EMPIRICAL FINDINGS ON THE DISTRIBUTION AND CONSTRUCTION OF LONGITUDINAL FAMILIES, PART I: MODELING THE UNIVERSE, INSCOPE CHANGES, AND STATIC AND DYNAMIC FAMILIES*

Marvin Dicker and Robert J. Casady, National Center for Health Statistics

INTRODUCTION

Two of the features of the National Medical Care Utilization and Expenditure Survey (NMCUES) which are of interest to survey methodologists are that it is both a family survey and a longitudinal survey. These design features have necessitated innovative methodologies of sample construction and unit measurement (Dicker, 1980; 1983; Dicker and Casady, 1982). Of major interest among these innovations is the reciprocal methodology and majority rule approach developed by Dicker and Casady (1982) for constructing a sample of longitudinal families out of an original cross-sectional sample of families and persons (Kasprzyk and Kalton, 1983; McMillan and Herriot, 1983). This methodology had two goals: the first was to construct a collection of families which would be a representative sample of the number and types of families that should be found within a longitudinal family universe at any given point in time. The second was to measure change occurring within the families during their time in the universe. This paper will present data on the sample of longitudinal families constructed using the above approach.

The universe for the sample was the population of civilian, noninstitutionalized families in the United States during the year 1980. The sample, when properly weighted, should give reliable and valid estimates of this population.

PART I: MODELING THE LONGITUDINAL UNIVERSE

A Sampling Approach to Modeling the Universe. In modeling a universe, it is necessary to define when an element is in or out of the universe, and, in the case of a longitudinal universe, how long it remained in the universe. Using this axiom as a guide, this section will develop a model of the types of families (elements) one should expect in a longitudinal universe based on the time the family entered and left the universe (i.e., the time the family was eligible for the survey). This model is found in Figure 1-1.

This Figure reveals four general types of families. These are indicated by solid black lines and are labeled A, B, C, and D.

Line A represents all longitudinal families that were continuously in existence in the universe the entire time covered by the survey. As this time period was a calendar year, these families will be labeled "initially sampled families, existing all year." They were present both at the beginning and end of the survey period.

Line B represents all longitudinal families that were present in the universe at the beginning of the survey but died as a family (either disintegrated as a continuing family or left the universe) before the end of the survey period. These families will be labeled "initially sampled families, existing a part of the year." They were present at the beginning of the survey period but not at the end.

Line C represents longitudinal families that were not present at the beginning of the survey year, but who came into existence as families sometime during the survey period and then continued as a family until the end of the period. These families will be labeled "new families, existing to the end of the year."

Line D also represents longitudinal families that were not present at the beginning of the survey, and who came into existince sometime during the year. However, they were not present at the end of the survey period. These families will be labeled "new families, not existing to the end of the year."

The above typology does not take into account the exact day (or month or quarter) the family came into existence nor the amount of time it existed in the universe. The addition of these additional characteristics would, of course, refine the typology. However, it is better to avoid such complexity at this time.

The four family types discussed above can be thought of as forming an index of change in the longitudinal universe. It measures change relative to the time families entered or left the universe. In the next section, we shall observe the frequencies in the sample with which the changes indicated by the typology took place.

PART II: CHANGE IN THE LONGITUDINAL UNIVERSE

An Index of Longitudinal Family Types. The distribution of longitudinal family types will be measured by the index of longitudinal family types developed in Part I above. This index has the following categories:

The Index of Longitudinal Family Types

1. Initially sampled families
   a. Existing all year
   b. Existing a part of the year
2. New families
   a. Existing to the end of the year
   b. Not existing to the end of the year.

Note that this index only categorizes families on the basis of when they began and when they ended relative to the time period of the longitudinal universe. The index assumes that we have correctly identified the beginnings and endings of families. In a subsequent report we will discuss the validity of this assumption and its implications for the data to be presented below. But for now, let us assume that the decisions have been correct.

Gross Change in 1980 in the Number of Longitudinal Families in the Sample. Gross change is the number of families that were born

in 1980 (came into existence after the beginning of the survey) plus the number of families that died in 1980 (went out of existence before the end of the survey). This is the process that is mapped by the index of longitudinal family types derived from the model of the longitudinal universe given in Figure 1-1. This accounting of gross change is found in Table 2-1.

