MEASURING LABOUR MARKET DYNAMICS: AN APPLICATION TO SPAIN

di

Concha Artola e Una-Louise Bell

Labour market research has, since the late 1980's, increasingly focused on the dynamic, as opposed to stocks, characteristics of the labour market. This more recent literature can be seen to have developed along two distinct paths: i) an analysis of gross labour market flows over the business cycle (see for example, Blanchard and Diamond (1990, 1992), Davis and Haltiwanger (1990), Burda and Wyplosz (1990, 1993), Konnings (1993), Antolìn (1994, 1995, 1996), Gomez and Dolado (1995)); and ii) microeconometric analysis of the determinants of individual labour market transitions (see Katz-Meyer (1990), Clark-Summers (1979), Narendranathan and Stewart. (1993), Peracchi and Welch (1994), Aralampulan and Stewart (1995), Alba (1996a, 1996b, 1997a, 1997b), Bover et al. (1996) and Addison and Portugal (1998)). Such studies have repeatedly shown that, in contrast to the traditional perception of labour market being characterised by large sluggish stocks, a considerable number of OECD labour markets are in fact characterised by a significant churning of both workers and jobs across the Business Cycle.

Many of the empirical studies focusing on the dynamics of the labour market have been conducted on national Labour Force Survey (LFS) data, using either: the set of *retrospective questions*; or exploiting the sample design of the survey which allows for the construction of panels of short duration, a strategy which is referred here to as the *matched files* approach. Both approaches suffer however, from a number of shortcomings; recall error and heterogeneous survey design being the main problems associated with the retrospective approach; misclassification, rotation group bias and non-random attrition being those associated with the rotating panel approach.

Whilst the use of LFS data to analyse labour market dynamics has generated an extensive volume of literature both in the US and Canada (see for example, Abowd and Zellner (1985), Fuller and Chua (1985), Poterba and Summers (1986, 1995), Meyer (1986), and Chua and Fuller (1987))[1], little attention has to our knowledge been paid to the reliability/appropriateness of the LFS data when used to identify transitions between labour market states in Europe. Yet, the consequences of the use of unadjusted data in an analysis of labour market dynamics can be of paramount importance: in an analysis of labour market turnover using individual records matched across consecutive waves of the LFS, the failure to correct for erroneous data caused for example, by misclassification, will typically result in a considerable over-estimation in labour force dynamics, since an individual who is correctly classified in one survey and incorrectly classified in the next, will be recorded as having made a transition, even though his labour market activity has not changed over the two periods. Thus if the objective of the study is an analysis of labour market dynamics,

[1]This work being a by-product of the general debate which took place towards the end of the 1970's as to the extent of labour market turnover. To summarise, a number of academics such as Feldstein (1973) and Hall (1972) argued that the US labour market was actually characterised by a considerable degree of mobility, with the majority of unemployment being largely frictional in nature and individuals experiencing very short unemployment spells. Others such as Akerlof and Main (1980), Clark and Summers (1979) and Poterba and Summers (1986) however, fervently rejected this idea, arguing instead that unemployment durations were actually much longer than that implied by simply looking at the raw labour force data. The reason being, that the use of national labour force surveys for the measurement of labour market turnover was in itself problematic. In that the use of these surveys, constructed to provide cross-sectional images of the economy at specific points in time, to create quasi-longitudinal data banks tends to result in a considerable amount of noise, due for example misclassification errors and mismatching of individual records over time, entering into the longitudinal data. Thus resulting in a distorted picture of the degree of dynamism of the market.

and a considerable number of individuals are actually misclassified, then research based on unadjusted flow data is likely to be subject to considerable error. In the case of duration analysis, the ramifications of such issues are equally worrisome. In their analysis of the effects of unemployment benefits on unemployment transitions for example, Poterba-Summers (1995) conclude that "the sensitivity of spell duration to UI benefits is substantially greater when the transition probabilities are corrected for classification error, in part because this correction reduces the level of the estimated unemployment escape rate".

This paper should be viewed as a first attempt to assess the magnitude and bias of these errors in labour markets which differ considerably from that of the US, where the majority of existing research in this area has to date been undertaken. The objective being to provide an overview of the potential pitfalls to avoid when trying to measure labour market dynamics, and to offer some evidence as to the size of the bias likely to be introduced into studies which fail to account for the error structure of the underlying data. In section one we evaluate the appropriateness of the methodologies frequently adopted in the literature to measure labour market turnover, highlighting in the course of this evaluation a number of difficulties which arise due to both the specific nature of the underlying LFS data and the respective characteristics of these alternative methodologies. In section two, an attempt to gauge the magnitude of the errors associated with the individual shortcomings of these alternative approaches is made using the Spanish labour force survey, in conjunction with previously under-utilised information obtained from Spanish Reinterview surveys. Finally in section three, we demonstrate that the available techniques used to adjust the observed labour market flows for misclassification errors in the underlying data cannot be applied to the Spanish data, since they result in negative gross flows across a number of labour market states. The results of this analysis draw into question the validity of the underlying assumptions of the methods currently available to correct the observed gross flows and underline the fact that access to individual re-interview data is needed in order to correctly adjust observed flows for classification errors.

1. Measuring Labour Market Turnover: the retrospective questions versus the matched file approach.

In this section we assess the appropriateness of the two alternative methodologies frequently adopted in the literature -the retrospective information and the matched file approaches- to obtain measures of labour market dynamics. Such an evaluation by its very nature, will at the same time allow us to assess the quality of the underlying data used in such studies.

1.1 The Retrospective Information Approach

The national labour force surveys typically contain a significant amount of retrospective information regarding an individual's labour market activity in the past. Some of the retrospective questions (such as unemployment duration) are included in every issue of the survey, whilst others (amongst which those relating to past labour market status or current job tenure) are only included in special supplements undertaken periodically, typically on an annual basis[2]. The quality of this retrospective information (and thus the subsequent ramifications for its use in empirical studies) which has been frequently used to obtain a measure of the flows across labour market, has been the focus of extensive literature, particularly in the US, where attention has tended to focus, given the differences that have emerged in unemployment and out of the labour force rates calcu-

[2]For example, the US Labour Force Survey includes in January of selected years, a supplementary section on occupational mobility and job tenure, which among others, contains a set of questions, usually addressed only to those currently employed, on their labour market status one year prior to the current survey.

lated from both contemporaneous and retrospective data, on the problems of recall error and heterogeneous sample design.

Recall Error: In view of its retrospective nature, the information used to construct labour market flows, and thus the retrospective question approach in itself, is particularly susceptible to the problem of *recall error*[3]. For, whilst it may be reasonable to expect a respondent to provide an accurate answer to his *current* labour market status at the moment in time in which the survey is being carried out, it is more questionable whether he will be able to accurately recall his labour market status of the previous year[4]. Studies by Horvarth (1982) and Mathiowetz and Duncan (1988) for example, have in fact not only highlighted the existence of a considerable degree of recall error in the recollection of unemployment spells, but also shown that errors of this nature tend to be an increasing function of the length of the recall period. The probability of inaccurate responses being recorded is likely to be accentuated further by the fact that the *LFS* questionnaires are often actually answered by any adult (individual over the age of 16) member of the household, who then answers on behalf of all the members of the family. A situation in which an elderly grandmother or young son (individuals whose reliability as conveyers of accurate information is highly questionable) responds for the entire household cannot therefore, be ruled out. When it comes to having to recall the situation of other household members, it is also highly likely that the respondent will only be able to recall "salient"

[3]See Akerlof & Yellen (1986) and Mathiowetz and Duncan (1998) for a more in-depth overview of these issues.

