cohort means (in an obvious notation)

$$\bar{x}_{ct} = \mu_t + \gamma_t \bar{z}_{ct} + \bar{v}_{ct} = \mu_t + \gamma_t z_c^* + \bar{v}_{ct}^* = x_{ct}^* + \bar{v}_{ct}^* \quad (30)$$

where

$$z_c^* = E\{z_i \mid i \text{ is a member of cohort } c\} \quad (31)$$

and

$$\bar{v}_{ct}^* = \bar{v}_{ct} + \gamma_t(\bar{z}_{ct} - z_c^*). \quad (32)$$

Furthermore, it follows from Assumption 3.1 for the aggregated individual effects  $\bar{\theta}_{ct}$  that

$$\bar{\theta}_{ct} = \lambda \frac{1}{T} (\bar{x}_{c1}^t + \bar{x}_{c2}^t + \dots + \bar{x}_{cT}^t) + \bar{\xi}_c, \quad (33)$$

where  $\bar{x}_{cs}^t$  is the average  $x$ -value in period  $s$  of all individuals observed in period  $t$  in cohort  $c$ . Notice that  $\bar{x}_{cs}^t$  is also an error-ridden measurement of  $x_{cs}^*$ , with the same properties as  $\bar{x}_{ct}$  except that it is not observed. To be able to derive the probability limit of  $\hat{\beta}_W$  we need expressions for the following probability limits.

$$\text{plim}_{C \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)^2 \quad (34)$$

and

$$\text{plim}_{C \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c) (\bar{\theta}_{ct} - \bar{\theta}_c). \quad (35)$$

For the evaluation of (34) we use that<sup>4</sup>

$$\begin{aligned} & E \left\{ \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)^2 \right\} = \\ & \frac{1}{T} \sum_{t=1}^T \left( \mu_t - \frac{1}{T} \sum_{s=1}^T \mu_s \right)^2 + \frac{1}{T} \sum_{t=1}^T \left( \gamma_t - \frac{1}{T} \sum_{s=1}^T \gamma_s \right)^2 \left( \frac{1}{C} \sum_{c=1}^C z_c^{*2} \right) \\ & + \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T E \left\{ \left( \bar{v}_{ct} - \frac{1}{T} \sum_{s=1}^T \bar{v}_{cs} \right)^2 \right\} \\ & + \tau n_c^{-1} \frac{1}{T} \sum_{t=1}^T \gamma_t^2 \frac{1}{C} \sum_{c=1}^C V \{ z_i \mid i \text{ in cohort } c \}, \end{aligned} \quad (36)$$

where  $V \{ z_i \mid i \text{ in cohort } c \}$  is the variance of  $z_i$  within cohort  $c$ . Because the total variance of  $z$  equals unity, increasing the number of cohorts implies that the distribution of  $z_c^*$  more and more resembles the distribution of  $z_i$ . Thus, the variance of  $z$  between the  $C$  cohorts satisfies

$$\lim_{C \rightarrow \infty} \frac{1}{C} \sum_{c=1}^C z_c^{*2} = 1 \quad (37)$$

---

<sup>4</sup>Convergence follows from applying Chebychev's weak law of large numbers.

while

$$\lim_{C \rightarrow \infty} \frac{1}{C} \sum_{c=1}^C V\{z_i \mid i \text{ in cohort } c\} = 1 - \lim_{C \rightarrow \infty} \frac{1}{C} \sum_{c=1}^C z_c^{*2} = 0. \quad (38)$$

Using these equalities one can easily derive that

$$\text{plim}_{C \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c)^2 = \omega_1 + \tau n_c^{-1} \sigma_v^2 = \omega_1 + \tau \omega_2 \quad (39)$$

For the derivation of (35) we use that Assumptions 3.2 and 3.3 imply

$$\lim_{C \rightarrow \infty} E \left\{ \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{cs}^t - x_{cs}^*) (\bar{x}_{cj}^t - x_{cj}^*) \right\} = \rho \omega_2, \quad j \neq s. \quad (40)$$

Now straightforward algebra shows that

$$\text{plim}_{C \rightarrow \infty} \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c) (\bar{\theta}_{ct} - \bar{\theta}_c) = \lambda [T^{-1} + \tau \rho] \tau \omega_2, \quad (41)$$

which proves (13).