Table 2-1 indicates that, over the year, there were 6,798 longitudinal families in the sample. Of these, 813 families were either births (new families) or deaths (families that ceased to exist) or both. These dynamic families accounted for 12% of the total number of families in the sample.

All in all, 6, 257 families, or 92% of the completed sample were initially sampled families. Of these, 5,985 families, or 88% of the completed sample, existed all year. Another 272 families, or 4% of the completed sample, were initially sampled but died (went out of existence) sometime during the year.

As the year passed, new families entered the sample. Over the year, there were 541 new families, equalling 8% of the completed sample. These new families, however, did not represent all the new families in the universe. Because of the nature of the sample, only new families derived from an initially sampled family could be in the sample. All other new families in the universe (single households, immigrant families, etc.) did not have a chance to enter the sample. Of the new families in the sample, 465 new families, or 7% of the completed sample, were still in existence at the end of the year. Another 76 new families, or 1% of the completed sample, died (went out of existence) before the end of the year.

To get the total number of families that went out of existence during the year, we must add the total number of families that started the year but did not finish to the total number of new families that also did not finish. Over the year, there were 348 families, or 5% of the completed sample, that died (went out of existence) before the end of the year. Of these, 272 families, 4% of the completed sample, were initially sampled families, and 76 families, 1% of the completed sample, were new families.

While only 813 families accounted for the dynamic aspects of the sample in terms of the number of families in existence at any given point in time during the year, these 813 families also accounted for a combined total of 889 sampling births and sampling deaths. This was because 76 of the families experienced both a sampling birth and a sampling death during the year. These families are counted twice when totalling change events but only once when totalling dynamic families.

Although from a sampling perspective, Table 2-1 gives an accounting of the types of sampling elements (families) that the index of longitudinal family types predicted would be found in the survey, the Table does not exhaust the dyanmic nature of the survey. All of the families may have had other changes that do not affect their status as a sampling element. That is, they may have maintained their identity as a longitudinal (continuing) family and still had changes in family membership. It turns out, however, that some of these changes in family memberships may also affect the definition of the families as sampling elements. To some extent, these definitions are, and must be, arbitrary. Therefore, the gross changes in this universe presented above are not to be considered an absolute picture of family distribution in 1980, but rather a distribution relative to a particular model of sampling elements.

Net Change in 1980 in the Number of Longitudinal Families in the Sample. Net change refers to the total increase or decrease in the number of families in the sample between the beginning and ending of the survey. Table 2-2 presents this data. This Table indicates that the longitudinal sample contained 6,257 responding families at the beginning of the survey, and 6,450 responding families at the end of the survey. This amounted to a net gain in the number of longitudinal families in the sample of 193 families. This is a 3.1% increase from the beginning to the end of the year. This increase, however, is the result solely of membership changes in the original sample of families. Therefore, this figure underrepresents the true amount of net increase. This was because the sample, as previously alluded to, was selected at one point in time (t_o, the beginning of the year), and there was no mechanism in the NMCUES for picking up families that entered the universe from out-of-scope after this initial sampling, unless they merged with an already existing family.

PART III: DISTRIBUTIONS FROM THE MODEL

In this section, we will examine how the dynamic families were constructed, and how this construction affected the distribution of families found in the survey sample.

Static and Dynamic Families. Table 3-1 gives the distribution of families according to whether they were a static family (defined as a family that existed all year without changes in family membership) or a dynamic family (defined either as a family that did not exist all year or as a family that had changes in family membership). A comparison is also made in this table with families found in the National Medical Care Expenditure Survey (NMCES), conducted by NCHSR in 1977.

Table 3-1 indicates, first, that the distribution of static and dynamic families found in the NMCUES is very similar to the distribution found in the NMCES. Although the methods used in the two surveys for constructing longitudinal families differed in important ways (see Dicker, 1981), both surveys show approximately three quarters of the families to be static and one quarter to be dynamic. To be precise, in the NMCUES, 76.5 percent of the families were static compared to 78.8 percent for the NMCES. This was a difference of only 2.3 percent.