[4]It is now widely accepted amongst for example, psychologists that, with the exception of 'salient' events, an individual's memory decays over time with the reported rate of the occurrence of an event being a decreasing function of the predetermined recall period. A salient event being one which is deemed to be of relatively greater importance to the individual. In terms of labour market activity, examples of such events are likely to be redundancy from a stable job or long periods of unemployment.

events, for example, the loss of employment of a father, or main income earner who is not normally subject to periods of unemployment.

Accurate or unbiased recollection of the duration of an event does not simply require however, that an event occurs, but that an individual is able to correctly recall the time duration of the event. Comparing individual responses over three consecutive quarters of the matched sample of the *CPS*, the US labour force survey, Poterba and Summers (1983) illustrate that unemployment duration data is in fact a particularly unreliable variable, with 3 out of every 4 survey respondents reporting inconsistent unemployment duration over consecutive surveys. Furthermore, inaccuracies in reported durations tend, as one might expect, to be an increasing function of unemployment duration itself[5]. The problem of recall error bias is therefore, likely to be even more significant in the case of responses relating to spell duration (unemployment or employment), tending to result in increases in the number of transitions being recorded, and simultaneous reductions in the length of continuous states.

Heterogeneous Survey Design: A further problem researchers often have to overcome, when using retrospective questions both to create retrospective measures of the labour market stock and to identify labour market transitions, is the inconsistency of both the survey design and the classification procedures adopted in the contemporaneous and retrospectives segments of the survey. More specifically, the restrictive nature of the retrospective questions often inhibits a consistent/homogenous identification of an individual's labour market status over the two time horizons under consideration.

It is not difficult to envisage how these differences in both the information and procedure utilised to identify current and previous labour market status are likely to result in a significant number of misclassifications of labour market status. For, whilst employment is

[5]Individuals with longer (greater than 20 weeks) unemployment durations tending in the Poterba and Summers study to report relatively higher levels of inaccuracies than those of shorter duration.

likely to be an easily identified state across the two surveys[6], the heterogeneous nature of current and retrospective assignment of unemployment and not in the labour force is likely to have more serious ramifications for those individuals with weaker levels of labour force attachment, such as women or youths. It is not unlikely for example, that such individuals will find themselves being classified as unemployed according to the contemporaneous classification process, and not in the labour force according to the retrospective one. One reason being that there is likely to be relatively more pressure, due to for example, issues of unemployment eligibility, to answer positively to job search requirement activity questions in the contemporary survey than in the retrospective one (see Ureta 1987). In contrast, the fact that the retrospective surveys often does not involve the same degree of detailed questioning with respect to job search activity, implies that a much higher level of job search is required before any form of retrospective unemployment is reported. In a similar vein, one can also argue that the retrospective definition of unemployment is itself much more lenient with respect to discouraged workers, with such individuals tending not to be classified as unemployed according to the contemporaneous definition, but may well be retrospectively. Work by Stevens et al. (1987) for example, illustrates that a considerable proportion of discouraged workers do in fact exhibit high levels of labour force attachment over the long run. When such individuals start looking for employment, they may tend therefore, to report job search activity over a much longer period retrospectively, than the actual weeks in which this activity took place.

In an attempt to assess the ramifications of such inconsistencies in survey procedure and heterogeneity of sample design Levine (1990) compares the percentage of individuals classified as unemployed in the first year, (*t*), with those who fail to retrospectively report in the second year (*t+1*), an unemployment spell in *t*, using

[6]Although even this classification can still be subject to recall error bias.

matched *CPS* individual records across consecutive March surveys. His results indicate that during the period 1978-1987, 35-50% of those individuals contemporaneously classified as being unemployed in t fail to report retrospectively periods of unemployment in $t+1$. Failure to report retrospective unemployment does however, differ considerably across demographic groups: with 58% of females and youths failing to report retrospective unemployment, compared to 32% of prime age males. Whilst some of this difference is undoubtedly due to the recall error bias, Levine finds that a large proportion of the documented difference between contemporaneous and retrospectives measures of unemployment can actually be attributed to the misclassifications arising from the heterogeneous nature of the design and the classification procedures used in the contemporaneous and retrospective definitions of unemployment. More specifically, he argues that the design of the *CPS*, which is used to obtain the contemporaneous measure of labour market stocks, is likely to result in a higher proportion of individuals with a much lower level of labour force attachment being included in the contemporaneous unemployment stock, than would be included in the measure obtained from the retrospective information.

The implications of these validation studies for the reliability of retrospective information contained in the national labour force surveys would appear to be quite ominous for studies of labour market dynamics based on unadjusted labour force data. For, they suggest that the use of such retrospective information in order to identify labour market turnover is likely to create a significant number of spurious transitions across labour market states, and in particular between unemployment and not in the labour force, due to the existence of significant levels of both recall error bias and misclassifications arising from the heterogeneous nature of the two surveys.

1.2 The Matched Files Approach.

The matched files approach exploits the rotating panel structure of the national labour force surveys, which results in a selected household remaining in the sample for a number of consecutive quarters before being replaced by a newly surveyed household. This rotating characteristic allows one then, to construct longitudinal data by matching records for the same individual across a number of consecutive periods (surveys), before he or she eventually leaves the sample. The majority of national labour force surveys currently adopt one of the following two forms of rotating schemes:

i) the *r-m-r* scheme in which each selected household unit (address) remains in the survey for *r* consecutive periods, is dropped for the next *m* surveys and then re-enters for the following *r* periods before leaving the sample. The Current Population Survey (CPS), the US labour force survey, is one example of this type of scheme, namely a 4-8-4 scheme carried out at monthly intervals. In any given month, researchers have in theory therefore, access to the previous month records for three quarters of the respondents, and are able to match files for 50% of the current months survey with the survey in same month in the next year.

ii) the *r-* scheme in which the unit is interviewed for *r* consecutive periods before leaving the sample. The Spanish LFS, *Encuesta de la Población Activa (EPA)* which is carried out on a quarterly basis, is one example of this type of scheme, with the household being interviewed for six consecutive quarters. Each survey is composed of six different rotation groups, thus in quarter *t* individual records belonging to rotation groups 2 to 6 can be matched with those of the previous quarter. In principle then, one is able to match 5/6 of the records between two consecutive quarters and 1/3 of the sample from one quarter to the same period in the following year.

In contrast to the retrospective question approach where one has access to retrospective information for almost the entire labour force survey, the matched files approach, by its very nature, suffers at the outset from a progressive loss of the panel component over time[7]. Despite this relative disadvantage, one would expect that the ability of the researcher using the matched files approach to track individuals over a specific length of time would eliminate the errors associated with the retrospective approach, thus producing a more reliable picture of the underlying dynamics of the labour market. Unfortunately, a sizeable literature in the US has demonstrated that, in practice, this tends not to be the case, with the matched files approach itself being subject to a number of specific problems, which unless corrected for, can have serious ramifications for the construction of labour market flows.