To derive the variance of  $\hat{\beta}_W$  we have to elaborate (19). Under the normality assumption of  $\bar{x}_{ct}$ ,  $\bar{\theta}_{ct}$  and  $\bar{\varepsilon}_{ct}$  the required fourth order moments can be written as functions of second order moments. In particular,

$$\begin{aligned} & V \left\{ \frac{1}{\sqrt{CT}} \sum_{c=1}^C \sum_{t=1}^T (\bar{x}_{ct} - \bar{x}_c) (\bar{\theta}_{ct} - \bar{\theta}_c + \bar{\varepsilon}_{ct} - \bar{\varepsilon}_c) \right\} = \\ & \frac{1}{CT} \sum_{c,d=1}^C \sum_{s,t=1}^T \left[ E\{\tilde{x}_{ct} \tilde{x}_{ds}\} \left( E\{\tilde{\theta}_{ct} \tilde{\theta}_{ds}\} + E\{\tilde{\varepsilon}_{ct} \tilde{\varepsilon}_{ds}\} \right) + E\{\tilde{x}_{ct} \tilde{\theta}_{ds}\} E\{\tilde{x}_{ds} \tilde{\theta}_{ct}\} \right], \end{aligned} \quad (42)$$

where  $\tilde{x}_{ct} = \bar{x}_{ct} - \bar{x}_c$  and analogously for the other variables. Using straightforward algebra one can derive the following equalities.

$$\begin{aligned} & \frac{T-1}{T} \left[ (\sigma_\xi^2 n_c^{-1} + \lambda^2 [T^{-1} + \tau \rho] \omega_2) \right] & \text{if } s = t, d = c \\ E\{\tilde{\theta}_{ct} \tilde{\theta}_{ds}\} = & -\frac{1}{T} \left[ (\sigma_\xi^2 n_c^{-1} + \lambda^2 [T^{-1} + \tau \rho] \omega_2) \right] & \text{if } s \neq t, d = c \\ & 0 & \text{elsewhere,} \end{aligned} \quad (43)$$

and,

$$E\{\tilde{x}_{ct}\tilde{\theta}_{ds}\} = \begin{cases} \frac{T-1}{T}\lambda[T^{-1} + \tau\rho]\omega_2 & \text{if } s = t, d = c \\ -\frac{1}{T}\lambda[T^{-1} + \tau\rho]\omega_2 & \text{if } s \neq t, d = c \\ 0 & \text{elsewhere.} \end{cases} \quad (44)$$

Using these equalities the variance  $V^*$  is readily obtained.

## References

- Blundell R, Browning M, Meghir C (1989) A Microeconomic Model of Intertemporal Substitution and Consumer Demand. Discussion Paper in Economics 89-11, University College London
- Browning M, Deaton A, Irish M (1985) A Profitable Approach to Labor Supply and Commodity Demands over the Life Cycle, *Econometrica*, 53: 503-543
- Deaton A (1985) Panel Data from Time Series of Cross-Sections, *Journal of Econometrics*, 30: 109-126
- Heckman JJ, MaCurdy ThE (1980) A Life Cycle Model of Female Labour Supply, *Review of Economic Studies*, 47: 47-74
- Heckman JJ, Robb R (1985) Alternative Models for Evaluating the Impact of Interventions: an Overview, *Journal of Econometrics*, 30: 239-267
- Hsiao, C (1985) Benefits and Limitations of Panel Data, *Econometric Reviews*, 4: 121-174
- MaCurdy, ThE (1981) An Empirical Model of Labor Supply in a Life-Cycle Setting, *Journal of Political Economy*, 89: 1059-1085
- Moffitt, R (1990) Estimating Dynamic Models with a Time Series of Repeated Cross Sections, mimeo, Brown University, Providence RI
- Nijman, ThE, Verbeek M (1990) Estimation of Time Dependent Parameters in Linear Models Using Cross Sections, Panels or Both, *Journal of Econometrics*, 46, forthcoming

