S E D A P

A PROGRAM FOR RESEARCH ON

SOCIAL AND ECONOMIC DIMENSIONS OF AN AGING POPULATION

A Longitudinal Study of the Residential Mobility of the Elderly in Canada

Yuri Ostrovsky

SEDAP Research Paper No. 78

For further information about SEDAP and other papers in this series, see our web site:
http://socserv2.mcmaster.ca/sedap

Requests for further information may be addressed to:
Secretary, SEDAP Research Program
Kenneth Taylor Hall, Room 426
McMaster University
Hamilton, Ontario, Canada
L8S 4M4
FAX: 905 521 8232
e-mail: qsep@mcmaster.ca

A LONGITUDINAL STUDY OF THE RESIDENTIAL MOBILITY OF THE ELDERLY IN CANADA

YURI OSTROVSKY

SEDAP Research Paper No. 78

June 2002

The Program for Research on Social and Economic Dimensions of an Aging Population (SEDAP) is an interdisciplinary research program centred at McMaster University with participants at the University of British Columbia, Queen's University, Univérsité de Montréal, and the University of Toronto. It has support from the Social Sciences and Humanities Research Council of Canada under the Major Collaborative Research Initiatives Program, and further support from Statistics Canada, the Canadian Institute for Health Information, and participating universities. The SEDAP Research Paper series provides a vehicle for distributing the results of studies undertaken by those associated with the program. Authors take full responsibility for all expressions of opinion.

A Longitudinal Study of the Residential Mobility of the Elderly in Canada

Yuri Ostrovsky*
York University

May 15, 2002

Abstract

An intensely debated question in the lifecycle literature is whether housing wealth is viewed by households as a financial asset that will be used to support general consumption after retirement. This paper uses the newly available longitudinal Canadian Survey of Labour and Income Dynamics (SLID) to investigate the factors influencing elderly households' residential mobility choices. A dynamic non-linear panel (longitudinal) data dynamic model is employed. I use the Bover-Arellano estimator (Chamberlain's class of estimators), based on reduced form predictions of the latent dependent variable. The residential mobility of the elderly appears to be affected mostly by moving costs, which are different for owners and non-owners.

*I thank Vincent Hildebrand, Sung-Hee Jeon, Gail Kalika, Barry Smith, Byron Spencer and Michael Veall for valuable suggestions and comments. I owe very special thanks to my supervisor Tom Crossley for encouragement, advice and many helpful discussions. Thanks also go to Cindy Cook and Tina Hotton for their help at the Statistics Canada Research Data Centre at McMaster University and University of Toronto. A SEDAP fellowship grant is gratefully acknowledged. All errors are my own. *E-mail address*: ostrovsk@dept.econ.yorku.ca

1 Introduction

Much of the debate about the housing decisions of the elderly concerns the role of housing wealth in the lifecycle consumption/savings decision of the household. A traditional theoretical framework for such analysis is the lifecycle theory, which in its simplest ('orthodox', certainty or certainty-equivalence) version assumes the systematic accumulation of assets during the working life and gradual decumulation during retirement. Formally, the problem is usually formulated as a discrete time dynamic programming problem in which households make sequential consumption decisions based on the information available to them at each period (Deaton [1992]). A reduction of wealth in later life has particular importance in the lifecycle literature

> "...since all standard consumption models predict that eventually households will start to dissave whether or not there is a bequest motive or uncertain lifetime. Thus this prediction can be considered a "critical experiment" for the lifecycle model in its most general form" (Browning and Lusardi [1996]).

One of the hotly debated questions is whether *housing wealth* is viewed by households as a financial asset that will be used to support general consumption after retirement. The certainty version of the lifecycle theory rests on the assumption of fungibility, which implies the equality of the marginal propensity to consume from different sources of wealth. If the assumption of fungibility holds then households view housing as a substitute for other financial assets and it should be included in the analysis of the adequacy of household savings. In this case, we should observe reduction in housing wealth after retirement unless households are unable - due to high financial and psychological costs of moving - to release their housing equity (Venti and Wise [2001]).

Earlier versions of the lifecycle theory and the assumption of fungibility have been sharply criticized by behavioral economists, who argue that households have a set of "mental accounts" with varying marginal propensities to consume for different assets (Thaler [1994]). They also argue that households view housing differently from other assets and will not use their housing wealth for general consumption in later life (Levine [1999]). If the assumption of fungibility of housing does not hold then it should be excluded from the analysis of the adequacy of household savings for retirement. The dynamics of housing markets, in this case, are a "sideshow" (Skinner [1996]) and the elderly will not generally consider moving regardless of moving costs or changes in housing prices.

It is important to point out that modern versions of the standard lifecycle theory recognize that agents may not treat all sources of wealth equally due to precautionary motive or liquidity constraints (Browning and Lusardi [1996], Browning and Crossley [2001]). The precautionary motive is particularly important in later life. Uncertainty about future medical expenses may prevent the elderly form "downsizing" (moving to a smaller house), particularly if housing is viewed as an asset of "last resort" to be used only to pay for nursing care or to support a surviving spouse (Venti and Wise [2001]).

The residential mobility (moving) of the elderly households is directly linked to the question of housing adjustment in later life. Changes in the amount of housing consumption are usually observed as a two step process: (1) a decision to move followed by (2) changes in housing consumption *conditional* on moving. In this study I concentrate on the first step and investigate the factors influencing household decisions to move in the context of a dynamic approach using panel (longitudinal) data from the newly available longitudinal Canadian Survey of Labour and Income Dynamics (SLID), which is discussed in more detail in Section 3.[1] A longitudinal approach is particularly relevant in the analysis of housing behavior over time as it reflects the evolution of household demographic, economic and lifecycle circumstances.

This is the first study of residential mobility study using Canadian panel data. Its a contribution to the international literature is that it deals explicitly with the issue of unobserved individual heterogeneity in a dynamic framework. The importance of controlling for individual effects (individual housing tastes) is emphasized by Feinstein and McFadden [1989] who test and reject the null of no individual heterogeneity in their study. They caution that their results may be seriously biased as they do not control for individual effects. Börsch-Supan and Pollakowski [1990] control for individual heterogeneity using a static fixed-effects multinomial logit model. VanderHart [1998] attempts to separate the effect of economic and non-economic variables on housing decisions using a dynamic multinomial logit model but does not explicitly control for individual heterogeneity. My approach is different from the abovementioned studies. I employ the two-step within-group and asymptotically efficient three-step GMM estimators for limited dependent variable models with unobserved individual effects developed by Bover and Arellano [1997], which are extensions to the minimum distance estimator suggested by Chamberlain [1984]. An important feature of all Chamberlain's class of estimators is that they allow for relaxing the

[1] Unfortunately, SLID does not provide information nesessary to investigate the issue of changes in housing consumption.

assumption that individual effects are uncorrelated with explanatory variables. To control for habit persistence I consider a dynamic specification. The inclusion of the lagged dependent variable usually poses considerable technical problems. The method used in this study is an attempt to avoid some of the pitfalls specific to dynamic models. The advantages of the method as well as possible drawbacks are discussed in Section 5.

Like previous authors, I find no strong support for the prediction of a simple (certainty-equivalence) lifecycle model that the residential mobility of the elderly is primarily motivated by the desire to consume out of housing wealth. In particular, one result is that *proportionally more* elderly who were *non-owners* at the beginning of the study *became owners* than vice versa. I also find that those with lower moving costs - non-owners, single people, and urban dwellers - are more likely to adjust their housing than those with higher moving costs. The results indicate a response to the changes in housing prices that is stronger than has previously been reported (Skinner [1996]), particularly for owners.

The paper is organized as follows. Section 2 offers a brief literature review. Section 3 familiarizes the reader with the structure of the Survey of Labour and Income Dynamics and describes the study sample. Section 4 presents some descriptive statistics concerning mobility. Section 5 outlines the relevant econometric issues in panel data analysis and discusses the Bover-Arellano estimator. Section 6 shows the Monte-Carlo simulations results. Section 7 describes the variables and presents the estimation results. Finally, Section 8 offers some conclusions and points to the directions for future research.

2 Literature

As mentioned above, a classical certainty ("stripped down") version of the lifecycle theory predicts a reduction in housing equity ("downsizing") in later life as a part of general wealth decumulation to keep the lifetime marginal utility of consumption constant. A desire for housing adjustment should lead to high mobility rates among the elderly, particularly among the owners, and high transition rates from ownership to renting.

In two influential studies Venti and Wise [1989] and Feinstein and McFadden [1989] found little evidence that the residential mobility of the elderly is influenced by the desire to consume out of housing equity. They also found that mobility rates do not particularly depend on the wealth of the seniors. Contrary to the logic of the lifecycle theory, those with high non-housing wealth and low housing equity are much more likely to move than those with low non-housing wealth but high housing equity (Venti and Wise [1989]). Generally, the mobility rates among the elderly are considerably lower than among the non-elderly and elderly renters are much more mobile than elderly owners (Venti and Wise [1989]). Also puzzling seems to be the finding that home equity has a negative effect on the probability that a homeowner will move or become a renter (Merrill [1984]).

While the impact of economic factors on the housing choices of the elderly appears to be muted, several studies have found that housing transitions are strongly related to *non-economic* lifecycle events, such as retirement and changes in family composition, in particular the loss of a spouse (Venti and Wise [1989], Feinstein and McFadden [1989], Ermisch and Jenkins [1999]). The elderly are interpreted to be reluctant movers who often move out of necessity rather than economic considerations. A sharp downturn in health status may leave an elderly person no choice other than to move to a health care facility or nursing home. This is in line with the fact that mobility falls with age but rises for the "older" elderly (Feinstein and McFadden [1989], Sheiner and Weil [1992], Megbolugbe *et al.* [1997]).

Browning and Lusardi [1996], and Browning and Crossley [2001] argue that the importance of non-economic variables is not inconsistent with modern versions of lifecycle theory that account for uncertainty, bequest motive, imperfect markets, and habit formation. On one hand, lifecycle changes such as changes in household size or health status have a direct effect by changing the marginal utility of housing consumption (Deaton [1992]). More importantly, however, non-economic events may have an indirect effect on housing choices by affecting future economic variables. As VanderHart [1998] put it:

> "The onset of a physical limitation will undoubtedly have a direct effect on housing decisions, but may also have an indirect effect via the expectation of higher health care costs, lower future income, faster depletion of financial assets in the future, and higher owner-occupied maintenance costs from the decontinuation of do-it-yourself repair."

Using a data sample from the Panel Study of Income Dynamics (PSID), VanderHart demonstrates that the role of economic factors such as income and financial assets for tenure transition after retirement may be more important than previously thought and that their effect may have been erroneously attributed to non-economic variables in other studies.

Health uncertainty alone may explain why the elderly are reluctant to deplete their housing wealth. A simple lifecycle model may overestimate the level of dissaving after retirement by ignoring the uncertainty caused by potentially large out-of-pocket medical expenses (Palumbo [1999]). Skinner [1996] suggests that housing wealth may play the role of a precautionary "buffer" that can be cashed out in the event of an income or health downturn, or widowhood, when the demand for housing services is likely to decline as well. Thus, it may be viewed as a form of self-insurance against retirement contingencies and potentially large out-of-pocket medical expenses, and thereby reduce the need for other precautionary savings.