Sample attrition: occurs when a household unit which in principle should be covered by the survey, in other words which is not a member of the incoming or outgoing rotating segments, does not respond to the survey. This non-response can be due either to migration (a member of the household or the entire household unit itself leaving the sampled address)[8] or refusal or absence ('no one being at home when the interviewer called'). The extent of sample attrition can be quite significant; work by Abowd and Zellner (1985) for example, illustrates that approximately 7.5% of the previous month's respondents belonging to ongoing rotation groups cannot be located in the current month's survey, and approximately 7.5% of the current

[7]If the rotation process is random, as it should be by design, this progressive loss of the panel component will not however, introduce any bias into the estimation of the labour market gross flows.

[8]When "migration" occurs, the new occupants of the sampled address are included in the survey, in the same rotation group as the previous occupants, but flagged as new entrants.

month's respondents cannot be located in the previous month's survey. Thus on average 15% of all those individuals who in theory should be matched across consecutive months cannot be matched at all. In Spain the fraction of attriters is also sizeable, and exhibits a sharp seasonal pattern. Between 6% and 8% of all the individuals who in theory should be able to be matched across consecutive quarters cannot actually be matched. This percentage rises to above 11% between the second and third quarters due to the vacation period involved [9].

If this attrition process is not random, in other words, if the behaviour of those in the rotation segments of the sample differs systematically from those who remain, estimates of labour market turnover based on the unadjusted data will be biased. A considerable amount of evidence does in fact suggest that attrition may not be random. Peracchi & Welch (1995) for example, conclude that "far from being random, success in matching persons across *CPS* files is systematically related to observable characteristics". More specifically, attrition is found to be concentrated amongst youths, being the result of household and individual mobility due to schooling, family formation and job search[10]. The available evidence for Spain also suggests that sample attrition is actually a non-random process. The results of the Moreno & Toharia (1998) study, in which the authors estimate the logistic regression of the probability of an individual being interviewed five to six times as a function of a series of personal characteristics, indicating that attrition is more frequent in: women; individuals aged between 25 and 34 years old; those with higher levels of education;

[9] INE (1996).

[10] It should be noted however, that the study by Pitt (1988) on the distribution of unmatched households illustrates that on average 42% of non-matchable individuals during the period 1979-1983 were in fact movers, whilst only 10% of the non-matchables were due to the fact that nobody was at home when the interviewer called. This result is however, likely to be very dependent on the patterns of mobility prevalent in the US and thus cannot be generalised to other countries.

household heads; and less frequent for unemployed individuals without previous work experience. The bias produced by non-random attrition will naturally be of particular importance when using the matched file data to estimate a behavioural model in which attriters and non-attriters systematically differ in a way that is not captured by observable characteristics.

Misclassification errors: which arise as a result of respondent, miscoding, interviewer errors or matching procedures[11], are however the major pitfall that researchers have to overcome when using the matched files approach to obtain a measure of labour markets turnover. The ramifications of misclassification for both the correct measurement of labour market flows and for the estimation of behavioural equations will however, depend on the time series nature of the errors themselves. In general, classification errors do not result in severely biased estimates of the population labour market stocks, but may result in biased estimates of the flows between states in two consecutive periods. This can be seen more clearly, if one represents labour market transitions using a 3x3 contingency table (such as that of table 1). For, if the true state of an individual who has remained in the same state for both of the time periods under consideration is observed with error in either *t* or *t+1*, then the individual will be recorded as having changed labour market state, and a transition, albeit spurious, will be observed; thus increasing the cell counts of the off-diagonal elements of the table. In contrast, misclassification on just one occasion for an individual who is truly changing states, can result in the individual being shifted to either the diagonal or to another off-diagonal cell of the table. As the majority of individuals remain in the

[11]In the majority of national LFS the unique identification code is only allocated at the household level, since these surveys are by their very nature, household based surveys. Researchers are forced then to create "linking procedures" in order to be able to match records of the individual members of the household across consecutive surveys. Errors in classification will therefore occur, when an incorrect matches are made.

same labour market state (in other words the diagonal cell counts far outweigh those of the off-diagonals), one will however, providing misclassification errors in the two periods are not perfectly correlated, observe more movements from the diagonal cells to the off-diagonals, as opposed to from the off-diagonal to the diagonals. Misclassification errors which are independent across individual surveys will therefore, tend to increase the number of reported changes. Only in the case in which misclassifications errors are perfectly positively correlated across time, will the calculated gross flows be unbiased.

Table 1:

Average Monthly Gross Labour Market Flows 1977-1982
(figures in 000's)

	Final State		
Initial State	Employed	Unem-ployed	NILF
Employed	91,865	1,652	3,157
Unemployed	1,857	3,899	1,521
Not in the Labour Force	2,805	1,610	55,541

Source: Adapted from Table IV Poterba and Summers (1986)

2. How big are the errors associated with these problems likely to be?: an evaluation using Spanish labour market data:

To evaluate the accuracy and thus the usefulness of the retrospective and matched files approaches for studies of labour market dynamics, information as to both the magnitude and the nature of the bias of the errors arising from the problems associated with the aforementioned methodologies is required. In this section the Spanish labour force survey *Encuesta de Población Activa* (*EPA*), is used to assess the ramifications of using the raw LFS data for studies of labour market dynamics. For, despite its repeated use (see for example, Antolìn (1995, 1996), Garcia Serrano (1996), Bover et al (1996), Alba (1996, 1997) and Moreno and Toharia (1998)), little is known about the error structure of *EPA* survey data. What little knowledge we have, is to be gained from the quality control tabulations published at regular intervals by *INE*.

2.1 The Retrospective Approach:

In addition to its standard format, the *EPA* has in every second quarter since 1987 included a small supplementary section of questions referring to the individual's labour market status 12 months prior to the current survey[12]. It is this retrospective information which has been used to identify an individuals transitions between the various labour market states between the two time intervals. Considerable differences with respect to both survey design and classification procedures exist, between the contemporaneous and retrospective Spanish surveys. In the regular *EPA* questionnaire, where responses to a much wider set of questions with respect to current labour market activity are used each quarter by the *interviewer* to assign labour market status, the *respondent* himself assigns the individual household members labour market status by classifying them into one of a number of predetermined categories in the retrospective supplement to the survey. More specifically, *current* labour market status is determined by the interviewer using declared information on a number of labour market characteristics, such as: labour market activity; job search activity and availability, contained in the main body of the survey to determine the individual's current labour market status. Previous labour market status on the other hand, is ascertained in an entirely different manner, simply by the respondent's allocation of individual household members to one of the following states: i) employed; ii) searching for employment; iii) available for work, but not actively searching; iv) military service; v) studying; and vi) other situations. Thus an individual is classified as: being out of the labour force if he is assigned to categories 3, 5 or 6; unemployed if he is assigned to category 2; and employed if assigned to category 1[13]. The Spanish

[12]See appendix 1 for a detailed description of the questions contained in the retrospective supplement relating to the respondent's situation one year prior to the current survey.