As a comparison cannot be made for the distribution of different types of dynamic families between the NMCUES and the NMCES, the

remainder of this section will focus only on the distributions of types of dynamic families in the NMCUES. (Footnote B in Table 3-1 addresses this issue.) Of the 1,599 dynamic families in Table 3-1, 813, or 12 percent of the total sample, did not exist the entire survey year. These are the same 813 families not existing the entire year found in Table 2-1. However, Table 3-1 also indicates that 786 initially sampled families (11.5 percent of the total sample) existed the entire survey year as the same family while experiencing changes in family membership. These families are also found in Table 2-1, but they are included in the 5,985 families listed as "initially sampled families existing all year." This was one of the categories of the Index of Longitudinal Family Types. This index was derived from the sampling approach to modeling the universe discussed in the text in Part I and illustrated in Figure 1. Table 3-1 indicates, as previously suggested, that the model represented by the Index fails to take into account all the dynamic aspects of families existing in the universe. This model only deals with changes in the number and type of families within the universe of families. In Table 3-1, we are also dealing with families that have membership change within families. Therefore, to get the total count of dynamic families in the universe, it was necessary to add the number of families that have membership change within families to the number of families that represent change in the universe of families. This gives a final count of 1,599 dynamic families and 5,199 static families.

As static families had the same family membership for the entire survey year, the longitudinal family construction method chosen would not have affected either their identities as the same or different over time, nor the family level values produced from their membership. For static families, the cross sectional family and the longitudinal family are the same social unit, differing only in its location in time. All that is needed for longitudinal research is repeated measurement of this unit. However, the 24 percent of the total sample represented by the 1,599 dynamic families involve a different set of considerations. For these families, membership exchanges between families raise the question of family identities over time. The remainder of this paper will deal only with these dynamic families.

Inscope and Out-of-Scope Changes. Membership exchanges between families result from internal changes in family membership. From a sampling perspective, there are two broad types of internal membership change that could affect the composition of the dynamic families in the sample. These are inscope membership change and out-of-scope membership change. Inscope membership change refers to changes in family membership that occur when either a family member moves from a nonsampled family in the universe to a sampled family, or when a family member moves from a sampled family to a nonsampled family also in the universe. These types of changes raise questions concerning both family identification procedures and family weighting procedures. Some examples of inscope changes are a marriage, or a relative (sibling, elderly parent, child, etc.) joining the family, or a divorce (separation, etc.), or, finally, a relative leaving the family. A relative that leaves the family can either join another ongoing family or set up a separate household.

Out-of-scope membership changes refer to changes in family membership that occur when a family member moves into or out of the universe as well as into our out of a family. In the NMCUES, out-of-scope changes could involve a family member entering the family from an institution, from the military, from overseas, or as a newborn child. He or she could also leave the family by going into an institution, into the military, to an overseas location, or leave as the result of death.

Table 3-2 indicates that of the 1,599 dynamic families in the sample, 1,197 families (17.6 percent of the total sample), were associated with at least one inscope membership change. Another 402 families (5.9 percent of the total sample), only had out-of-scope membership changes. As the families with out-of-scope membership changes were considered not to have longitudinal construction problems (out-of-scope membership changes were considered not to affect either the identity of the family or the weighting of the sample), the NMCUES model for constructing longitudinal families was not applied to them. Therefore, the remainder of this section will only deal with families with inscope membership changes.

The Gross Index of Inscope Change. Every incidence of inscope membership change involves at least two simultaneous change events in the universe of families. These simultaneous events are a deletion (or split) of a member (or members) from one family in the universe and an addition (or merger) of a member (or members) into one or more other families in the universe. One approach to measuring the incidence of such events is to conceptualize the longitudinal universe along a time line going from the beginning of the survey period to the end of the period. Simultaneous events of inscope family membership change occur at single time points along this line. These events always involve the deletion of one or more family members from one family and the simultaneous addition of one or more family members into one or more other families. Although three or more families can be involved, most of these situations involve only two families. Although the deletions and additions always occur at the same point in time, for accounting purposes they may be coded as happening at two adjacent points in time. For example, family A may have a split on day 95. Family B, which began as a result of this split, is coded as beginning on day 96. This coding indicates two events occurring at two points in time to two different families; however, it is really one event occurring simultaneously.