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
8938	G. Marini	Monetary Shocks and the Nominal Interest Rate
8939	W. Güth and E. van Damme	Equilibrium Selection in the Spence Signaling Game
8940	G. Marini and P. Scaramozzino	Monopolistic Competition, Expected Inflation and Contract Length
8941	J.K. Dagsvik	The Generalized Extreme Value Random Utility Model for Continuous Choice
8942	M.F.J. Steel	Weak Exogeneity in Misspecified Sequential Models
8943	A. Roell	Dual Capacity Trading and the Quality of the Market
8944	C. Hsiao	Identification and Estimation of Dichotomous Latent Variables Models Using Panel Data
8945	R.P. Gilles	Equilibrium in a Pure Exchange Economy with an Arbitrary Communication Structure
8946	W.B. MacLeod and J.M. Malcomson	Efficient Specific Investments, Incomplete Contracts, and the Role of Market Alternatives
8947	A. van Soest and A. Kapteyn	The Impact of Minimum Wage Regulations on Employment and the Wage Rate Distribution
8948	P. Kooreman and B. Melenberg	Maximum Score Estimation in the Ordered Response Model
8949	C. Dang	The $D_3$ -Triangulation for Simplicial Deformation Algorithms for Computing Solutions of Nonlinear Equations
8950	M. Cripps	Dealer Behaviour and Price Volatility in Asset Markets
8951	T. Wansbeek and A. Kapteyn	Simple Estimators for Dynamic Panel Data Models with Errors in Variables
8952	Y. Dai, G. van der Laan, D. Talman and Y. Yamamoto	A Simplicial Algorithm for the Nonlinear Stationary Point Problem on an Unbounded Polyhedron
8953	F. van der Ploeg	Risk Aversion, Intertemporal Substitution and Consumption: The CARA-LQ Problem
8954	A. Kapteyn, S. van de Geer, H. van de Stadt and T. Wansbeek	Interdependent Preferences: An Econometric Analysis

No.	Author(s)	Title
8955	L. Zou	Ownership Structure and Efficiency: An Incentive Mechanism Approach
8956	P.Kooreman and A. Kapteyn	On the Empirical Implementation of Some Game Theoretic Models of Household Labor Supply
8957	E. van Damme	Signaling and Forward Induction in a Market Entry Context
9001	A. van Soest, P. Kooreman and A. Kapteyn	Coherency and Regularity of Demand Systems with Equality and Inequality Constraints
9002	J.R. Magnus and B. Pesaran	Forecasting, Misspecification and Unit Roots: The Case of AR(1) Versus ARMA(1,1)
9003	J. Driffill and C. Schultz	Wage Setting and Stabilization Policy in a Game with Renegotiation
9004	M. McAleer, M.H. Pesaran and A. Bera	Alternative Approaches to Testing Non-Nested Models with Autocorrelated Disturbances: An Application to Models of U.S. Unemployment
9005	Th. ten Raa and M.F.J. Steel	A Stochastic Analysis of an Input-Output Model: Comment
9006	M. McAleer and C.R. McKenzie	Keynesian and New Classical Models of Unemployment Revisited
9007	J. Osiewalski and M.F.J. Steel	Semi-Conjugate Prior Densities in Multivariate t Regression Models
9008	G.W. Imbens	Duration Models with Time-Varying Coefficients
9009	G.W. Imbens	An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-Based Sampling
9010	P. Deschamps	Expectations and Intertemporal Separability in an Empirical Model of Consumption and Investment under Uncertainty
9011	W. Glth and E. van Damme	Gorby Games - A Game Theoretic Analysis of Disarmament Campaigns and the Defense Efficiency-Hypothesis
9012	A. Horsley and A. Wrobel	The Existence of an Equilibrium Density for Marginal Cost Prices, and the Solution to the Shifting-Peak Problem
9013	A. Horsley and A. Wrobel	The Closedness of the Free-Disposal Hull of a Production Set



No.	Author(s)	Title
9014	A. Horsley and A. Wrobel	The Continuity of the Equilibrium Price Density: The Case of Symmetric Joint Costs, and a Solution to the Shifting-Pattern Problem
9015	A. van den Elzen, G. van der Laan and D. Talman	An Adjustment Process for an Exchange Economy with Linear Production Technologies
9016	P. Deschamps	On Fractional Demand Systems and Budget Share Positivity
9017	B.J. Christensen and N.M. Kiefer	The Exact Likelihood Function for an Empirical Job Search Model
9018	M. Verbeek and Th. Nijman	Testing for Selectivity Bias in Panel Data Models
9019	J.R. Magnus and B. Pesaran	Evaluation of Moments of Ratios of Quadratic Forms in Normal Variables and Related Statistics
9020	A. Robson	Status, the Distribution of Wealth, Social and Private Attitudes to Risk
9021	J.R. Magnus and B. Pesaran	Evaluation of Moments of Quadratic Forms in Normal Variables
9022	K. Kamiya and D. Talman	Linear Stationary Point Problems
9023	W. Emons	Good Times, Bad Times, and Vertical Upstream Integration
9024	C. Dang	The $D_2$ -Triangulation for Simplicial Homotopy Algorithms for Computing Solutions of Nonlinear Equations
9025	K. Kamiya and D. Talman	Variable Dimension Simplicial Algorithm for Balanced Games
9026	P. Skott	Efficiency Wages, Mark-Up Pricing and Effective Demand
9027	C. Dang and D. Talman	The $D_1$ -Triangulation in Simplicial Variable Dimension Algorithms for Computing Solutions of Nonlinear Equations
9028	J. Bai, A.J. Jakeman and M. McAleer	Discrimination Between Nested Two- and Three-Parameter Distributions: An Application to Models of Air Pollution
9029	Th. van de Klundert	Crowding out and the Wealth of Nations
9030	Th. van de Klundert and R. Gradus	Optimal Government Debt under Distortionary Taxation