[13]The *EPA* regularly classifies men in the military service in a special category denoted *Poblacion Contada Aparte*.

retrospective section contains no further information with respect to search activity, despite the fact that this issue is of fundamental importance to the contemporaneous classification of unemployment; the non-employed individual simply being asked in the retrospective part of the survey to declare whether he was looking for work, with no reference being made to the quality/extent of job search activity.

In this sub-section a combination of two different sets of information: i) the aforementioned retrospective supplement of the labour force questionnaire; and ii) data obtained by matching individual records across various quarters of the labour force survey are used to illustrate the magnitude of errors arising from the problems associated with the use of retrospective information, and thus ultimately the appropriateness of this approach as a means of measuring labour market flows.

This type of validation procedure is feasible, since the longitudinal characteristics of the *EPA* ensures that: i) in the second quarter of each year the *EPA* contains information on the labour market status 12 months ago of the *N* individuals interviewed in the current quarter; and ii) the rotating panel nature of the *EPA*, ensures that, in the absence of attrition, approximately one third of the current sample would have also been interviewed in the second quarter of the previous year, year *1*. Thus for *N/3* individuals, one has access to two different sources of information as to their labour market status in the second quarter of year *1*: i) the *reported* status during the interview which took place in year 1; and ii) the status which the individual *recalls* in year *2*, he was in 12 months prior to the current interview. A comparison of these two sources of information provides us therefore, with a measure of the size and bias of the errors one encounters when attempting to construct flow data using the retrospective survey.

Table 2:

Actual and Recalled Labour Market Status of *EPA* Individuals

in the Second Quarter 1992.
(figures in 000's.)

Recalled LMS **Actual LMS**	Working	Searching	Not in the Labour Force	Military Service(*)	Total
Employed	3350.4	70.4	77	5.1	3502.9
Unemployed	69	549.4	122.8	2.5	743.7
Not in the Labour Force	51.8	128.4	4268	4.5	4452.7
Military Service (*)	3.2	6.2	4.7	43.4	57.5
Total	3474.4	754.4	4472.5	55.5	8756.8

Source: Table II.1 INE (1996). (*) *Población Contada Aparte*

In table 2, we reproduce statistics published in *INE* (1996) on *recalled* and *reported* labour market status (LMS) for the matched files of the second quarters of 1992 and 1993. If we assume that the reported LMS reflects the individual's *true state*, whilst *recalled* LMS is taken as an indicator of the true LMS, the fundamental question is then, how good an indicator is *recalled* LMS of an individual's *true* LMS? One simplistic measure of the overall accuracy of this indicator is the global index of agreement $\prod_0 = \sum \prod_{ii}$, where $\prod_{ii}$ is the proportion of individuals which are classified in state i both in the actual and recalled status. In the absence of errors, all the off-diagonal terms in the recalled-actual LMS matrix should be equal to zero, and $\prod_0$ will be 1. Using the data in table 2, a value of 0.94 is obtained for this index of agreement, implying then that for the period second quarter 1992 to second quarter 1993 recalled labour market status would in fact appear to be a reasonable indicator of true labour market status[14].

[14]*INE* in fact state that the *"La proporcion de identicamente clasificados es elevada, por lo que se pueden considerar fiables los resultados derivados de la*

A better measurement of agreement is however, the index *kappa* defined as:

$$Kappa = \frac{\Pi_o - \Pi_e}{1 - \Pi_e},$$

where Π_e represents the degree of agreement which is expected purely by chance. Thus if the two classifications (actual and recalled status) were statistically independent one would obtain that:

where

$$\Pi_{ii} = \Pi_{i+} \cdot \Pi_{+i} \ ,$$

$$\Pi_{i+} = \sum_j \Pi_{ij} \qquad \text{and} \qquad \Pi_{+i} = \sum_j \Pi_{ji}$$

In this case, the index Π_e defined as:

$$\Pi_e = \sum_i \Pi_{i+} \, \Pi_{+i} \, ,$$

will be equal to Π_o.

The numerator in the definition of kappa provides us therefore, with the excess of agreement over the degree of agreement expected purely by chance (i.e. Π_e).

In the denominator Π_o is replaced by 1, its maximum possible value. Thus the index kappa relates the observed excess of agreement (over that which is produced merely by chance) to the theoretical maximum value of the excess of agreement, i.e. (1-Π_e).

---

preguntas reprospectivas. Refuerza aún más esta afirmación la comparación de las clasificaciones marginales, que son extermadamente parecidas", INE (1996).

If one considers the two extreme values of this index, we have that: i) in the case of total agreement (in other words, zero inconsistencies), $\prod_0 = 1$, and thus $Kappa = 1$; ii) in the case of statistical independence, where the degree of agreement equals that expected by chance, $Kappa = 0$.

Using the information on recall and actual labour market status presented in table 2, an estimated value of $Kappa = 0.89$ is obtained, with the associated standard error being equal to 0.0063[15].

Although the fact that one obtains: i) values close to 1 for the two measures of overall agreement we have considered; and ii) very similar marginal distribution of observations across the three states, would seem to suggest that *recalled* LMS is actually quite a good proxy of the *true* LMS at any point in time, such measures are of little value when one attempts to assess the appropriateness of *recalled* LMS as the indicator to be used in the construction of labour market flows data. For, as is evident from table 2, in a number of cases there are in fact quite large discrepancies between *true* LMS and *recalled* LMS. More specifically, of those individuals who were actually employed 12 months prior to the current survey, 4.3% recall having been in another situation, and of those who were really unemployed, 26% recall being in another labour market state.

The ramifications of these inconsistencies for an analysis of labour market dynamics becomes more apparent if one compares the actual and spurious transitions across labour market states over the two sample periods presented in table 3. In panel A we report the observed transition probabilities between the three states. Panel B, calculated using the data of table 2, contains the spurious transitions caused by recall errors. It is evident that, particularly in the case of transitions from unemployment, recall errors are relatively important.

[15]Note that in the case of multinomial sampling, the estimated *Kappa* has a large-sample normal distribution for which we can compute its asymptotic variance (see Agresti (1990), pp. 366-367).

For, whilst the observed rate of labour force withdrawal from unemployment from panel A, is equal to 8.5%, one sees from panel B that 17% of those whose initial state was considered to be unemployment were really out of the labour force. It is also worth emphasising that a number of transitions, other than those between the traditionally ambiguous labour market states of unemployment and not in the labour force, would also appear to be subject to relatively important errors rates. The probability of an individual who is actually observed at *t-12 months* as being unemployed and as employed 12 months later is for example, 27%. Note however, that the data in panel B indicate that almost 10% of the people initially classed as unemployed were in fact working.

Table 3:
Actual and Spurious Transitions over the Period 2/1992 - 2/1993

Recalled LMS 2/93	Employed	Unemployed	Not in the Labour Force
Employed	0.875	0.082	0.043
Unemployed	0.272	0.643	0.085
Not in the Labour Force	0.028	0.034	0.938
Recalled LMS	Employed	Unemployed	Not in the Labour Force
Employed	0.965	0.020	0.015
Unemployed	0.094	0.734	0.172
Not in the Labour Force	0.017	0.027	0.955

Source: adapted from Tables II.1 and II.4 *INE* (1996)[16].