Several studies have investigated the importance of moving costs - including psychological costs of separation - on older households' housing decisions. Venti and Wise [1990] argue that the high transaction cost of moving do not explain the lack of "downsizing" in later life. Feinstein [1996], on the other hand, shows that mobility costs considerably reduce the residential mobility of the elderly in response to changes in their health status. Megbolugbe *et al.* [1997] suggest that the elderly can potentially benefit from "reverse mortgage" programs that would allow homeowners to "unlock" their housing equity while continuing to live in their homes. Such programs usually offer several different options including lump sum payments, payments on a monthly basis for a fixed term, payments on a monthly basis for as long as the elderly live in their home, or as a line of credit (Fratantoni [1999]). So far there has been little demand for "reverse mortgages" although that may be explained by the lack of awareness of such programs among the elderly.

Many of these issues can not be analyzed in the context of a static cross-sectional framework. An importance of a dynamic approach in studies of residential mobility has long been recognized (Venti and Wise [1989], Feinstein and McFadden [1989], Megbolugbe *et al.* [1997], VanderHart [1998]). As Henderson and Ioannides [1987] point out, an obvious problem with static analysis is that, typically, it attributes historical choices to today's condition. Housing is a very expensive good and, for most families, housing expenditures are the largest part of total family expenditures. A transition from renting to owning that the majority of households make some time along their lifecycle often requires a large downpayment and long term financing (mortgage), so the decision to become an owner usually comes as a result of long term planning [2].

The difference between age and cohort effects underscores another important advantage of a dynamic approach. Such an approach based on longitudinal data would present an opportunity to separate age and cohort effects that are indistinguishable in the context of cross-section data (Börsch-Supan and Pollakowski [1990], Deaton [1997], Myers [1999], Dieleman [2001]). The panel data models became particularly attractive in the past decade as the speed of computing has increased tremendously and thus the handling of data has become much easier (Dieleman [2001]). The theoretical aspects of panel data models will be discussed in more detail in Section 5.

There have been several studies that introduced some dynamic element and used panel data. However, these studies were often confined to the descriptive analysis of hazard rates or tenure transitions (Venti and Wise [1989], Feinstein and McFadden [1989], Megbolugbe et al. [1997]). VanderHart [1998] attempts, in a dynamic framework, to disentangle direct and indirect effects of lifecycle events in a study of residential mobility of the elderly. He uses a conventional

[2] Lifecycle housing dynamics are often described by the concept of "housing career." Becoming an owner is one of the stages in such a career.

multinomial logit model that ignores individual effects. Börsch-Supan and Pollakowski [1990], on the other hand, use a conditional fixed-effects multinomial logit model (based on Chamberlain [1984]) which allows them to control for individual effects but is not explicitly dynamic.

I am not aware of any dynamic analysis of residential mobility among the elderly based on Canadian data. Jones [1996] studied residential mobility among the elderly in the context of a simple static model based on a single cross-sectional survey. The results based on the Canadian Family Expenditure Survey (FAMEX) suggest that households that have diminished in size (for instance, widowed homeowners) are more likely to cease homeownership. As Jones himself admits, given the small sample and the simplicity of his model, his results should be treated with caution.

3 SLID and Study Sample

3.1 SLID

SLID was designed to have rotating overlapping panels, with each new six-year panel starting half way through the life-span of the previous one. The first labor and income interviews were conducted by Statistics Canada in January and May 1994 for reference year 1993. The second six-year panel was introduced in 1996. When the first panel ended in 1999, a third one began.

The main objective of the survey is to support research on income and labor market dynamics. However, in addition to the very detailed information on income and personal and family characteristics, it also includes information on residential mobility and housing tenure. Combined with the longitudinal aspect of SLID this creates a powerful tool for the analysis of lifecycle events including residential mobility. Longitudinal respondents are followed for six year after the preliminary interview, even if they move. Of course, there are some practical constraints and operational limits to SLID's capacity to follow people. For persons moving outside Canada and the continental United States, tracing will be done to identify only those who will subsequently return. If they return interviewing is resumed.

The sample for each SLID panel is a subset of the sample selected for the Canadian Labour Force Survey (LFS). The SLID income concepts are very similar to those in the Canadian Survey of Consumer Finances (SCF), which was replaced by SLID (after five years of overlap) in 1998. Like LFS, SLID covers the population of the ten provinces (residents of Yukon and Northwest Territories are excluded) with the exception of Indian reserves. Also excluded are the residents of institutions[3] (unless under six months) and military barracks. The size of the first six-year panel is 15,000 households which includes about 31,0000 persons aged 16 and over.

Over the life span of a panel up to 13 interview are conducted. First, a preliminary interview (SLID uses computer-assisted interviewing) is conducted at the beginning of each panel to collect background information. Labor interviews are conducted every January for six years and refer to experience in previous calendar year. For example, the first annual interview took place in January 1994, and the reference period was 1993. Income interviews are conducted each May for six years in a similar manner. The income interviews are deferred until May to take advantage of income tax time (the deadline for filing tax return

[3]The concept of institutions includes childrens group homes and orphanages, nursing homes, chronic care hospitals, residents for senior citizens, hospitals, psychiatric institutions, treatment centres and institutions for the physically handicapped, correctional and penal institutions, young offenders facilities and jails.

forms in Canada is April 30). The respondents have an option of giving Statistics Canada permission to access their tax information instead of responding to detailed questions about each year's income.

To insure better representation the urban centres are divided into strata which, in turn, are divided into "clusters" constituting the primary sampling units. A sample of units in large apartment buildings is selected based on information supplied by Canada Mortgage and Housing Corporation. Primary sampling units in rural areas are selected on the basis of well-defined physical features such as rivers, roads, etc.

Proxy response is accepted in SLID. Usually only one member of the household answers questions for all members. The response rates are quite high, with the highest in Newfoundland (around 90% in the first year and 87% in the second) and the lowest in British Columbia (85% and 77%).

3.2 Study Sample

The data used in this study are drawn from the first four years (1996-1999) of the second SLID panel, which started in 1996 and will continue until 2001. It would be desirable to use a longer panel; unfortunately, some of the important variables were introduced only when the second panel started, which was 1996, so the first three years of the first panel (which began in 1993) could not be used. 1999 was the last year available at the time of the study.

Due to the overlapping structure of SLID the choice of a four year span is essentially a trade off between a longer panel and larger sample size. A three-period panel (1996-1998) would take advantage of a larger sample drawn from both the first and second panels. The decision to have a longer panel is motivated by the fact that even if the sample is drawn only from the second panel the sample size after imposing age and other restrictions is still considerable (2822 household heads), so the benefit of having an additional period in the sample outweighs the cost of reducing the sample size.

One of the issues in longitudinal studies is the choice of the unit of analysis. SLID data are collected on the individual level. For the purpose of cross-sectional studies, Statistics Canada has constructed family variables. However, the construction of family variables is more straight forward in cross-sectional studies than in longitudinal studies (Duncan and Hill [1985], Butlin [1994]). In cross-sectional studies a family or a household are identified at certain point in time. It is implicitly assumed that their composition is unchanged during the reference period.

In longitudinal studies such an assumption is implausible, since the composition of many families and households will change over the period of study. A household constructed during the first year of the panel may split into several households by the last year. Following all the "splits" may not only be extremely difficult but also misleading as their socio-economic attributes may be very different from the socio-economic attributes of the original household.

Duncan and Hill [1985] proposed an "attributional" approach. They have noticed that households can be conveniently followed by tracking, for instance, household heads and attributing to them the characteristics of the household in which they live. Changes in the household composition in this case are regarded as changes in individual's characteristics and treated as a demographic variable. This approach is recommended by SLID (Butlin, 1994) and used in this study.

At the first stage only major income earners[4] (MIE) in the households (both

[4]If more than one person has the same income, the major income earner is defined to be the oldest.

men and women) aged 60 and over in 1996 were selected. The sample consists of 1622 male and 1200 female MIEs. Table 1 presents the proportion of male and female owners (and, by implication non-owners) in the sample. There is a noticeable difference in the proportion as between households with male and female MIEs; there is a higher proportion of male owners. Over 80% of male MIEs are owners while only about 60% of female MIEs are owners. There is a small reduction in ownership among both groups by 1999 although the reduction among female MIEs is slightly more pronounced.

	1996		1997		1998		1999	
	male	female	male	female	male	female	male	female
owned	1349	752	1350	748	1346	737	1331	711
%	83.2	63.7	83.2	62.3	83.0	61.4	82.1	59.3

Table 1: Sample size and proportion of owners, by sex

Table 2 shows the income distribution in 1996. Median income was $19269. Most respondents, over 60%, were retired (Table 3). There is, however, a considerable degree of inconsistency in the answers. For example, some respondents (mostly women) gave their status as “retired” or “keeping house” interchangeably in different years.

Two hundred and thirty two household MIEs (8.22% of the sample) held university degrees. A considerable proportion of MIEs (357 or 12.65%) have never had or raised a child.

percentile	20	40	50	60	80
income	$12621	$16172	$19269	$23108	$34930

Table 2: Income Distribution in 1996

Major activity (unavailable for 1999)	1996	1997	1998
Working, looking for work or in school	435	347	287
Retired	2039	2218	2273
Other (keeping house, caring for disabled, etc.)	348	257	262

Table 3: Major Activity

4 Descriptive Analysis

SLID was designed to record the number of moves during the reference year for each individual. In practice, however, the SLID variable “nbmov” (number of times moved) is set to 1 even if the number of moves is greater than one. Thus, we know only whether a person moved at least once during the reference year. Another problem is that although the date of the (last) move is available, the date of other life cycle events that might influence the decision to move, such as the death of a spouse, are not. Hence, it is impossible to know exactly whether the individual moved before or after such event (more on this in the next section).

The number of ‘movers’ is 133, 116, 125 and 119 in 1996, 1997, 1998, and 1999 respectively. Overall, 415 (14.7%) moved at least once during the 4 year period. In the vast majority of cases the whole household moved. The average income of ‘movers’ was $22987, while the average income of ‘stayers’ was $25185. Unfortunately, SLID does not have asset information. I used the information about investment income as a proxy for the household assets. The average investment income of ‘movers’ was lower ($9823) than that of ‘stayers’ ($11883). Table 4 shows the average annual proportion of ’movers’ for five different percentile group. The mobility rate for the highest investment income group is the lowest while the mobility rate for the lowest investment income group is the highest.

Percentile group	Average annual % of 'movers' By total income	By investment income
1st (low)	4.78	6.25
2nd	4.92	3.96
3rd	4.56	3.72
4th	4.43	4.26
5th	3.14	3.42

Table 4: Income groups and average annual proportion of movers

The residents of British Columbia had the highest level of residential mobility. At least 22% of those who resided in BC have moved at least once over the 4 year period (Table 5). Nova Scotia had the lowest rate (10%). In general, the Maritime Provinces, particularly Nova Scotia and Prince Edward Island, had considerably lower mobility rates than other provinces. Most of the moves took place within a province. The overwhelming majority of 'movers' (an average of 78.6 %) moved within 50km from the place where they had lived before.

Province (t-1)	Moved at least once
Newfoundland	14%
Prince Edward Island	12%
Nova Scotia	10%
New Brunswick	16%
Quebec	19%
Ontario	18%
Manitoba	17%
Saskatchewan	18%
Alberta	16%
British Columbia	22%

Table 5: Mobility by province

Table 6 shows the relationship between residential mobility and health. Those on the periphery of the health spectrum ('excellent' or 'poor') are more likely to move than those who describe their health 'good' or 'very good'. This supports a proposition that there may be two health related reasons for mobility: those who feel 'excellent' are likely to migrate for reasons of amenity, while those who are in 'poor' health are more likely to move out of necessity (Hayward [2000]).

health in '96	excellent	very good	good	fair	poor
moved in '97 (%)	5.69	3.70	3.29	4.40	5.73
health in '97	excellent	very good	good	fair	poor
moved in '98 (%)	4.43	3.70	3.70	5.85	7.29
health in '98	excellent	very good	good	fair	poor
moved in '99 (%)	4.21	3.62	3.70	5.49	5.76

Table 6: Health and mobility

As mentioned above, a *decline* in health may play an important role in the housing choices of the elderly. I defined a decline in health status as a transition to 'poor'. Table 7 shows that approximately 5.6% of those whose health had declined in period *t-1* moved in period *t* compared to 4.3% of those whose health had not.