When we approach the phenomenon of inscope membership change from this perspective, the families in the sample can be divided into two broad categories. The first category includes all the sampled families whose members were only involved in one inscope, simultaneous change

event over the year. The second category includes all the sampled families whose members were involved in multiple and sequential, inscope, simultaneous change events over the year. This bivariate measure will be called the gross index of inscope family change.

Table 3-3 presents the distribution of families with inscope change by the gross index of inscope family change and a collapsed version of the index of longitudinal family types found in Table 2-1. Table 34 indicates, first, that the families with inscope changes are almost evenly divided between initially sampled families and new families. (This is not surprising as each simultaneous inscope change event must involve two or more families, and, by definition, most of the "second plus" families will be new.) Of the 1,197 families with inscope changes, 656 families (9.7 percent of the total sample) were initially sampled families. The remaining 541 families (8 percent of the total sample) were new families. As these 541 families represented all the new families generated from the initial sample, a comparison of Table 3-3 and Table 2-2 indicates that only 10.5 percent of the 6,257 initially sampled families accounted for all the new families generated in the sample.

Table 3-3 indicates, second, that most of the families with inscope change were only involved with one change event (66 percent to 34 percent). But, more important, this was also true of the initially sampled families with inscope change. Of the 656 initially sampled families that experienced inscope change, 467 families (6.9 percent of the total sample) had family members who were involved in only one inscope change event over the year, compared to 189 families (2.8 percent of the total sample) who had family members who were involved in multiple, sequential inscope change events over the year.

Finally, Table 3-3 indicates that initially sampled families associated with multiple, sequential change events generated new families at a greater rate per initially sampled family than initially sampled families associated with only one inscope change event. The rate for the former was 1.15 new families per each initially sampled family compared to .69 for the latter.

Summary and Conclusion. This paper has presented some findings from the NMCUES on family construction using reciprocal methodology and a majority population counting rule. It presented one possible approach to modeling the universe and the distribution of NMCUES families according to that model. It demonstrated, however, that this initial model was not sufficient to account for change within families as well as change in the universe of families. As part of this demonstration, the sample was dichotomized into static and dynamic families. The dynamic families were further dichotomized into families with inscope and out-of-scope changes. Finally, the families with inscope changes were dichotomized into those with only one simultaneous inscope change and into those with multiple, sequential, simultaneous inscope change. Future reports will deal with specific types of inscope changes (equal splits, unequal splits, etc.) and the sociodemographic characteristics of the individuals and families involved.

*This paper was written by the above authors in their private capacity. No official support or endorsement by the National Center for Health Statistics is intended or should be inferred.

REFERENCES

Dicker, M. (1980): "A Statistical Method for Classifying Families in Longitudinal, Large Scale, Panel Surveys." A Working Paper of the National Medical Care Utilization and Expenditure Survey. National Center for Health Statistics, U.S. Department of Health and Human Services, Washington, DC.

Dicker, M. (1981): "An Analysis of Differences in the Longitudinal Family Models of the National Medical Care Expenditure Survey and the National Medical Care Utilization and Expenditure Survey." A Working Paper of the National Medical Care Utilization and Expediture Survey. National Center for Health Statistics, U.S. Department of Health and Human Services, Washington, DC.

Dicker, M., and Casady, R.J. (1982): "A Reciprocal Rule Model for Defining Longitudinal Families for the Analysis of Panel Survey Data." 1982 Proceedings of the American Statistical Association, Social Statistics Volume. American Statistical Association, Washington, DC, Pp. 532-537.

Dicker, M. (1983): "Panel Surveys and Family Level Measures: Problems and Solutions." 1983 proceedings of the American Statistical Association. Social Statistics Volume. American Statistical Association, Washington, DC, Pp. 390-394.

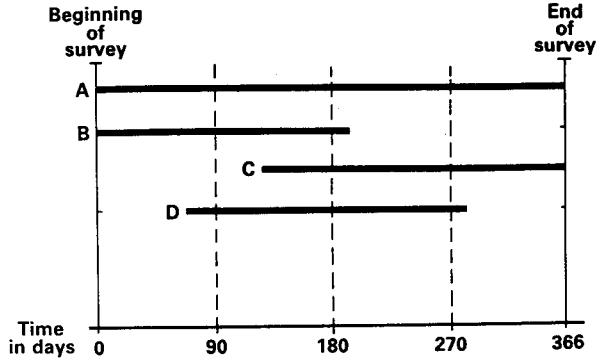
Cohen, S.B. (1982): "Family Unit Analysis in the National Medical Care Expenditure Survey." 1982 Proceedings of the American Statistical Association. Survey Research Methods Volume. American Statistical Association, Washington, DC, Pp. 561-566.