No.	Author(s)	Title
9031	A. Weber	The Credibility of Monetary Target Announcements: An Empirical Evaluation
9032	J. Osiewalski and M. Steel	Robust Bayesian Inference in Elliptical Regression Models
9033	C. R. Wichers	The Linear-Algebraic Structure of Least Squares
9034	C. de Vries	On the Relation between GARCH and Stable Processes
9035	M.R. Baye, D.W. Jansen and Q. Li	Aggregation and the "Random Objective" Justification for Disturbances in Complete Demand Systems
9036	J. Driffill	The Term Structure of Interest Rates: Structural Stability and Macroeconomic Policy Changes in the UK
9037	F. van der Ploeg	Budgetary Aspects of Economic and Monetary Integration in Europe
9038	A. Robson	Existence of Nash Equilibrium in Mixed Strategies for Games where Payoffs Need not Be Continuous in Pure Strategies
9039	A. Robson	An "Informationally Robust Equilibrium" for Two-Person Nonzero-Sum Games
9040	M.R. Baye, G. Tian and J. Zhou	The Existence of Pure-Strategy Nash Equilibrium in Games with Payoffs that are not Quasiconcave
9041	M. Burnovsky and I. Zang	"Costless" Indirect Regulation of Monopolies with Substantial Entry Cost
9042	P.J. Deschamps	Joint Tests for Regularity and Autocorrelation in Allocation Systems
9043	S. Chib, J. Osiewalski and M. Steel	Posterior Inference on the Degrees of Freedom Parameter in Multivariate-t Regression Models
9044	H.A. Keuzenkamp	The Probability Approach in Economic Methodology: On the Relation between Haavelmo's Legacy and the Methodology of Economics
9045	I.M. Bomze and E.E.C. van Damme	A Dynamical Characterization of Evolutionarily Stable States
9046	E. van Damme	On Dominance Solvable Games and Equilibrium Selection Theories
9047	J. Driffill	Changes in Regime and the Term Structure: A Note

No.	Author(s)	Title
9048	A.J.J. Talman	General Equilibrium Programming
9049	H.A. Keuzenkamp and F. van der Ploeg	Saving, Investment, Government Finance and the Current Account: The Dutch Experience
9050	C. Dang and A.J.J. Talman	The $D_1$ -Triangulation in Simplicial Variable Dimension Algorithms on the Unit Simplex for Computing Fixed Points
9051	M. Baye, D. Kovenock and C. de Vries	The All-Pay Auction with Complete Information
9052	H. Carlsson and E. van Damme	Global Games and Equilibrium Selection
9053	M. Baye and D. Kovenock	How to Sell a Pickup Truck: "Beat-or-Pay" Advertisements as Facilitating Devices
9054	Th. van de Klundert	The Ultimate Consequences of the New Growth Theory; An Introduction to the Views of M. Fitzgerald Scott
9055	P. Kooreman	Nonparametric Bounds on the Regression Coefficients when an Explanatory Variable is Categorized
9056	R. Bartels and D.G. Fiebig	Integrating Direct Metering and Conditional Demand Analysis for Estimating End-Use Loads
9057	M.R. Veall and K.F. Zimmermann	Evaluating Pseudo- $R^2$ 's for Binary Probit Models
9058	R. Bartels and D.G. Fiebig	More on the Grouped Heteroskedasticity Model
9059	F. van der Ploeg	Channels of International Policy Transmission
9060	H. Bester	The Role of Collateral in a Model of Debt Renegotiation
9061	F. van der Ploeg	Macroeconomic Policy Coordination during the Various Phases of Economic and Monetary Integration in Europe
9062	E. Bennett and E. van Damme	Demand Commitment Bargaining: - The Case of Apex Games
9063	S. Chib, J. Osiewalski and M. Steel	Regression Models under Competing Covariance Matrices: A Bayesian Perspective
9064	M. Verbeek and Th. Nijman	Can Cohort Data Be Treated as Genuine Panel Data?

P.O. BOX 90153 5000 LE TILBURG THE NETHERLANDS

**Bibliotheek K. U. Brabant**



17 000 01117534 7