One of the reasons the elderly may not "downsize" is the bequest motive. Venti and Wise [1989] and Feinstein and McFadden [1989] present evidence that the bequest motive is not a major factor that causes the elderly not to withdraw wealth from housing. To assess its role I compared the mobility rates for households with or without children (Table 7). Interestingly enough, the rates are almost the same - 4.48% for the households in which the head have never had or raised a child compared 4.35% for the households with children. This result is similar to Venti and Wise [1989] who show that the change in housing equity at the time of sale by the elderly persons without children is about the same as the change for those with children.

Households in which household heads are married[5] have a considerably lower proportion of 'movers' than those in which household heads are not married. This is also in line with the suggestion that the elderly often move out of necessity, when they no longer can take care of themselves. The proportion of 'movers' among urban household is almost twice the proportion of 'movers' among rural households. I will return to the role of tenure, poor health and urban residency in the econometric analysis.

	Yes	**No**
Health declined at *t-1 (*average annual % of 'movers'*)*	5.56	4.26
Ever had or raised a child	4.35	4.48
Married at t	3.24	5.52
Urban dweller at *t-1*	4.87	2.49

Table 7: Proportion of movers with specified status

Table 8 also shows a much higher proportion of 'movers' among non-owners. The average percentage of 'movers' among owners is 2.72 compared to 8.66 percent among non-owners. The result echoes Venti and Wise [1989] who report that renters are almost three times as likely to move as owners.

	status in 1996		status in 1997		status in 1998		status in *t-1*	
moved (%)	owners	non-own.	owners	non-own.	owners	non-own.	owners	non-own.
in 1997	2.62	8.46						
in 1998			2.86	8.98				
in 1999					2.69	8.53		
in t (ave.)							2.72	8.66

Table 8: Mobility by tenure

During the preliminary interview in 1995, 75.7% of households owned their dwelling. By 1999, the proportion of owners decreased to 72.4%. However, *proportionally more non-owners have become owners* than vice versa (Table 9).

[5] Including common law partners

7.07% of those who were owners in 1995 became non-owners by 1999, while 8.31% of those who were non-owners in 1995 became owners by 1999. The higher transition rate to ownership is puzzling. If indeed the major reason for mobility in later life was the desire to "downsize" we would expect to see the opposite trend, at least to some degree. This is consistent with Venti and

		1995	1996	1997	1998	1999
All	owners	2136	2101	2098	2083	2042
	%	75.69	74.45	74.34	73.81	72.36
Owners in 1995	owners	2136	2069	2056	2031	1985
	non-owners	-	67	80	105	151
Non-owners in 1995	owners	-	32	42	52	57
	non-owners	686	654	644	634	629

Table 9: Changing Tenure

Wise [1989] who point out that the reason for this result is that the renters are much more likely to move, not that they are more likely than owners to switch tenures. Another possible explanations that a renter is more likely to drop out the sample.

Some information about the motivations for a move could be drawn from the reasons given by the respondents themselves. Unfortunately, the design of the survey question about the reasons for moving is not very helpful. One of the answer options is "moved to a new residence." This answer is chosen by the vast majority of the 'movers'. Clearly, it does not add much to our understanding of the reasons for moving.

5 Econometric Issues

A general version of life cycle theory asserts that in each period individuals (households) choose an optimal consumption plan based on their assessment of the expected present value of the remaining life time utility. Applied to residential mobility, the theory suggests that individuals (households) compare their expected present value of the remaining life time utility if they move EU_m with the remaining life time utility if they stay EU_s and make a decision to move if

$$Ey^* = EU_m - EU_s - C_m > 0,$$

where y^* is the net benefit of moving and C_m is the utility cost of moving. The expected utility in both cases (moving and staying) will be a function of variables that describe the present period individual's (household's) wealth, characteristics of current dwelling and demographics (Feinstein and McFadden [1989]).

Formally, if the decision to move is described by a binary $[0,1]$ $(1 = move, 0 = stay)$ random variable y then

$$Ey_{it} = \Pr(y_{it} = 1) = F(Ey^*_{it}),$$

where Ey^*_{it} is the expected net benefit of moving of the i^{th} household in period t and F is a probability function. In a standard random utility approach, y^*_{it} is a continuous random variable that can be regarded as an index function

$$y^*_{it} = w_{it}\delta + \eta_i + \varepsilon_{it},$$

where w_{it} set of explanatory variables, δ is a vector of coefficients and η_i is unmeasured individual effect. Applied to residential mobility one can think about individual effects as a person's psychological cost of moving or simply "taste" for mobility[6].

[6]Feinstein and McFadden [1989] tested their data for the null of "no individual heterogeneity." The hypothesis was strongly rejected.

There are two basic frameworks for modelling unobserved heterogeneity[7]:

- fixed effects models in which η_i are assumed to be unit (individual) - specific and fixed over time,

- random effects models in which η_i are random variables such that η_i ˜$N(0, \sigma_\eta^2)$.

  In the first case we condition on η_i, in the second case we consider the distribution of η_i.

A fixed effect static linear model can easily be estimated by differencing out η_i. For example, we can transform the data into deviations from the cluster mean, which allows us to treat unobserved heterogeneity as a nuisance parameter and avoid its estimation altogether while producing an efficient and consistent estimator of β, called the within group estimator or, simply, within-estimator.

A random effects model is often referred to as a variance-component model because it implies that $\sigma_y^2 = \sigma_\eta^2 + \sigma_\varepsilon^2$. The OLS covariance estimator of δ is not BLUE in finite samples but the GLS (generalized-least-squares) estimator is. An important extension to the random effects panel data model has been introduced by Chamberlain [1984]. He discusses the case when individual effects are correlated with explanatory variables and there exist some joint distribution for $(x_{i1}, ..., x_{iT}, \eta_i)$. The necessary assumption in his analysis is that $E(u_{it}|x_{i1}, ..., x_{iT}, \eta_i) = 0$, which implies that x_{it} are exogenous. It is not restrictive to write the linear predictor of η_i as

$$E(\eta_i|x_{i1}, ..., x_{iT}) = a_0 + a_1 x_{i1} + a_2 x_{i2} + +a_T x_{iT},$$

[7] The benefits of using panel data have been widely explored in the context of linear models (exellent book-length reviews are Hsiao [1986], Baltagi [1995], Matyas and Sevestre [1996], Arellano and Honore [1999]).

where $\mathbf{a} = V^{-1}(x_i)Cov(x_i, \eta_i)$.

The linear predictor of y can now be written in reduced form as

$$E(y_{it}|x_{i1}, ..., x_{iT}) = \beta x_{it} + E(\eta_i|x_{i1}, ..., x_{iT}) = \xi_i + \pi_{i1}x_{i1} + ... + \pi_{iT}x_{iT}.$$

Given the exogeneity of x, the coefficients π form matrix Π, which has a distinctive structure:

$$\Pi = \beta\mathbf{I} + \iota a',$$

where I is a $T \times T$ identity matrix and ι is a $T \times 1$ vector of ones. An efficient estimator of β can be obtained by using minimum-distance approach (or GMM) that is by minimizing the distance between the estimated Π and the matrix of structural coefficients. It is possible that some time-invariant variables reflecting measured heterogeneity such as sex, race, union status and so on, can be absorbed in Π. A solution to this problem is discussed in Hausman and Taylor [1981].

Fewer results are available for non-linear discrete choice models although some significant progress has been made in recent years. Consider the following (static) binary choice probit model for N individuals (households) observed T consecutive time periods:

$$y_{it} = \mathbf{1}(\beta' x_{it} + \varepsilon_{it}),$$

where $\mathbf{1}(.)$ denotes an indicator function, which takes values 0 and 1 depending on whether (.) is negative or positive. If $\varepsilon_{it}\tilde{}N[0,1]$ then the panel nature of the data is irrelevant and we can consider a pooled regression. If we want to model individual heterogeneity explicitly then we should consider a different set of assumptions.

As in the linear case, suppose that ε_{it} can be decomposed as

$$\varepsilon_{it} = \eta_i + u_{it},$$

where η_i is the group specific effect independent of u_{it}, $E(u_{it}|x_{i1},...,x_{iT},\eta_i) = 0$ and (Var of u_{it} is normalized to one because in binary models the scale factor is unidentifiable). Then,

$$y_{it} = \mathbf{1}(\beta' x_{it} + \eta_i + u_{it})$$

A simple random effects probit model in which individual effects are assumed to be uncorrelated with the explanatory variables can be defined as

$$E(y_{it}) = \Phi(\alpha + \beta_x' x_{it} + \beta_d' d_{it} + \eta_i),$$

where y_{it} is a binary choice variable indicating whether the household moved during the reference year, x_{it} is a vector of exogenous explanatory variables that include household total earnings, investment income and housing price index/consumer price index ratio for the province in which the household resides, d_{it} is a vector of demographic variables, η_i is an individual specific random variable such that $\eta_i \tilde{} N(0, \sigma_\eta^2)$ and Φ is the cumulative normal distribution function. The estimation of the random effects model is based on Butler and Moffitt [1982] and requires Gauss-Hermite quadrature for computation.

As in linear models, the assumption that η_i's are uncorrelated with x_{it}'s has serious limitations. To relax this assumption we need to assume a specific functional form of η_i. One of the possibilities is to assume (following Chamberlain) that η_i is linearly dependent on $x_i = (x_{i1}'...x_{iT}')'$, that is η_i=$\lambda' x_i + \omega_i$, where ω_i $\tilde{} IN[0, \sigma_\omega^2]$. Then,

$$y_{it} = \mathbf{1}(\beta' x_{it} + \lambda' x_i + \omega_i + u_{it}),$$

where $E(\omega_i + u_{it}) = 0$. If we assume that $Var[u_{it}] = 1$, then $Var[\omega_i + u_{it}] = 1 + \sigma_\omega^2$ and the variance-covariance matrix is $\mathbf{I}_T + \sigma_\omega^2 \iota\iota'$. The inclusion of $\lambda' x_i$ produces cross-equation restrictions on the coefficients.

We can consider probit equation for each t. The distribution for y_{it} conditional on x_i but marginal on η_i has a probit form:

$$Prob(y_{it} = 1) = F[(1 + \sigma_\omega^2)^{-1/2}(\beta' x_{it} + \lambda' x_i)]$$

The matrix of the multivariate probit coefficients is:

$$\mathbf{D} = diag(1 + \sigma_\omega^2)^{-1/2}[(\beta'\mathbf{I} + \iota\lambda'].$$

This can be solved for $(1 + \sigma_\omega^2)^{-1/2}\beta$ and $(1 + \sigma_\omega^2)^{-1/2}\lambda$. We can estimate β and λ by running T probit equations separately and imposing $\mathbf{D} = diag(1 + \sigma_\omega^2)^{-1/2}[(\beta'\mathbf{I} + \iota\lambda']$ as a constraint using a minimum distance estimator (see Chamberlain [1984]).[8]

Although Chamberlain's minimum distance estimator is consistent and efficient, the retrieval of the structural form parameters is computationally involved. Bover and Arellano [1997] proposed a simple alternative to the Chamberlain's minimum distance estimator, which does not require non-linear optimization. It is discussed in the next subsection.