Kasprzyk, D. and Kalton, G. (1983): "Longitudinal Weighting in the Income Survey Development Program." Pp. 155-171 in Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP), Martin H. David (Editor). Social Science Research Council, New York, New York.

McMillan, D.B., and Herriot, R.A. (1983): "Towards a Longitudinal Definition of Households." A paper presented before the Washington Statistical Society, October 1983.

Figure 1

Schematic illustration of types of longitudinal families expected in a longitudinal universe when using a criterion based on time in the universe



A Initially sampled families existing all year.
B Initially sampled families existing a part of the year.
C New families existing to the end of the year.
D New families not existing to the end of the year.

Table 2-1: Gross Changes Over the Year 1980 in the Number of Longitudinal Families in the Sample

	Number of Families		%
Initially Sampled Families			
Existing all year	5,985		88
Existing a part of the year	272		4
New Families		813	
Existing to the end of the year	465		7
Not existing to the end of the year	76		1
Total	6,798		100

Source: Family History File, National Medical Care Utilization and Expenditure Survey, NCHS, 1980

Table 2-2: Net Changes Over the Year 1980 in the Number of Longitudinal Families in the Sample

Time Period	Number of Families
Beginning of Survey	6,257
End of Survey	6,450
Difference between beginning and end of survey	+ 193
Percent increase or decrease	+ 3.1%

Source: Family History File, National Medical Care Utilization and Expenditure Survey, 1980

Table 3-1. The Distribution of Static and Dynamic Families in the NMCUES Sample with a Comparative Distribution from NMCES

	NMCUES[a] (1980)		NMCES[b] (1977)	
	N	%	N	%
Static Families (The initially sampled family existed the whole year without membership change)	5,199	76.5	11,653	78.8
Dynamic Families				
Existing All Year (The initially sampled family existed the whole year, but experienced membership change)	786	11.5	N.A.[b]	N.A.[b]
Not Existing All Year (The initially sampled family did not exist the whole year and/or a new family was generated from it)[c]	813	12.0	N.A.[b]	N.A.[b]
All Families	6,798	100	14,789	100

[a]Source: Family History File, National Medical Care Utilization and Expenditure Survey, NCHS, 1980

[b]Extrapolated from S. Cohen (1982), "Family Unit Analysis in the National Medical Care Expenditure Survey." Table I. As Cohen's analysis is done in terms of RU's rather than families, it is not clear how the author determined that an RU was the same or different at different points in time when there were membership changes in the RU. However, those RU's that did not have membership changes over 5 rounds are equivalent to initially sampled families without membership changes over the whole year.

[c]See Table 2-1

Table 3-2. The Distribution of Dynamic Families in the NMCUES Sample by Whether They Were Categorized by Inscope or Out-of-Scope Membership Change

	(N)	% This Table	% of Total Sample
Families with Inscope Membership Change	1,197	74.9	17.6
Families with Out-of-Scope Membership Change Only	402	25	5.9
All Families with Membership Change	1,599	100	23.5

Source: Family History File, National Medical Care Utilization and Expenditure Survey, NCHS, 1980

Table 3-3. The Distribution of Dynamic Families with Inscope Change by Gross Index of Inscope Change and a Collapsed Version of the Index of Longitudinal Family Types

The Gross Index of Inscope Change	The Collapsed Index of Longitudinal Family Types[a]								
	All Families			Initially Sampled Families			New Families		
	(N)	% this table	% of total sample	(N)	% this table	% of total sample	(N)	% this table	% of total sample
Families associated with only one inscope change event	791	66.1	11.6	467	71.2	6.9	324	59.9	4.8
Families associated with sequential inscope change events	406	33.9	6.0	189	28.8	2.8	217	40.1	3.2
All families	1,197	100.0	17.6	656	100.0	9.7	541	100.0	8.0

Source: Family History File: National Medical Care Utilization and Expenditure Survey, NCHS, 1980.

[a]See Table 2-1 for full index.