The use of the retrospective part of the labour force survey for both the construction of labour market gross flows and for analysing labour market dynamics would in light of these preliminary results appear therefore, to be somewhat problematic. The combination of heterogeneous sample design and recall error being likely to generate a significant number of spurious transitions, particularly between the stock of unemployed and out of the labour force. In addition to this, when using the retrospective information to construct a measure of labour market dynamics one also has to overcome the problem of missing data. In the second quarter of 1991 for example, almost 3% of the respondents in the regular *EPA* did not respond to the set of retrospective questions. As previously discussed, if these observations are not missing at random, a further bias will be introduced into the gross flows data.

2.2 The Matched Files Approach:

[16]Those individuals recorded as being in the military in either of the two periods have been dropped from the analysis.

As already mentioned in section 1.2, the panel structure of the Spanish *LFS*, is such that one is able to construct longitudinal data by matching records for the same individual across six consecutive quarters (surveys) before his or her's household eventually leaves the sample. Thus if we have *N* individuals answering the survey at time *t*, we are, in principle, able to follow the labour market history of *N/6* of the sample for 1.5 years or alternatively *(5/6)*N* for two consecutive quarters[17].

In this section we use the second quarter 1994 wave of the Spanish Labour Force Survey, together with previously unexploited information from the *EPA* re-interview survey, to obtain additional information as to the characteristics and magnitude of misclassification errors in *EPA* data, and the subsequent ramifications for gross flows data constructed using unadjusted data from the longitudinal sections of the *EPA*[18].

In common with the reinterview system of the US Current Population Survey (*CPS*), the Spanish reinterview survey introduced in 1970, is used principally as a means of controlling the quality of both the underlying data and the work of the interviewers conducting the survey. As in the US, this survey consists of reinterviewing a subsample of the original *EPA* respondents approximately 15 days after the original interview takes place[19]. Respondents of the reinterview

---

[17]It is important to note however, that *INE* argues that: *"En la practica, sin embargo, por razones de representatividad de la muestra, con el fin de asegurar que la parte renovada de la misma tenga iguales caracteristicas que la parte que permanece, el seguimiento debe limitarse a cinco trimestres".* In other words, in order to maintain the representatives of the sample, one should limit oneself to five quarters only.

[18]We do not therefore, address the problem of sample attrition briefly discussed in section 1.2, since it is widely accepted that in these types of rotating panel surveys, misclassification is a more important problem than sample attrition.

[19]In both the *CPS* and *EPA* the reinterview survey is carried out on approximately 5% of the sample of the regular survey. In the *CPS*, the reinterview procedure takes place approximately one week after the original survey

survey are once again asked to respond to the same set of questions used in the regular labour force survey, although the responses refer to labour force activity in the original survey period, and not to the reinterview period. The reinterview serves two main purposes: i) to evaluate field work (e.g. to identify interviewers who either misunderstand or falsify the data); and ii) to estimate error components (e.g. to estimate both simple response variance and response bias). In contrast to the *CPS* procedure however, the Spanish reinterview procedure does not involve any *reconciliation process* in order to identify the correct response when inconsistencies appear between the original and reinterview surveys[20]. As a result, one is unable to assume that the Spanish reinterview process correctly identifies the individual's true labour market state. This inability, as discussed in due course, does tend to complicate attempts to calculate the magnitude of errors present in the original survey somewhat. As the reinterview is carried out by more experienced interviewers, one can however, assume that the data are of superior quality to that collected in the original survey, and thus that the extent of error is smaller in the follow-up survey.

[20]In the US, the reinterview procedure is divided into 2 sub-samples: the reconciled and unreconciled components. In the reconciled case, which involves approximately 75% of the reinterview survey, interviewers are provided with the original interview. Thus after having conducted the second interview, the reinterviewer compares the responses given in both the original and reinterview surveys. Any discrepancies uncovered are then reconciled *with the respondents* before leaving the sampling unit. In the unreconciled sub-sample on the other hand, the interviewer has no such additional information, thus no attempt is made to compare results across the two surveys.

Survey Response Inconsistencies in the *EPA.*

1. Observed Transition Probabilities First to Second Quarter 1994				
	Employed	Unemployed	Not in the Labour Force	Total
Employed	0.947	0.035	0.018	1.000
Unemployed	0.139	0.786	0.075	1.000
NILF	0.020	0.036	0.944	1.000
2. Spurious Transition Probabilities from the Interview and Reinterview Surveys				
	Employed	Unemployed	Not in the Labour Force	Total
Employed	0.935	0.030	0.035	1.000
Unemployed	0.082	0.770	0.148	1.000
NILF	0.029	0.049	0.923	1.000
3. Survey Response Inconsistencies. Reinterview Survey				
	Employed	Unemployed	Not in the Labour Force	Total
Employed	0.327	0.011	0.012	0.350
Unemployed	0.010	0.092	0.018	0.119
NILF	0.015	0.026	0.490	0.531
Total	0.352	0.128	0.520	1.000

Source: INE (1995).

Information from the reinterview survey enables one then to conduct a cross check of a number of key labour market variables, and can be used to provide an estimation of the magnitude of erroneous responses contained in the regular LFS[21]. Cross-tabulating the distribution of recorded labour market status for the subset of the individuals who respond to both the *EPA* interview and the reinterview surveys for example, provides one with an idea of the incidence and magnitude of the misclassification errors present in the original survey. In the absence of misclassification errors, the respondents labour market status in the original *EPA* and the reinterview survey should be identical: i.e. the off-diagonal elements of the interview-reinterview table should be equal to zero. The off-diagonal elements corresponding therefore, to inconsistent responses.

In the first instance then, it is evident, that the use of raw LFS data does not pose a significant problem for the estimation of populations stocks: in that the differences in the marginal distributions across the two surveys are minimal. In contrast, the off-diagonal values of panel 3, which in this case indicate that more than 9% of respondents are classified in a different labour market state across the two surveys, suggest that misclassification error can seriously overstate the degree of labour market turnover across two time periods. This is clearly evident from the data presented in panel 2, where a comparison of individual responses for labour market status in the original and reinterview surveys, illustrate the magnitude of the misclassification problem for the accurate measurement of the flows between labour market states. These figures are particularly alarming, when one compares them to the observed quarterly transitions be-

[21]It is important to note however, that the percentage of *true* misclassifications is likely to be higher than that detected by the reinterview procedure, since quality control procedures of this nature are unable to detect, and therefore correct for, misclassifications which arise either as a result of incorrect information supplied to the interviewer or because of ambiguity of responses due to poor survey design.

tween two consecutive *EPA*, since the transitions in panel 2 all of which are spurious, are very large in size relative to observed quarterly transitions reported in panel 1.