5.1 Bover-Arellano Estimators

Bover and Arellano [1997] proposed a simple two-step estimator for limited dependent variable models, which may include lags of the dependent variable, other endogenous explanatory variables and unobserved individual effects. The

[8]Recent developments in the field have produced altogether different estimation methods for limited dependent variable panel data models based on simulated moments (see Keane [1994]).

estimator is based on reduced form predictions of the latent endogenous variables and can be viewed as a member of Chamberlain's class of random effects minimum distance estimators. As such, it is consistent and asymptotically normal for a fixed number of periods. Although the within-group Bover-Arellano estimator is not asymptotically efficient since it implicitly uses a non-optimal weighting matrix, it is possible to obtain, in one more step, a linear GMM estimator that is asymptotically efficient.

A particularly interesting feature of the Bover-Arellano estimator is the separation of the specification searches at each level of estimation. The primary goal at the level of the reduced form is to find satisfactory estimates of the latent variables and asymptotic variance-covariance matrix at each period. Once this goal is achieved, different models can be tried and tested on the structural form level.[9]

5.1.1 Static Model

Consider, first, a static random effects limited dependent variable (LDV) model

$$y_{it}^* = x_{it}'\beta + u_{it}$$

$$u_{it} = \eta_i + \nu_{it} \ (i = 1, ..., N; \ t = 1, ..., T)$$

where x_{it} is an exogenous variable such that $E(v_{it}|x_{i1}, x_{i2}, ..., x_{iT}, \eta_i) = 0$, η_i is an unobservable individual effect potentially correlated with x_{it}, and y_{it}^* is a dependent latent variable which is not directly observable. Chamberlain [1984] suggests the following parametrization of η_i

$$E(\eta_i|x_{i1}, x_{i2}, ..., x_{iT}) = \lambda_0 + \lambda_1' x_{i1} + \lambda_2' x_{i2} + ... + \lambda_2' x_{iT} + \lambda_*' r_i$$

[9]The only application of the Bover-Arellano estimator known to me is a study of the demand for tobacco in Spain (Labeaga [1999]).

where r_i is a vector of variables that includes nonlinear (quadratic, cubic and so on) terms in the x's as well as time invariant variables. Hence, y_i^* can be rewritten in the reduced form as

$$y_i^* = \Pi z_i + \varepsilon_i$$

where $z_i = (x'_{i1}, x'_{i2}, ..., x'_{iT}, r'_i)$.

The transformation of the variables into the deviations from the mean eliminates the η_i's.

$$y_{it}^+ = X_i^+ \beta + \nu_i^+$$

where $y_{it}^+ = \mathbf{Q} y_i^*$, $X_i^+ = \mathbf{Q} X_i$, $\nu_i^+ = \mathbf{Q} \nu_i$ and $\mathbf{Q} = \mathbf{I}_T - \iota\iota'/T$. If y_i^* is directly observed, the OLS regression gives us the within-estimator of β. However, even if y_i^* is not directly observed, the following expression for the restrictions

$$X_i^+ \beta = \mathbf{Q} \Pi z_i$$

implies that

$$\beta = \left(\sum_{i=1}^{N} X_i^{+\prime} X_i^+ \right)^{-1} \sum_{i=1}^{N} X_i^{+\prime} \Pi z_i$$

To estimate β, we can replace Π with a consistent estimator $\widehat{\Pi}$ and y_i^* with a consistent reduced form predictor $\widehat{y_i}$. Bover and Arellano [1997] show that if $\widehat{\Pi}$ is a consistent and asymptotically normal estimator of Π, then $\widehat{\beta}$ is also consistent and asymptotically normal. The asymptotic variance of can be consistently estimated as

$$AVAR(\widehat{\beta}) = \left(\sum_{i=1}^{N} X_i^{+\prime} X_i^+ \right)^{-1} M' \widehat{V} M \left(\sum_{i=1}^{N} X_i^{+\prime} X_i^+ \right)^{-1},$$

where $M = \sum_i (X_i^+ \nwarrow z_i)$ and $\widehat{V}$ is a consistent estimator of V.

A consistent estimator of Π can be obtained using a simple probit specification based on the assumption that the errors ε_{it} in the reduced form are independent of z_i and $\varepsilon_{it}|z_i\tilde{}N(0,\sigma^2)$. Using $\sigma^2 = 1$ as a normalization, we have

$$Pr(y_{it} = 1|z_i) = \Phi(\pi_t' z_i)$$

where $\Phi(.)$ is the $N(0,1)$ cdf and π_t is the t-th row of Π.

This model can be extended to the case of time-series heteroskedasticity in which $\varepsilon_{it}|z_i\tilde{}N(0,\sigma_t^2)$ so that $\Pr(y_{it} = 1|z_i) = \Phi(\pi_t^{*\prime} z_i)$, where $\pi_t^* = \pi_t/\sigma_t$.

The asymptotic variance matrix $\widehat{V}$ can be calculated as follows. Let be $\widehat{\pi}$ an estimate of π such that $\widehat{\pi} = vec(\widehat{\Pi})$ minimizes $s(\pi) = \sum_{t=1}^{T} s_t$, where s_t is a differentiable criterion (for instance, a log likelihood function) and $s_t = \sum_{i=1}^{N} s_{it}(y_{it}, z_i, \pi_t)$. Under usual regularity conditions, the Taylor series expansion of $\frac{\partial s(\widehat{\pi})}{\partial \pi}$around the true value of π suggests an estimate of $\widehat{V}$ of the form

$$\widehat{V} = \frac{1}{N}\widehat{H}^{-1}\Psi\widehat{H}^{-1}$$

where $\widehat{H} = diag\left(N^{-1}\frac{\partial^2 \widehat{s}_t}{\partial \pi_t \partial \pi_t'}\right)$ and $\widehat{\Psi} = N^{-1}\sum_{i=1}^{N} \frac{\partial \widehat{s}_{it}}{\partial \pi_t}\frac{\partial \widehat{s}_{is}}{\partial \pi_s}$.

5.1.2 Within Group Dynamic Estimator

Generally, the inclusion of the lagged dependent variable into a fixed-effects model makes OLS within-group inconsistent due to the correlation of the lagged dependent variable with the residual. The problem can be avoided by including $y_{i(t-1)}^*$ as opposed to $y_{i(i-1)}$. Bover and Arellano point out that by conditioning on $y_{i(t-1)}$ one is conditioning on past choices while by conditioning on $y_{i(t-1)}^*$ one is specifying distributed lagged effects of past exogenous variables and past errors on current choices. Consider a dynamic model of the form

$$y_{it}^{*} = \alpha y_{i(t-1)}^{*} + x_{it}'\beta + u_{it} = w_{it}^{*\prime}\delta + u_{it}$$

$$u_{it} = \eta_i + \nu_{it}$$

As above we assume that $E(\eta_i|z_i) = \lambda' z_i$. We also assume that $E(y_{i1}^{*}|z_i) = \mu' z_i$ so the reduced model is the same as in the case of the static model. For the set of the last $(T-1)$ equations, it can be rewritten as

$$(I_0 - \alpha L)y_i^{*} = X_i'\beta + \eta_i \iota + \nu_i$$

where I_0 is the trim operator $I_0 = (0 \vdots I_{T-1})_{(T-1)\times T}$, L is the lag operator $L = (I_{T-1} \vdots 0)_{(T-1)\times T}$, X_i is the last $(T-1)$ observations for each unit, and ν_i are $(T-1)\times 1$ vectors. Letting $B = I_0 - aL$ and Q be of order $(T-1)$ we have

$$QBy_i^{*} = X_i^{+}\beta + \nu_i^{+}.$$

We can rewrite the restrictions in the form $X_i^{+}\beta = QB\Pi z_i$ or $X_i^{+}\beta = Q(I_0 - \alpha L)\Pi z_i$. If $W_i = (L\Pi z_i \vdots X_i)$ and $W_i^{+} = QW_i$, then

$$W_i^{+}\delta = QI_0\Pi z_i,$$

where $\delta = [\alpha \vdots \beta']'$. Hence,

$$\delta = \left(\sum_i W_i^{+\prime} W_i^{+}\right)^{-1} \sum_i W_i^{+\prime} I_0 \Pi z_i.$$

Similarly to the argument for the static model estimator, Bover and Arellano show that the estimator

$$\widehat{\delta} = \left(\sum_i \widehat{W}_i^{+\prime} \widehat{W}_i^{+}\right)^{-1} \sum_i \widehat{W}_i^{+\prime} \widehat{y}_{i0}^{+},$$

where $\widehat{y}_{i0} = I_0\widehat{\Pi} z_i$ and $W_i = (L\Pi z_i \vdots X_i)$ is consistent and asymptotically normal given the consistency and asymptotic normality of $\widehat{\Pi}$. The variance of $\widehat{\delta}$ can be consistently estimated as

$$AVAR(\widehat{\delta}) = \left(\sum_i \widehat{W}_i^{+} \widehat{W}_i^{+} \right)^{-1} \widehat{M}' \widehat{V}^* \widehat{M} \left(\sum_i \widehat{W}_i^{+} \widehat{W}_i^{+} \right)^{-1},$$

where $M = \sum_i \widehat{W}_i^{+} \nwarrow z_i$ and $\widehat{V^*} = (\widehat{B} \nwarrow I_m)\widehat{V}(\widehat{B}' \nwarrow I_m)$.

Bover and Arellano show that this estimator is inefficient relative to Chamberlain's minimum distance estimator as it is essentially the minimum distance estimator which uses a non-optimal weighting matrix. They also point out that the robustness of $\widehat{\delta}$ depends directly on the robustness of $\widehat{y}_{it}$. The good news, however, is that specification searches at the level of the reduced form are separated from the specification searches at the level of structural equation, which means that the functional form can be tested in the reduced form level until the satisfactory $\widehat{y}_{it}$'s are available.

5.1.3 GMM Estimator

Finally, Bover and Arellano suggest a linear GMM estimator that can be obtained in one more step and is asymptotically efficient relative to the minimum distance estimator. Consider again a dynamic model given by (8). Given that within-groups equation errors are uncorrelated with the conditioning variables z_i, we can write down the optimal GMM estimator of δ based on the following moment conditions

$$E\left[Z_i' Q u_i\right] = 0,$$

where $Z_i = (I \nwarrow z_i')$, Q is a $(T-1) \times (T-1)$ within-groups operator and $u_i = (u_{i2}, ..., u_{iT})$. The estimator of δ based on the sample orthogonality condition will be given by

$$b_N(\delta) = \frac{1}{N} \sum_{i=1}^{N} Z_i' \left(\widehat{y}_{i0}^{+} - \widehat{W}_i^{+} \delta \right),$$

where $\widehat{y}_{i0}=I_0\widehat{\Pi}z_i$, $\widehat{W}_i = \left(L\widehat{\Pi}z_i \vdots X_i\right)$ and symbol "+" denotes within-groups transformations. A GMM estimator of δ is

$$\left[\left(\sum_i \widehat{W}_i^{+\prime} Z_i\right) A_N \left(\sum_i Z_i^{\prime}\widehat{W}_i^{+}\right)\right]^{-1}\left(\sum_i \widehat{W}_i^{+\prime} Z_i\right) A_N \left(\sum_i Z_i^{\prime}\widehat{y}_{i0}^{+}\right)$$

where A_N is a weighting matrix.