It is interesting to note that the magnitude of inconsistencies with respect to labour market status does however, differ across gender, with, as is evident from table 5, 8% of men and 10% of women being misclassified in one of the two surveys. It is also evident that the degree of inconsistency of response between the two surveys is, as one might expect, higher the more complex is the actual concept being measured: the size of the error between unemployment and not in the labour force being particularly high for women, a group known to have a lower level of labour force attachment[22].

[22]Since exactly the same questionnaire is used in both the original and reinterview surveys, it is in fact in principle possible to check for the consistency of a very large number of variables. INE does in fact publish on an annual basis a number of consistency indicators, which are computed for a sub-set of the labour force survey using the original interview and re-interview data. As one might expect the degree of inconsistency tends to increase with the complexity of the concept being measured. If one considers the Global index of agreement, $\prod_0$, as the simplest consistency measure, in 1994 the value of this index ranges from 0.98 for marital status (the variable with the highest degree of consistency) to 0.60 for the elapsed time searching for a job (which according to
the re-interview survey is the variable measured with the greatest amount of the error). The values of the Globax index of agreement are very similar from one year to another.

Table 5:
Labour Market Status in Original and Reinterview Survey (First quarter 1994)[23]

	Reinterview			
Original Interview	Employed	Unem-ployed	Not in the Labour Force	Total
Employed	2007 (0.3270)	65 (0.0106)	75 (108)	2147 (0.3498)
Unemployed	60 (0.0098)	562 (0.0916)	108 (0.0176)	730 (0.1190)
Not in the Labour Force	93 (0.0152)	159 (0.0259)	3008 (0.4901)	3260 (0.5312)
Total				6137 (1.000)
Men	Employed	Unem-ployed	Not in the Labour Force	Total
Employed	1367 (0.4691)	41 (0.0141)	31 (0.0106)	1439 (0.4938)
Unemployed	36 (0.0124)	322 (0.1105	39 (0.01338	397 (0.1362
Not in the Labour Force	39 (0.01338)	50 (0.0172)	989 (0.3393)	1078 (0.3699)
Total				2914 (1.000
Women	Employed	Unem-ployed	Not in the Labour Force	Total
Employed	624 (0.1964)	24 (0.0076)	44 (0.1385)	692 (02178)
Unemployed	22 (0.0069)	237 (0.0746)	68 (0.0214)	327 (0.1029)
Not in the Labour Force	54 (0.0170)	108 (0.0340)	1996 (0.6283)	2158 (0.6793)

[23] Figures in parenthesis refer to sample proportions.

Total	3177 (1.00[illegible]

Source: INE (1995).

These findings clearly underline the need to adjust the underlying "raw" flow data obtained from matching consecutive waves of LFS, before such data can be used in an analysis of labour force dynamics. Adjustments using information from reinterview surveys have in fact been undertaken in a number of studies in the US, (see for example, Abowd & Zellner (1985), Chua & Fuller (1987), Fuller & Chua (1985) and Poterba & Summers (1985)) to correct for the presence of misclassification errors in the *observed* gross labour market flows between two points in time. Underlying the general approach to adjustment taken in the US is the key assumption that classification errors are independent (*ICE*). In other words that the probability of being misclassified depends only on the individual's current classification. In such circumstances, the true labour market flows, F^*, between $t-1$ and t are related to the observed flows, F, in the following manner:

$$\mathrm{Vec}\hat{F}^*_{(t-1,t)} = \left[\hat{K}^{-1}_{(t)} \otimes \hat{K}^{-1}_{(t-1)}\right]\mathrm{Vec}F_{(t-1,t)} \qquad 1)$$

where: $F_{(t-1,t)}$ denotes the observed gross flows matrix between period *(t-1)* and *t*. The *ij*th entry of which is F_{ij}: i.e. the number of people observed in state i in period *(t-1)* and in state *j* in period *t*; and $F^*_{(t-1,t)}$ the true unobserved gross flow matrix. The vector, *Vec F*, denotes the (9*1) vector obtained from stacking the rows of matrix *F* one below the other, the same applies to $Vec\ F^*$. $\hat{K}_{(t)}$ is the response probability matrix, whose *ij*th entry, K_{ij}, is the probability in period *t* that an individual truly in state *j* is classified in state *i*.

In such a framework, the observed flows from, for example, unemployment to employment will be the result of the following nine different possible true transitions[24]:

$$
\begin{aligned}
F_{UE} = {} & K_{UE} K_{EE} F_{EE}^{*} + K_{UE} K_{EU} F_{EU}^{*} + K_{UE} K_{EN} F_{EN}^{*} \\
& + K_{UU} K_{EE} F_{UE}^{*} + K_{UU} K_{EU} F_{UU}^{*} + K_{UU} K_{EN} F_{UN}^{*} \\
& + K_{UN} K_{EE} F_{NE}^{*} + K_{UN} K_{EU} F_{UN}^{*} + K_{UN} K_{EN} F_{NN}^{*}
\end{aligned}
\qquad 2)
$$

According to equation (2), regardless of the actual labour market state in periods t and *(t-1)*, any individual can be seen, with a certain probability, as having made a transition between employment and unemployment between these two periods. It is important to note however, that some of these probabilities will be negligible: the *ICE* assumption for example, results in a very small number of individuals being misclassified in the two periods.

The existing studies of the adjustment of gross flows data do however, differ substantially in their estimation of the probability K of being misclassified. More specifically, in using only the information from the reconciled sample of the *CPS* reinterview survey Abowd and Zellner (1985) are implicitly assuming in their adjustment procedure, that the correct labour market status is revealed by the reinterview survey. Thus their matrix K is simply the interview-reinterview table (*after reconciliation*) for the reconciled sub-sample, normalised so that each column adds up to unity. There is some concern however, as to the quality of the reconciled sub-sample, which casts doubts on its ability to determine the individual respondents true labour market status. These doubts suggest that adjustments based solely on the use of information taken from the reconciled sub-sample may produce a downward bias in the incidence of misclassifications.

[24]For simplicity, all time indices have been dropped in this section unless otherwise stated.

For, if the procedure specified for the reconciled reinterview survey is correctly carried out, one should observe that the incidence of discrepancies in the unreconciled sub-sample are similar to the incidence of discrepancies in the reconciled (before reconciliation). If this were indeed the case, then the Abowd-Zellner adjustment procedure would correctly account for the incidence of misclassifications. In practice however, the incidence of discrepancies between the original and reinterview surveys is much greater in the unreconciled than in the reconciled sample (before reconciliation); this anomaly indicates that reinterviewers “cheat” by using information from the original survey to complete the reconciliation process. In order to allow for such differences in the incidence of discrepancies between the reconciled and unreconciled samples, Poterba and Summers (1986) also use, in addition to the *CPS* error classification rates obtained from the reconciled sample, information from the unreconciled sample to estimate the incidence of errors. The elements of the *K* matrix in the Poterba and Summers' study are therefore, composed of two components: the probability of an error occurring obtained from the *CPS* reconciled reinterview survey and the incidence of errors estimated from the unreconciled sub-sample.