If $A_N = \left(\sum_i Z_i^{\prime}Z_i\right)^{-1}$, then the GMM estimator coincides with the WG estimator

$$\widehat{\delta} = \left(\sum_i \widehat{W}_i^{+\prime}\widehat{W}_i^{+}\right)^{-1}\sum_i \widehat{W}_i^{+\prime}\widehat{y}_{i0}^{+}$$

Arellano and Bover [1995] showed that the $(T-1)\times(T-1)$ within-groups operator Q can be replaced by the $(T-2)\times(T-1)$ first difference operator K, so that a generic estimator takes form

$$\left[\left(\sum_i \widehat{W}_i^{\prime}K^{\prime}Z\right) A_N \left(\sum Z^{\prime}K\widehat{W}_i\right)\right]^{-1}\left(\sum_i \widehat{W}_i^{\prime}K^{\prime}Z\right) A_N \left(\sum_i Z^{\prime}K\widehat{y}_{i0}^{+}\right)$$

Using K instead of Q allows to eliminate redundant moment conditions.

The optimal choice of A_N is given by a consistent estimate of the inverse of the covariance matrix of the orthogonality conditions. Let $M_{zz} = \sum_i Z_i^{\prime}Z_i$ and $V^{*} = (B\otimes I_m)V(B^{\prime}\otimes I_m)$. The estimator of $\widetilde{\delta}$ that uses $A_N = \widehat{V}_b^{-1}$, where

$$\widehat{V}_b = (K\otimes I_m)\, M_{zz}\widehat{V^{*}}M_{zz}\,(K^{\prime}\otimes I_m)$$

is asymptotically efficient and asymptotically equivalent to the optimal minimum distance estimator. Furthermore, a consistent estimate of the asymptotic variance of is given by

$$\begin{aligned} AVAR(\widetilde{\delta}) = {} & \left[\left(\textstyle\sum_i \widehat{W}_i^{\prime}K^{\prime}Z_i\right) A_N \left(\textstyle\sum_i Z_i^{\prime}K\widehat{W}_i\right)\right]^{-1} \times \\ & \times \left(\textstyle\sum_i \widehat{W}_i^{\prime}K^{\prime}Z_i\right) A_N\widehat{V}_b A_N \left(\textstyle\sum_i Z_i^{\prime}K\widehat{W}_i\right)\left[\left(\textstyle\sum_i \widehat{W}_i^{\prime}K^{\prime}Z_i\right) A_N \left(\textstyle\sum_i Z_i^{\prime}K\widehat{W}_i\right)\right]^{-1} \end{aligned}$$

The estimate of $\widehat{V}$ depends on α through matrix B, so the calculation of the efficient estimator will require a preliminary consistent estimate of α. This estimate can be obtained by estimating δ (WG estimator) which produces inefficient but consistent estimate of α.

To capture the asymptotic behavior of the Bover-Arellano estimators, I conducted Monte-Carlo simulations for two different sample sizes N=250 and N=1000 with T=5. The number of repetitions in most of the simulations was 100, however a set of simulations was performed with 1000 repetitions. Some of the results of these simulations are presented in the next section.

6 Monte Carlo Simulations

For simplicity, I assume that there is only one independent variable. This variable is generated according to a formula similar to the Nerlove (1971) process so that the elements of the dependent variable X are correlated overtime. X is generated as

$$x_{it} = 0.1t + 0.5x_{i(t-1)} + xd_{it} + d_t + U[-0.5, 0.5],$$

where t is a time trend, xd_{it} is a dummy variable that equals 1 with probability 0.5 and d_t is the time effects assumed to be -1 in all periods (see Lechner [1995]) and U[-0.5,0.5] is a uniformly distributed random variable that takes values from -0.5 to 0.5. The value of x_{i0} is assumed to be zero. To reduce the effect of the initial condition 7 waves of x were generated, however only the last 5 were used.

The important assumption of the model is that individual effects can be approximated to any degree by a polynomial expansion. Generally, a linear specification for individual effects may include quadratic or cubic terms. For the purpose of simulations, I assumed that a combination of only linear terms provides a good approximation for the individual effects

$$\eta_i = \lambda' z_i + N[0, 0.01],$$

where $\lambda' = (1, 0.5, 0.3, 0, -0.2)$ and $z_i = (x_{i1}, x_{i2}, ...x_{i5})'$. To ensure good fit, the standard deviation of the random term was chosen to be small (0.1). Another consideration in choosing the values of the coefficients is to make the mean of the latent dependent variable approximately 0.5 for $\beta = 1$.

For the static model the dependent variable Y is generated according to

$$y_{it} = \beta x_{it} + \eta_i + u \cdot N[0, 1].$$

In the dynamic specification, the lagged dependent variable was added

$$y_{it} = \alpha y_{i(t-1)} + \beta x_{it} + \eta_i + u \cdot N[0, 1], \;\; y_{i0} = 0.$$

Simulations were performed for different β and u.

Table 10 summarizes the results for the static model where 'nrep' (the number of repetitions) = 100 (It would be extremely time consuming to perform, say, 1000 repetitions when N=1000. However, later I will present comparative results for N=250 and nrep=1000). Table 10 contains the mean of β estimates for N (the number of individuals in the sample) =1000 and N=250, the mean of estimated variance of β, the number of rejections for the 95% level of significance, the variance of β in the obtained sample of estimates and the confidence interval for the mean of β. "True" β takes values 1, -1 and 0.8. Setting u equal to 1 allows for the identification of the coefficients. A set of simulation has been performed for $u = 0.8$.

Even for N=250 (250 individuals) the estimates of β and variance of β are close to the true values although the number of rejections (especially for $\beta = 1$

	$\beta = \mathbf{1}, \mathbf{u} = \mathbf{1}$		$\beta = -\mathbf{1}, \mathbf{u} = \mathbf{1}$	
	N=1000	N=250	N=1000	N=250
Mean estimate of β	1.0078	1.0230	-1.0115	-1.0177
Mean $\widehat{var(\widehat{\beta})}$	0.0021	0.0087	0.0021	0.0088
Number of rejections (95%)	4	10	2	3
Sample variance of $\widehat{\beta}$	0.0021	0.0136	0.0016	0.0072
	$\beta = \mathbf{0.8}, \mathbf{u} = \mathbf{1}$		$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.8}$	
Mean estimate of β	0.0951	0.1015	1.2501	1.2920
(”normalized” β)	-	-	(1.0001)	(1.0336)
Mean $\widehat{var(\widehat{\beta})}$	0.0015	0.0064	0.0027	0.0115
Number of rejections (95%)	6	9	2	6
Sample variance of $\widehat{\beta}$	0.0018	0.0112	0.0022	0.0126
(”normalized” value)	-	-	(0.0014)	(0.0081)

Table 10: Monte-Carlo: static model (nrep=100)

and $\beta = 0.1$) reaches 10% for the 95% level of significance. Increasing the number of individuals in the sample to 1000 brings considerable improvement, especially in the estimates of the variance of β. The number of rejections is very close to the 5% level, as expected. When $u = 0.8$ (fourth column) the system is not directly identified and the estimates of β must be normalized. The normalization is straightforward. The obtained probit coefficients are π^*=π/u. Hence, the predicted value of the latent dependent variable based on π^* is y^*/u. Therefore, to obtain the correct estimates of β's we have to multiply β by u. In our experiment the corrected/normalized estimates of β's are β=1.2501*0.8=1.0001 (N=1000) and β=1.2920*0.8=1.0336 (N=250). “Normalized” values of estimates and variances of estimates are given in the brackets.

The next set of simulations was designed to compare properties of two estimators: 1) GMM estimator which uses optimal weighting matrix and, thus, asymptotically as efficient as minimum distance estimator, and 2) less efficient GMM estimator which is numerically equivalent to the within-groups estimator.

The results are summarized in Table 11 (N=1000). In addition to the estimates of β and variance of β the mean and variance of α estimates are also presented. The system is identified at $u \approx 0.85$. The reason for this is the additional variance component in the lagged dependent variable. Finding a precise value of u at which the system is identified is complicated by this additional time series component.

As u decreases from 1 to 0.8, the estimates of β increase from 0.85 for GMM (0.8806 for WG) to 1.056 (1.1013 for WG), reaching 1.0046 (1.037 for WG) for $u = 0.85$. When $u = 0.85$ the GMM estimator clearly outperforms the WG estimator. The estimates of α are also much closer to the true values for the GMM estimator. The estimates of variances are smaller and closer to the true values for the GMM estimator. The number of rejections, however, is higher for the GMM estimator for higher u's (u=1 and u=0.9) and lower for lower u's.

All previous experiments were conducted with 100 repetitions. Undoubtedly, it would be desirable to conduct simulations for a much larger number of repetitions. For N=1000, however, simulations are considerably more time consuming. Less time is needed for N=250. Table 12 compares the estimation results for nrep=1000 when N=250.

Comparing the results for N=1000 and N=250 it is clear that estimation results improve considerably when the sample size increases. Given the asymptotic properties of the Bover-Arellano estimator this was expected. The comparison between the WG and GMM estimators also shows that the GMM estimator produces smaller variances although the number of rejections is approximately the same.

To summarize the results, it is clear that the Bover-Arellano estimator per-

$\alpha = 0.5, N = 1000, nrep = 100$	$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.85}$		$\beta = -\mathbf{1}, \mathbf{u} = \mathbf{0.85}$	
	WG	GMM	WG	GMM
Mean estimate of β	1.037	1.0005	-1.005	0.9894
Mean estimate of α	0.4441	0.4853	0.4993	0.4971
Mean $\widehat{var(\widehat{\beta})}$	0.0070	0.0053	0.0045	0.0038
Number of rejections for β (95%)	5	3	5	4
Sample variance of $\widehat{\beta}$	0.0056	0050	0.0047	0.0039
	$\beta = \mathbf{1}, \mathbf{u} = \mathbf{1}$		$\beta = \mathbf{1}, \mathbf{u} = \mathbf{1.1}$	
Mean estimate of β	0.8876	0.9441	0.7864	0.8466
(”normalized” β)	(0.9862)	(1.0490)	(0.9616)	(1.0347)
Mean estimate of α	0.4629	0.3593	0.4580	0.3221
Mean $\widehat{var(\widehat{\beta})}$	0.0053	0.0226	0.0046	0.0194
Number of rejections (95%)	7	4	9	6
Sample variance of $\widehat{\beta}$	0.0046	0.0324	0.0052	0.0225
(”normalized” value)	(0.0057)	(0.040)	(0.0078)	(0.0336)
	$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.8}$		$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.7}$	
Mean estimate of β	1.1090	1.1967	1.2638	1.3220
(”normalized” β)	(0.9858)	(1.0637)	(0.9830)	(1.0282)
Mean estimate of α	0.4513	0.3329	0.4304	0.3397
Mean $\widehat{var(\widehat{\beta})}$	0.0077	0.0315	0.0095	0.0397
Number of rejections (95%)	4	4	9	5
Sample variance of $\widehat{\beta}$	0.0077	0.0437	0.0089	0.0445
(”normalized” value)	(0.0061)	(0.0345)	(0.0054)	(0.0269)

Table 11: Monte-Carlo: dynamic model (WG vs. GMM , nrep=100, N=1000) (nrep=100)

$\alpha = 0.5, N = 250, nrep = 1000$	$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.85}$		$\beta = -\mathbf{1}, \mathbf{u} = \mathbf{0.85}$	
	WG	GMM	WG	GMM
Mean estimate of β	1.1073	0.9873	-1.053	-0.9782
Mean estimate of α	0.3390	0.4127	0.4583	0.4581
Mean $\widehat{var(\widehat{\beta})}$	0.0287	0.0190	0.0190	0.0143
Number of rejections for β (95%)	93	103	71	74
Sample variance of $\widehat{\beta}$	0.0316	0.0263	0.0206	0.0176
	$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.8}$		$\beta = \mathbf{1}, \mathbf{u} = \mathbf{0.9}$	
Mean estimate of β	1.1684	1.029	1.0349	0.9294
("normalized" β)	(1.0998)	(0.9685)	(1.0958)	(0.9841)
Mean estimate of α	0.3359	0.4157	0.3472	0.4118
Mean $\widehat{var(\widehat{\beta})}$	0.0326	0.0204	0.0265	0.0178
Number of rejections (95%)	110	117	102	115
Sample variance of $\widehat{\beta}$	0.0394	0.0310	0.0300	0.0242
("normalized" value)	(0.0349)	(0.0275)	(0.0336)	(0.0271)

Table 12: Monte-Carlo: dynamic model (WG vs. GMM, nrep=1000, N=250)

forms very well in the context of a static model. The results are straightforward and easy to interpret. On the other hand, the estimators obtained in the context of the dynamic model are less reliable and inference is complicated by the presence of the additional variance component in the lagged dependent variable. The use of an optimal weighting matrix improves the results, although so the performance of the WG estimator can be considered satisfactory.