As *INE* does not have carry out a reconciliation procedure, it cannot be assumed here that the reinterview process correctly identifies an individual's true labour market status. Estimates of the error response probabilities are therefore obtained using the third method available, proposed by Fuller-Chua (1985, 1987), which provides a means for estimating the error matrix *K*, when the true labour market state of the individual cannot be readily identified. In view of the aforementioned concern as to the quality of the reconciled reinterview data, Fuller and Chua designed a parametric procedure to estimate the response matrix, K, based on the unreconciled sub-sample only[25].

Estimating the probability of being misclassified with unreconciled data:

It is assumed that at any point in time, each individual in the sample can be classified into one of the three labour market states: employment (E), unemployment (U) and not in the labour force (NILF), and the classification process is subject to error. Let $K(t)$ represent the response probability matrix at time t. The probability of being misclassified in state i given that the individual's true labour market state is j, is the ijth entry of the matrix K. By construction, the sum of the individual columns of this probability matrix sum to unity.

Denote: the (3*1) vector of *observed* marginal proportions of the population in each state (E, U, NILF) in a sample of n observations by $\hat{P}$; and the (3*1) vector of the *true* marginal proportions by P. It is assumed that the misclassification process only affects the distribution of the interior cells, thus that the marginal proportions are an unbiased estimator of the true proportions: $E(\hat{P}) = P$. Such an assumption would not appear that unreasonable, given that the LFS are actually designed to measure labour market stocks, the construction of flows being in a sense, a by-product. Moreover, the assumption of biased response errors would itself be equivalent to assuming that the existing estimates of the stocks of E, U and $NILF$ are themselves biased. The assumption of *Unbiased Response Error (URE)* implies that:

[25]In this section, only a brief overview of the correction methodology proposed by the authors is presented. For a more in-depth analysis the reader is referred to the 1985 and 1987 papers cited in the bibliography.

$$E\left(\hat{P}\right) = K\ P\ = P$$

Assuming unbiased response errors, the matrix of error response probabilities, *K*, can be estimated in the following manner. Let $\hat{A}$ be the observed interview-reinterview table (e.g. the third panel in Table 4). Â is therefore, the two-way contingency table in which individuals are classified on two separate occasions into 3 mutually exclusive classes (E, U, NILF). As previously mentioned, it is known that the classification process is itself subject to potential errors in both the original interview and the reinterview. It is assumed in the first instance that: i) the misclassification process in the interview is independent from the classification process in the reinterview; and ii) that the error response probabilities are identical in the two trials. In this set-up one has that the *population* interview-reinterview table, the matrix *A*, can be expressed as:

$$\mathrm{A} = \mathrm{E}\{\hat{\mathrm{A}}\} = \mathrm{K}\left[\mathrm{diag\ P}\right]\mathrm{K}'. \qquad 3)$$

Note however, that given that: i) *K* is a transition probability matrix (i.e. $K' \cdot i = i$, where i is the (3*1) unit vector); and ii) the assumption of *URE* ($KP = P$), it follows that the marginal distributions of *A* are equal to the population marginal distribution *P*. Thus:

$$\begin{aligned} Ai &= K\left[diagP\right]K'i \\ &= K\left[diagP\right]i \\ &= KP \\ &= P \end{aligned}$$

Note also that if equation 3 holds, then the matrix *A* should be symmetric.

An estimation of the error response probabilities, K, can therefore be obtained directly from the estimation of equation 3. Note however, that: i) as each individual column of the matrix K sums to unity, one needs only to estimate the elements of two of the three rows of the K matrix; and ii) as $KP = P$, only two columns in K need to be estimated. This in total 4 parameters of the matrix K need to be estimated. As the matrix A is symmetric and its row (and column) marginals equal the population marginal distributions, given P the matrix A contains only three free parameters. In other words, it does not contain sufficient information to estimate the response error matrix K. Some additional parameterisation is therefore, required. Thus the following parametric form is assumed for K:

$$K_{ij} = \left[1 - \sum_{1=1}^{r} a_{tj} P_t \left(P_t + P_j\right)^{-1}\right] d_{ij} + a_{ij} P_j \left(P_j + P_i\right)^{-1}, \; i, j = 1{,}2{,}3, \quad 4)$$

where: d_{ij} is Kronecker's delta; i.e. $d = 1$ if $i = j$, and $d = 0$ if $i \neq j$; P are the true marginal proportions; $a_{ii} = 0$; $\alpha_{ij} = \alpha_{ji}$; and a_{ij} are constants, taking a value of between 0 and 1.

The probability of being misclassified is therefore a function of the parameters a and the marginal proportions, P. The assumption of symmetry, i.e. $a_{ij} = a_{ji}$, implies that the probability of misclassifying in state i, an individual who is really in state j, is offset by the probability of classifying in state j an individual who is actually in state i. The parameters a can therefore, be considered to represent an index of the probability of making a misclassification error.

The matrix of error probabilities K is therefore, such that the probability of misclassifying an individual is proportional to the true proportions between states i and j. For example, the probability of being incorrectly classified in state 1 when the individual's true labour market state is 2, is given by:

$$K_{12} = a_{12} \frac{P_1}{P_1 + P_2}.$$

In other words, the greater the number of individuals in state 1 relative to state 2, the higher will be the probability of misclassifying someone into state 1 when he is really in state 2.

The functional form chosen for K also implies that this probability is, in a specific sense, symmetric to the probability of misclassifying in state 2 someone who is really in state 1:

$$K_{21} = a_{12} \frac{P_2}{P_1 + P_2}.$$

Taken together these two assumptions imply that the greater the number of individuals classified in state 1, relative to state 2, then: (i) the higher will be the probability of misclassifying someone in state 1 when he is really in state 2; and (ii) the lower will be the probability of misclassifying someone in state 2 when he is really in state 1. An alternative way of viewing these assumptions is that the relative probability K_{12}/K_{21} is *equal* to the relative proportions P_1/P_2,

Finally, note that by construction, the probability of being classified in states 1, 2 or 3 given that the true state is 2, is $K_{12}+K_{22}+K_{32}=1$.

Before proceeding to the estimation of the model itself, it is also important to note that the assumption of equal response probabilities across independent samples implicit in the K formulation, would appear to be rejected by the data presented in the interview-reinterview table (table 4)[26]. This rejection implies that in our case

[26]One interesting phenomenon which appears to be true in both the US and Spain, is that the percentage of people classified as unemployed in the reinter-

then, the A matrix will not be not symmetric and that equation 3 no longer holds.

To accommodate the presence of a non-symmetric interview-reinterview table, the error response probabilities are allowed to vary across the two trials by assuming a different functional form for the error rates in the two surveys. More specifically, the lack of symmetry in the observed interview-reinterview table is assumed to be due to a shift in the proportion of individuals falling into the three labour market states. The model is therefore, expanded to allow for the possibility of differing marginal proportions across the two trials. More specifically, the error response matrix K, as given in equation 4, is assumed to hold for the original interview, and the following additional matrix of error response probabilities, G, is considered for the reinterview survey[27]:

$$\begin{aligned} G_{ij} &= J_{ji} K_{ij} && \text{if } A_{ij} < A_{ji} \text{ and } i \neq j \\ G_{ij} &= K_{ij} && \text{if } A_{ij} \geq A_{ji} \text{ and } i \neq j \qquad 5) \\ G_{ij} &= 1 - \sum_{i=1}^{r} G_{ri} && \text{if } i = j, \end{aligned}$$

where $J_{ij} \geq 1$.