7 Econometric Results

7.1 Static specification

As mentioned above, SLID does not collect asset information. I used the "investment income$_t$" variable as a proxy for household's assets. Another income variable - "total earnings$_t$" - includes wages and salaries as well as government transfers (social security, old age security pension and Canadian and Quebec Pension Plan benefits) and private pension benefits. Given that the majority of the household heads in the sample are retired (about 72% in 1996) and will depend on the same government transfers and pensions in the future, this variable and the "investment income$_t$" variable should give a good idea about household's permanent income.

There is no information on house prices in the SLID. Here, I also have to rely on a proxy variable constructed as the ratio of the provincial housing price index and consumer price index ("hpi/cpi$_{t-1}$") to capture the effect of housing prices. The vector of demographic variables includes the variables that may change over time - "household size$_{t-1}$" and "marital status$_{t-1}$" as well as time invariant variables - "age in 1996" and "university degree." Also included are "dwelling tenure$_{t-1}$" and a dummy variable indicating poor health ("poor health$_t$").

	coef	std. err.	z	P>|z|
total earnings	0.003	0.022	0.15	0.878
investment income	-0.014	0.018	-0.8	0.422
hpi/cpi (t-1)	-0.563	0.392	-1.43	0.152
tenure (t-1)	-0.585	0.056	-10.52	0.000
urban/rural (t-1)	0.145	0.064	2.26	0.024
household size (t-1)	0.087	0.037	2.37	0.018
marital status	-0.208	0.064	-3.25	0.001
poor health	0.198	0.084	2.37	0.018
*age in 1996	-0.006	0.003	-1.70	0.089
*university degree	0.181	0.094	1.92	0.054
"rho" (fraction of variance due to η_i)	0.159	0.035		

Table 13: Random Effects Model (uncorrelated individual effects)

The results from a benchmark random effects model in which individual effects are uncorrelated with explanatory variables are presented in Table 13. As with any probit model, only the ratios of the coefficients and standard errors of the residuals can be identified. Urban and larger size households appear to be more likely to change residence, as do households in which the major income earner is in poor health. Higher mobility of those in poor health is consistent with previous findings for British households. Ermisch and Jenkins [1999] find, based on data from the British Household Panel Survey, that people with health problems are more likely to move. I also tried a specification that includes changes in health status. Similar to Ermisch and Jenkins [1999] I found no association between changes in health status and mobility. Owning the residence and being married is negatively related to mobility. Income variables are not individually significant and neither is provincial hpi/cpi ratio.

It is possible to test whether the panel estimator is different from an estimator that may be obtained from a pooled regression by testing the significance of the panel-level variance component. The null of no panel effect is rejected by the likelihood ratio test ($\chi^2 = 22.9$).

The next model relaxes the assumption that individual effects are uncorrelated with the explanatory variables and uses the two-step Bover-Arellano estimator for a static model with and without heteroskedasticity. The estimated matrix of the reduced form coefficients and the covariance matrix from the first stage are used to predict the values of the latent dependent variable and estimate the coefficients from a linear fixed effects model (within-groups) on the second stage. A particular challenge is to find a good fit for the reduced form probit specification on the first stage. I have tried several different specifications. Following Chamberlain [1984] I assume a linear parametrization for individual effects, which includes lags and leads of all exogenous variables as well as a vector of time-invariant dummy variables such as age in 1996, university degree, sex of head, visible minority and a dummy variable indicating whether the person ever had or raised a child. The goodness-of-fit for the reduced form probit varies considerably for each period. For example, the pseudo-R^2 for the first period probit is about twice the pseudo-R^2 for the fourth period. The inclusion of quadratic and cubic terms has not resulted in significant changes in the log-likelihood. Part of the problem may be attributed to the measurement error in the dependent variable ("number of moves"). As mentioned above, the variable indicates only whether the household moved at least once during the year but does not record the actual number of moves. The measurement error is absorbed in the error term, which inevitably reduces the "goodness-of-fit."

	coef	std. err.	z	P>|z|
total earnings /10000	0.088	0.085	1.04	0.298
investment income/10000	-0.016	0.029	-0.56	0.287
hpi/cpi (t-1)	0.273	0.999	0.27	0.787
tenure (t-1)	0.180	0.190	0.95	0.342
urban/rural (t-1)	0.573	0.274	2.09	0.037
household size (t-1)	0.007	0.084	0.08	0.936
marital status	0.048	0.245	0.20	0.841
poor health	0.128	0.118	1.08	0.280

Table 14: Bover-Arellano estimator (static model, no heteroskedasticity)

Table 14 presents the results from a static model that assumes no heteroskedasticity. Only the "urban/rural$_{t-1}$" variable is significant at the 95% level and the effect of the variable appears to be stronger. Generally, the variance estimates of the coefficients seem to be larger than in the random effects model.

The within-groups estimators of β for a static model with heteroskedasticity can be obtained based on reduced form model estimates of π_t^*, such that

$$\Pr(y_{it} = 1|z_i) = \Phi(\pi_t^{*\prime} z_i),$$

where $\pi_t^* = \pi_t/\sigma_t$ and $\sigma_1 = 1$. The results presented in Table 15 are very similar to the results from the previous model. The model also allows one to estimate the variances of residuals in each period assuming that the first year variance is normalized to 1. The estimates of standard errors are 0.951, 0.942 and 0.967 for the second, third and fourth period respectively. The joint F-test of $\sigma_2 = \sigma_3 = \sigma_4 = 1$ rejects the null.

	coef	std. err.	z	P>|z|
total earnings/10000	0.080	0.079	1.01	0.314
investment income/10000	-0.018	0.028	-0.63	0.529
hpi/cpi (t-1)	-0.361	1.075	-0.34	0.734
tenure (t-1)	0.177	0.186	0.96	0.337
urban/rural (t-1)	0.553	0.268	2.07	0.038
household size (t-1)	0.008	0.082	0.10	0.920
marital status	0.026	0.239	0.11	0.912
poor health	0.132	0.114	1.16	0.246

Table 15: Bover-Arellano estimator (static model with heteroskedasticity)

Urban households appear to be more mobile, perhaps because they have more housing choices than rural households. Their separation costs are probably also lower since rural households often have deeper roots in the community, are more attached to ancestral land, and have fewer job opportunities.

The estimated effects of some variables on residential mobility decisions are quite different for owners and non-owners. Table 16 shows the results separately for owners and non-owners. All three models produce similar results with respect to the relationship between owner's mobility and health status. Poor health appear to be an important determinant of the housing decisions of the elderly owners. The estimated effect of marital status is very similar in all three models although it is significant only in the random effects model in which individual effects are assumed to be uncorrelated with the regressors. Urban non-owners appear to be more mobile than rural non-owners although the results are not robust to different model specifications.

	owners		non-owners	
	coef.	st. err.	coef.	st. err.
Random effects model				
total earnings/10000	-0.020	0.028	0.054	0.040
investment income/10000	-0.009	0.022	-0.033	0.035
hpi/cpi (t-1)	-0.832	0.522	-0.350	0.660
urban/rural (t-1)	0.143	0.079	0.234	0.131
household size (t-1)	0.077	0.047	0.126	0.071
marital status	-0.385	0.084	0.109	0.110
poor health	0.263	0.124	0.134	0.118
*age in 1996	0.001	0.005	-0.016	0.005
*university degree	0.117	0.125	0.413	0.166
Bover-Arellano estimator (no heteroskedasticity)				
total earnings/10000	0.007	0.121	0.075	0.170
investment income/10000	-0.026	0.031	-0.180	0.113
hpi/cpi (t-1)	-0.292	1.235	2.522	1.717
urban/rural (t-1)	0.029	0.370	1.257	0.388
household size (t-1)	0.117	0.108	-0.250	0.153
marital status	-0.382	0.332	0.366	0.362
poor health	0.315	0.141	-0.003	0.178
Bover-Arellano estimator (with heteroskedasticity)				
total earnings/10000	0.009	0.113	-0.034	0.117
investment income/10000	-0.026	0.031	-0.125	0.067
hpi/cpi (t-1)	-0.155	1.298	-2.954	1.361
urban/rural (t-1)	0.021	0.367	0.682	0.319
household size (t-1)	0.115	0.109	-0.083	0.106
marital status	-0.382	0.326	0.121	0.254
poor health	0.317	0.139	-0.004	0.122

Table 16: Owners vs. Renters

I now turn to the dynamic specification.

7.2 Dynamic Specification

Table 17 presents the results of the dynamic model with non-optimal weighting matrix. The coefficient on hpi/cpi$_{t-1}$ is now positive and significant. The only other significant variable is "household size$_{t-1}$." An optimal weighting matrix in the GMM model produces more significant variables (Table 18).

	coef	std. err.	z	P>|z|
$\widehat{y}^*_{t-1}$	-0.156	0.103	-1.515	0.607
total earnings/10000	0.001	0.093	0.012	0.990
investment income/10000	0.030	0.032	0.947	0.343
hpi/cpi (t-1)	2.721	1.321	2.060	0.039
tenure (t-1)	-0.462	0.355	-1.301	0.193
urban/rural (t-1)	0.560	0.394	1.421	0.155
household size (t-1)	0.331	0.143	2.315	0.021
marital status	0.171	0.344	0.497	0.619
poor health	0.068	0.138	0.492	0.623

Table 17: Bover-Arellano estimator (dynamic model)

The lagged dependent latent variable is significant and negative. Somewhat surprisingly the "total earnings" variable is significant on the 95% level and positive, while the "investment income" variable - a proxy for household's assets - is not significant and negative. This result, however, is in line with Merrill [1984] who observed a similar relationship between housing assets and mobility. The effect of changes in the hpi/cpi$_{t-1}$ ratio appears to be much stronger than in the static models. The "urban/rural$_{t-1}$" variable is on the border of significance. The "household size$_{t-1}$" variable is significant on the 90% level and *negative*.

	coef	std. err.	z	P>|z|
$\widehat{y}^*_{t-1}$	-0.278	0.082	-3.390	0.001
total earnings/10000	0.131	0.057	2.298	0.021
investment income/10000	-0.037	0.025	-1.502	0.133
hpi/cpi (t-1)	3.856	1.197	3.221	0.001
tenure (t-1)	-0.391	0.301	-1.299	0.194
urban/rural (t-1)	0.628	0.329	1.909	0.056
household size (t-1)	-0.231	0.125	-1.848	0.064
marital status	-0.195	0.301	-0.647	0.517
poor health	0.024	0.119	0.202	0.840

Table 18: Bover-Arellano estimator (GMM model)

The significance and the sign of the lagged dependent variable suggest that if the net benefit of moving in the previous period is increasing, the likelihood of moving in period t is decreasing. An obvious explanation for this result is that if the net benefit of moving in period $t-1$ is high then the household is most likely to adjust its housing in that period. Once the adjustment took place, the net benefit of staying is higher.

	WG			
	owners		non-owners	
	coef.	st. err.	coef.	st. err.
$\widehat{y}^*_{t-1}$	-0.033	0.109	-0.340	0.100
total earnings/10000	0.043	0.132	0.364	0.207
investment income/10000	0.038	0.032	-0.196	0.100
hpi/cpi (t-1)	2.335	1.582	3.683	2.508
urban/rural (t-1)	0.360	0.501	1.052	0.409
household size (t-1)	-0.131	0.190	-0.856	0.227
marital status	-0.060	0.535	0.250	0.416
poor health	0.053	0.175	0.131	0.208

Table 19: Bover-Arellano estimator: Owners vs. Renters(WG)

	GMM			
	owners		non-owners	
	coef.	st.err.	coef.	st. err.
$\widehat{y}^*_{t-1}$	-0.242	0.086	-0.263	0.057
total earnings/10000	0.660	0.559	0.319	0.162
investment income/10000	-0.008	0.021	-0.077	0.059
hpi/cpi (t-1)	2.997	1.341	1.376	2.059
urban/rural (t-1)	0.846	0.403	0.320	0.236
household size (t-1)	-0.021	0.121	-0.798	0.185
marital status	-0.334	0.447	-0.150	0.309
poor health	0.013	0.134	0.016	0.155

Table 20: Bover-Arellano estimator: Owners vs. Renters (GMM)

Tables 19 and 20 show estimation results separately for owners and non-owners. Clearly, it appears that the mobility of owners and non-owners is affected very differently by different variables. While the effects of 'hpi/cpi$_{t-1}$' and 'urban/rural$_{t-1}$' are significant for owners, they are not significant for non-owners (GMM model). The opposite is true for 'household size$_{t-1}$'. These results have an intuitive appeal. A priori, we would expect that the effect of housing price changes is more important for those with housing wealth (owners) than for those without (non-owners). It appears that owners' mobility is encouraged by higher house prices, which contradicts the results of the static model and the finding by Ermisch and Jenkins [1999] who observed that older owners are discouraged by tighter housing markets.

The mobility costs (including separation costs) for rural owners are expected to be higher than for urban owners, but a rural renter is hardly more attached to his dwelling than an urban one. The estimates are significant for owners but not for non-owners. Non-owners, on the other hand, appear to be more sensitive to the changes in household size. The relationship between "household size$_{t-1}$" and mobility in both WG and GMM models is negative, which is in line with the previous observation that the moving costs for non-owners living alone are probably lower that for any other category. With respect to the question of whether the elderly use housing wealth for general consumption, there is little evidence to support that hypothesis. A positive association between "total earnings$_t$" and mobility is significant for non-owners but not for owners, while asset holding does not appear to have strong effect on mobility for either group. The model can be tested for overidentifying restrictions. The null of no overidentifying restrictions is not rejected at the 95% level.

8 Conclusions

As Deaton [1997] points out, "attempts to disentangle heterogeneity, on one hand, and dynamics, on the other hand, have a long and difficult history in various branches of statistics and econometrics" (p.111). In this paper, I present an attempt to deal with both issues in a study of the residential mobility of the elderly in Canada by employing the Bover-Arellano [1997] estimator for panel data in both static and dynamic frameworks. Although compositionally straightforward, the method appears to be very sensitive to finding a good fit for the latent variable (the net benefit of moving in this case). The limitations of the data provide an additional constraint. Future analysis will benefit from more detailed data on the number of moves during a particular year and information on household assets, mortgage payments and house prices.

The random effects approach used in this study provides a flexible framework that can be improved in several directions. In particular, it may be possible to relax the assumption of the linear specification of individual effects and generalize the model to a non-parametric specification of individual effects. Some steps in this direction have already been made (Newey [1994]; see Arellano [2000] for a detailed overview of recent contributions in the field).

With respect to the lifecycle aspect of housing decisions, there is some evidence that household income and wealth may play a role in housing mobility decisions although the link does not appear to be strong. I also find higher transition rates from non-ownership to ownership than vice versa. The results of the dynamic model (GMM estimator) suggest that mobility consideration are quite different for owners and non-owners. It appears, however, that mobility for both groups is mostly related to transaction costs. Those for whom transaction costs are lower are more likely to adjust their housing. Once the adjustment took place future mobility is less likely. The role of the transaction cost is consistent with "richer" versions of the lifecycle theory and can be further explored in a study of "reverse mortgages," an arrangement offered by banks and government agencies under which a bank or an agency purchases the home and pays to the family that continues to live in it.

The results must be considered with caution. First, the SLID does not provide information on the actual number of moves during a given year, only whether the person has moved *at least once*. In addition, relying on proxy variables for household assets and house prices reduces the precision of the estimates. Second, the two(three)-stage Bover-Arellano panel data estimator used in this study assumes a particular specification of the conditional distribution of individual effects. This assumption is restrictive (Chamberlain [1984]) and can be relaxed. Finally, both Monte-Carlo simulation and estimation confirm that the results are not very robust when both individual heterogeneity and dynamics are present in the model.

References

Arellano, M., 2000. "Discrete Choices with Panel Data," Mimeo, CEMFI.

Arellano, M. and Bover, O., 1995. "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68, pp. 29-51.

Arellano, M. and Honore, B., 2000. "Panel Data Models: Some Recent Developments," CEMFI, Working Paper # 0016.

Baltagi, B., 1995. *Econometric Analysis of Panel Data*, John Wiley and Sons Ltd.

Butler, J. and Moffitt, R., 1982. "A Computationally Efficient Quadrature Procedure for the One-factor Multinomial Probit Model," *Econometrica,* 50(3), pp. 225-238.

Börsch-Supan, A. and Pollakowski, H., 1990. "Estimating Housing Consumption Adjustments from Panel Data," *Journal of Urban Economics*, 27, pp. 131-150.

Bover, O. and Arellano, M., 1997. "Estimating Dynamic Limited Dependent Variable Models From Panel Data," *Investigaciones Economicas*, vol. XXI(2), 1997, 141-165.

Browning, M. and Crossley, T., 2001. "The Life Cycle Model Of Consumption and Saving," *Journal of Economic Perspectives*, 15(3), pp. 3-22.

Browning, M. and Lusardi, A., 1996. "Household savings: micro theories and macro facts," *Journal of Economic Literature*, 34(4), pp.1797-1855

Butlin, G., 1994. "SLID Household and Family Variables," SLID Research Paper Series, cat. no. 94-06.

Chamberlain, G., 1984. "Panel Data," in Z. Griliches and M. D. Intrilligator (eds.), *Handbook of Econometrics*, vol. II, Elsevier Science.

Deaton, A., 1992. *Understanding Consumption*. Oxford: Clarendon Press.

________, 1997. *The Analysis of Household Surveys*, The Johns Hopkins University Press, Baltimore, Maryland.

Dieleman, Frans M., 2001. "Panel Data in Housing Research: A Reaction to the Paper by Myers on Cohort Longitudinal Estimation of Housing Careers," *Housing Studies*, Vol. 16, No. 1, 115-121.

Duncan, G. and Hill, M., 1985. "Conceptions in Longitudinal Households: Fertile and Futile?" *Journal of Economic and Social Measurement*, 13, 361-375.

Ermisch, J. and Jenkins, S., 1999. "Retirement and Housing Adjustment in Later Life: Evidence from the British Household Panel Survey," *Labor Economics* 6, pp. 311-333

Feinstein, J., 1996. "Elderly Health, Housing and Mobility," in *Advances in the Economics of Aging*, (D. Wise, ed.). Chicago and London: University of Chicago Press, pp. 275-320.

Feinstein, J. and McFadden, D., 1989. "The Dynamics of Housing Demand by the Elderly," in *The Economics of Aging* (D. Wise, Ed.), Chicago: Univ. of Chicago Press.

Fratantoni, M., 1999. "Reverse Mortgage Choices: A Theoretical and Empirical Analysis of the Borrowing Decisions of Elderly," *Journal of Housing Research*, 10(2), pp. 189-208.

Hausman, J. and Taylor, W., 1981. "Panel Data and Unobservable Individual Effects," *Econometrica*, 49, 1377-1398.

Hayward, L., 2000. "Health and Residential Mobility in Later Life: A New Analytical Technique to Address an Old Problem," SEDAP Research Paper No. 34.

Hsiao, C., 1986. *Analysis of Panel Data*, Cambridge University Press.

Jones, L., 1996. "Housing Tenure Transitions and Dissaving by the Elderly," *Canadian Journal of Economics*, XXIX, Special Issue, April.

Labeaga, J., 1999. "A double-hurdle rational addiction model with heterogeneity: Estimating the demand for tobacco," *Journal of Econometrics*, 93, pp.49-72.

Lechner, M., 1995. "Some Specification Tests for Probit Models Estimated on Panel Data," *Journal of Business & Economic Statistics*, Vol. 13, No. 4.

Levin, L., 1998. "Are Assets Fungible? Testing Alternative Theories of Life Cycle Saving," *Journal of Economic Behaviour and Organization*, 36(1), pp.59-83.

Matyas, L. and Sevestre, P., 1996. "The econometrics of panel data: A handbook of the theory with applications," in *Advanced Studies in Theoretical and Applied Econometrics*, vol 33. Dordrecht; Boston and London: Kluwer Academic, 1996, pages xxii, 922.

Megbolugbe, I., Sa-Aadu, J., and Shilling, J., 1997. "Oh, Yes, The Elderly Will Reduce Housing Equity under the Right Circumstances," *Journal of Housing Research*, Vol. 8, Isuue 1.

Merrill, S., 1984. "Home Equity and the Elderly," in *Retirement and Economic Behavior* (H. Aaron and G. Burtless, eds.), Brookings Institution.

Myers, D. 1999. "Cohort Longitudinal Estimation of Housing Careers," *Housing Studies*, Vol. 14, No. 4, 473-490.

Nerlove, M., 1971. "Further evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections," *Econometrica*, 39, pp. 359-382.

Newey, W., 1994. "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 55, 357-362.

Palumbo, M., 1999. "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle," *Review of Economic Studies*, 66, pp. 395-421.

Skinner, J., 1996. "Is Housing Wealth a Sideshow?" in *Advances in the Economics of Aging* (D. Wise Ed.). Chicago: University of Chicago Press, pp. 241-271.

Sheiner, L. and Weil, D., 1992. "The Housing Wealth of the Aged," NBER Working Paper 4115.

Thaler, R., 1994. "Psychology and Savings Policies," *American Economic Review*, 84(2), p.184-192.

VanderHart, P., 1998. "The Housing Decisions of Older Households: A Dynamic Analysis," *Journal of Housing Economics*, 7, pp. 21-48

Venti, S. and Wise, D., 1989. "Aging, Moving, and Housing Wealth," in *The Economics of Aging* (D. Wise, Ed.), Chicago: Univ. of Chicago Press.

__________________, 1990. "But They Don't Want to Reduce Housing Equity, in *Issues in the Economics of Aging*, ed. D. Wise. Chicago: University of Chicago Press.

__________________, 2001. "Aging and Housing Equity: Another Look," NBER Working Paper 8608.

SEDAP RESEARCH PAPERS

Number	Title	Author(s)
No. 1:	Population Aging and Its Economic Costs: A Survey of the Issues and Evidence	F.T. Denton B.G. Spencer
No. 2:	How Much Help Is Exchanged in Families? Towards an Understanding of Discrepant Research Findings	C.J. Rosenthal L.O. Stone
No. 3:	Did Tax Flattening Affect RRSP Contributions?	M.R. Veall
No. 4:	Families as Care-Providers Versus Care-Managers? Gender and Type of Care in a Sample of Employed Canadians	C.J. Rosenthal A. Martin-Matthews
No. 5:	Alternatives for Raising Living Standards	W. Scarth
No. 6:	Transitions to Retirement: Determinants of Age of Social Security Take Up	E. Tompa
No. 7:	Health and Individual and Community Characteristics: A Research Protocol	F. Béland S. Birch G. Stoddart
No. 8:	Disability Related Sources of Income and Expenses: An Examination Among the Elderly in Canada	P. Raina S. Dukeshire M. Denton L.W. Chambers A. Scanlan A. Gafni S. French A. Joshi C. Rosenthal
No. 9:	The Impact of Rising 401(k) Pension Coverage on Future Pension Income	W.E. Even D.A. Macpherson
No. 10:	Income Inequality as a Canadian Cohort Ages: An Analysis of the Later Life Course	S.G. Prus
No. 11:	Are Theories of Aging Important? Models and Explanations in Gerontology at the Turn of the Century	V.L. Bengtson C.J. Rice M.L. Johnson
No. 12:	Generational Equity and the Reformulation of Retirement	M.L. Johnson
No. 13:	Long-term Care in Turmoil	M.L. Johnson L. Cullen D. Patsios

Number	Title	Author(s)
No. 14:	The Effects of Population Ageing on the Canadian Health Care System	M.W. Rosenberg
No. 15:	Projections of the Population and Labour Force to 2046: Canada	F.T. Denton C.H. Feaver B.G. Spencer
No. 16:	Projections of the Population and Labour Force to 2046: The Provinces and Territories	F.T. Denton C.H. Feaver B.G. Spencer
No. 17:	Location of Adult Children as an Attraction for Black and White Elderly Migrants in the United States	K.-L. Liaw W.H. Frey J.-P. Lin
No. 18:	The Nature of Support from Adult *Sansei* (Third Generation) Children to Older *Nisei* (Second Generation) Parents in Japanese Canadian Families	K.M. Kobayashi
No. 19:	The Effects of Drug Subsidies on Out-of-Pocket Prescription Drug Expenditures by Seniors: Regional Evidence from Canada	T.F. Crossley P. Grootendorst S. Korkmaz M.R. Veall
No. 20:	Describing Disability among High and Low Income Status Older Adults in Canada	P. Raina M. Wong L.W. Chambers M. Denton A. Gafni
No. 21:	Parental Illness and the Labour Supply of Adult Children	P.T.Léger
No. 22:	Some Demographic Consequences of Revising the Definition of #Old& to Reflect Future Changes in Life Table Probabilities	F.T. Denton B.G. Spencer
No. 23:	Geographic Dimensions of Aging: The Canadian Experience 1991-1996	E.G. Moore D. McGuinness M.A. Pacey M.W. Rosenberg
No. 24:	The Correlation Between Husband's and Wife's Education: Canada, 1971-1996	L. Magee J. Burbidge L. Robb
No. 25:	The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1988 Tax Flattening in Canada	M.-A. Sillamaa M.R. Veall

Number	Title	Author(s)
No. 26:	The Stability of Self Assessed Health Status	T.F. Crossley S. Kennedy
No. 27:	How Do Contribution Limits Affect Contributions to Tax-Preferred Savings Accounts?	K. Milligan
No. 28:	The Life Cycle Model of Consumption and Saving	M. Browning T.F. Crossley
No. 29:	Population Change and the Requirements for Physicians: The Case of Ontario	F.T. Denton A. Gafni B.G. Spencer
No. 30:	Nonparametric Identification of Latent Competing Risks and Roy Duration Models	G. Colby P. Rilstone
No. 31:	Simplified Estimation of Multivariate Duration Models with Unobserved Heterogeneity	G. Colby P. Rilstone
No. 32:	Structural Estimation of Psychiatric Hospital Stays	G. Colby P. Rilstone
No. 33:	Have 401(k)s Raised Household Saving? Evidence from the Health and Retirement Study	G.V. Engelhardt
No. 34:	Health and Residential Mobility in Later Life: A New Analytical Technique to Address an Old Problem	L.M. Hayward
No. 35:	2 ½ Proposals to Save Social Security	D. Fretz M.R. Veall
No. 36:	The Consequences of Caregiving: Does Employment Make a Difference	C.L. Kemp C.J. Rosenthal
No. 37:	Fraud in Ethnocultural Seniors' Communities	P.J.D. Donahue
No. 38:	Social-psychological and Structural Factors Influencing the Experience of Chronic Disease: A Focus on Individuals with Severe Arthritis	P.J. Ballantyne G.A. Hawker D. Radoeva
No. 39:	The Extended Self: Illness Experiences of Older Married Arthritis Sufferers	P.J. Ballantyne G.A. Hawker D. Radoeva
No. 40:	A Comparison of Alternative Methods to Model Endogeneity in Count Models. An Application to the Demand for Health Care and Health Insurance Choice	M. Schellhorn

Number	Title	Author(s)
No. 41:	Wealth Accumulation of US Households: What Do We Learn from the SIPP Data?	V. Hildebrand
No. 42:	Pension Portability and Labour Mobility in the United States. New Evidence from SIPP Data.	V. Andrietti V. Hildebrand
No. 43:	Exploring the Effects of Population Change on the Costs of Physician Services	F.T. Denton A. Gafni B.G. Spencer
No. 44:	Reflexive Planning for Later Life: A Conceptual Model and Evidence from Canada	M.A. Denton S. French A. Gafni A. Joshi C. Rosenthal S. Webb
No. 45:	Time Series Properties and Stochastic Forecasts: Some Econometrics of Mortality from the Canadian Laboratory	F.T. Denton C.H. Feaver B.G. Spencer
No. 46:	Linear Public Goods Experiments: A Meta-Analysis	J. Zelmer
No. 47:	Local Planning for an Aging Population in Ontario: Two Case Studies	L.M. Hayward
No. 48:	Management Experience and Diversity in an Ageing Organisation: A Microsimulation Analysis	T. Wannell M. Gravel
No. 49:	Resilience Indicators of Post Retirement Well-Being	E. Marziali P. Donahue
No. 50:	Continuity or Change? Older People in Three Urban Areas	J. Phillips M. Bernard C. Phillipson J. Ogg
No. 51:	Intracohort Income Status Maintenance: An Analysis of the Later Life Course	S.G. Prus
No. 52:	Tax-Preferred Savings Accounts and Marginal Tax Rates: Evidence on RRSP Participation	K. Milligan
No. 53:	Cohort Survival Analysis is Not Enough: Why Local Planners Need to Know More About the Residential Mobility of the Elderly	L.M. Hayward N.M. Lazarowich

Number	Title	Author(s)
No. 54:	Unemployment and Health: Contextual Level Influences on the Production of Health in Populations	F. Béland S. Birch G. Stoddart
No. 55:	The Timing and Duration of Women's Life Course Events: A Study of Mothers With At Least Two Children	K.M. Kobayashi A. Martin-Matthews C.J. Rosenthal S. Matthews
No. 56:	Age-Gapped and Age-Condensed Lineages: Patterns of Intergenerational Age Structure Among Canadian Families	A. Martin-Matthews K. M. Kobayashi C.L. Rosenthal S.H. Matthews
No. 57:	The Relationship between Age, Socio-Economic Status, and Health among Adult Canadians	S.G. Prus
No. 58:	Measuring Differences in the Effect of Social Resource Factors on the Health of Elderly Canadian Men and Women	S.G. Prus E. Gee
No. 59:	APOCALYPSE NO: Population Aging and the Future of Health Care Systems	R.G. Evans K.M. McGrail S.G. Morgan M.L. Barer C. Hertzman
No. 60:	The Education Premium in Canada and the United States	J.B. Burbidge L. Magee A.L. Robb
No. 61:	Student Enrolment and Faculty Recruitment in Ontario: The Double Cohort, the Baby Boom Echo, and the Aging of University Faculty	B.G. Spencer
No. 62:	The Social and Demographic Contours of Contemporary Grandparenthood: Mapping Patterns in Canada and the United States	C.L. Kemp
No. 63:	Changing Income Inequality and the Elderly in Canada 1991-1996: Provincial Metropolitan and Local Dimensions	E.G. Moore M.A. Pacey
No. 64:	Mid-life Patterns and the Residential Mobility of Older Men	L.M. Hayward
No. 65:	The Retirement Incentive Effects of Canada's Income Security Programs	M. Baker J. Gruber K. Milligan

Number	Title	Author(s)
No. 66:	The Economic Well-Being of Older Women Who Become Divorced or Separated in Mid and Later Life	S. Davies M. Denton
No. 67:	Alternative Pasts, Possible Futures: A “What If” Study of the Effects of Fertility on the Canadian Population and Labour Force	F.T. Denton C.H. Feaver B.G. Spencer
No. 68:	Baby-Boom Aging and Average Living Standards	W. Scarth M. Souare
No. 69:	The Invisible Retirement of Women	L. McDonald
No. 70:	The Impact of Reference Pricing of Cardiovascular Drugs on Health Care Costs and Health Outcomes: Evidence from British Columbia – Volume I: Summary	P.V. Grootendorst L.R. Dolovich A.M. Holbrook A.R. Levy B.J. O'Brien
No. 71:	The Impact of Reference Pricing of Cardiovascular Drugs on Health Care Costs and Health Outcomes: Evidence from British Columbia – Volume II: Technical Report	P.V. Grootendorst L.R. Dolovich A.M. Holbrook A.R. Levy B.J. O'Brien
No. 72:	The Impact of Reference Pricing of Cardiovascular Drugs on Health Care Costs and Health Outcomes: Evidence from British Columbia – Volume III: ACE and CCB Literature Review	L.R. Dolovich A.M. Holbrook M. Woodruff
No. 73:	Do Drug Plans Matter? Effects of Drug Plan Eligibility on Drug Use Among the Elderly, Social Assistance Recipients and the General Population	P. Grootendorst M. Levine
No. 74:	Living Alone and Living with Children: The Living Arrangements of Canadian and Chinese-Canadian Seniors	M.A. Pacey
No. 75:	Student Enrolment and Faculty Recruitment in Ontario: The Double Cohort, the Baby Boom Echo, and the Aging of University Faculty (Revised and updated version of No. 61)	B.G. Spencer
No. 76:	Gender Differences in the Influence of Economic, Lifestyle, and Psychosocial Factors on Later-life Health	S.G. Prus E. Gee
No. 77:	Asking Consumption Questions in General Purpose Surveys	M. Browning T.F. Crossley G. Weber

Number	Title	Author(s)
No. 78:	A Longitudinal Study of the Residential Mobility of the Elderly in Canada	Y. Ostrovsky