To reiterate then, the idea behind this formulation is simply that the assumption of symmetry of the interview-reinterview table breaks down because of a pure shift in the proportion of individual classified between these two states across the two samples. Implicit in this functional form is the fact that we are only focusing on the possibility that, for example, an individual classified as unemployed in the interview and as employed in the re-interview is truly either em-

view is larger than in the original interview. For detailed figures see *INE* (1995) pp 47-48.

[27]The functional form chosen for the errors probabilities of the reinterview survey, G, is obviously not the only possible one, but it is the one which produces smaller estimates for alpha(i,j) than other models considered.

ployed or unemployed. Thus it is assumed that an individual is always correctly classified in one of the two surveys, and that it can never be the case that the individual of the previous example was actually in the third state: out of the labour force.

As an example, consider the case in which $A_{12} < A_{21}$. The corresponding elements of G will be such that:

$$G_{12} = J_{12} a_{12} \frac{P_1}{P_1 + P_2},$$

and

$$G_{21} = a_{12} \frac{P_2}{P_1 + P_2}.$$

Thus the 'symmetry' that existed between K_{12} and K_{21} is broken by shifting up the probability of misclassifying as being in state 1, an individual who is truly in state 2 in the re-interview, but keeping constant the probability of misclassifying as 2 a true 1.

In defining G, in this manner we are in fact assuming that if more individuals are placed in the ji-th cell than in the ij-th cell of the interview-reinterview table, the probability that an individual in true category j reports category i on the reinterview is larger than the probability that an individual in true category j reports category i on the original interview.

Incorporating this extension into the model, equation 3 becomes:

$$E\{\hat{A}\} = K\,[diag.P]\,G'. \tag{6}$$

6)

The parameters a, J, p, embedded in equation 6 are then estimated using a generalised non-linear least squares procedure (described in greater detail in appendix 2), assuming that the proportions A_{ij} follow a multinomial sampling.

Table 6:
Estimated Parameters.

Parameter	Estimated Value	Standard Error
Males and Females		
P1	0.3498	0.0044
P2	0.1189	0.0030
a_{21}	0.0588	0.0053
a_{31}	0.0292	0.0024
a_{32}	0.0987	0.0064
J_{12}	1.1502	0.2598
J_{13}	1.4715	0.2743
J_{32}	1.8693	0.2397
Males		
P1	0.4938	0.0069
P2	0.1363	0.0046
a_{21}	0.0612	0.0070
a_{31}	0.0252	0.0033
a_{32}	0.0724	0.0080
J_{12}	1.2628	0.3497
J_{13}	1.5290	0.4354
J_{32}	1.5240	0.3627
Females		
P1	0.2178	0.0053
P2	0.1029	0.0039
a_{21}	0.0536	0.0081
a_{31}	0.0427	0.0046
a_{32}	0.1324	0.0106
J_{12}	1.1722	0.4384
J_{13}	1.4468	0.3523
J_{32}	2.0691	0.3169

Substituting the estimated parameters reported in table 5 into equation 4 an estimation $\hat{K}$ of the error response matrix K is obtained. In table 7 we present the estimated error response rates relating to the first quarter 1994 *EPA* survey. Initially error rates were estimated for the entire sample population. From these results, reported in the top panel of the table, it is evident that the highest error rates are obtained for the unemployed: the estimated probability that an individual who is actually unemployed, being incorrectly classified as not in the labour force being 8.1%. Somewhat surprisingly however, the incidence of misclassification of those individual truly unemployed, as being employed, is also significant: an estimated 4.4% of the truly unemployed are found to be classified as employed.

Table 7:

Estimated Error Probability Rates in the *EPA*:
(First Quarter 1994)

True Class			
Reported Class	Employed	Unemployed	NILF
Employed	0.9675	0.0439	0.0116
Unemployed	0.0149	0.8754	0.0181
NILF	0.0176	0.0807	0.9703
Total	1.00	1.000	1.000
Males	Employed	Unemployed	NILF
Employed	0.9760	0.0480	0.0144
Unemployed	0.0132	0.8991	0.0195
NILF	0.0108	0.0529	0.9661
Total	1.00	1.000	1.000
Females	Employed	Unemployed	NILF
Employed	0.9505	0.0364	0.0104
Unemployed	0.0172	0.8486	0.0174
NILF	0.0323	0.1150	0.9722
Total	1.00	1.000	1.000

The results of the disaggregated analysis presented in the lower panels of the table, do suggest a number of important differences in the incidence of errors across the sexes. In particular, as one might expect given the lower labour force attachment of this group: i) the probability of an unemployed individual being misclassified as not in the labour is significantly higher for women than men: the respective error probabilities being 11.5% and 5.3%; and ii) the probability of a female being misclassified as not in the labour force, when she is actually employed is also considerably higher (3.2%) than that for males (1.1%). Such differences tend to emphasise the necessity of using disaggregate information on the error structure of the underlying data, when working with unadjusted gross flows data.

Table 8:

A Comparison of the Estimated Error Response Matrices in the *CPS* and *EPA*

EPA				
		0,967	0,044	0,012
Artola-Bell	K=	0,015	0,875	0,018
(March-June 1994)		0,018	0,081	0,970
		1,000	1,000	1,000
CPS				
		0,977	0,038	0,012
Poterba-Summers	K=	0,005	0,848	0,006
(January-June 1981)		0,017	0,115	0,982
(Table III)				
		1,000	1,000	1,000
		0,984	0,055	0,019
Fuller-Chua	K=	0,003	0,842	0,008
January 1979		0,013	0,103	0,974
(Table 5)				
		1,000	1,000	1,000
Abowd-Zellner		0,988	0,019	0,005
I/1977-4/1982	K=	0,002	0,886	0,003
(Table 6)		0,010	0,095	0,992
		1,000	1,000	1,000

In table 8 we compare the aggregate response error matrix estimated with *EPA* data with those obtained using *CPS* data in the Poterba & Summers (1986), Chua & Fuller (1985) and Abowd & Zellner (1985). In the first instance, it is interesting to note that despite considerable differences in the methodology employed in the US studies, the estimated error response probability matrices are surprisingly similar[28]. Secondly, there does not appear to be significant differences between the estimated error rates for the *CPS* and those computed using Spanish data. In that two main types of errors are consistently invariant across all the reported estimated error probability matrices. Firstly, the highest error obtained in each of the studies, relates to those individuals who are really unemployed, but incorrectly

[28]Note however, that the Abowd-Zellner results are in general, due to the methodology adopted for the computation of their *K* matrix, smaller than the rest.

classified as not in the labour force: the size of this error being relatively similar across studies, ranging from 8.1% in the *EPA* data to 12% in the Poterba & Summers study. Secondly, the misclassification of unemployed individuals as employed ranges from 5.5% in the Chua & Fuller paper to 1.9% in Abowd & Zellner. The values obtained for the estimated error rates would tend to suggest then, that when analysing labour market flows one should be particularly careful when dealing with transitions from and to unemployment.

3. Adjusting the Observed Flows:

The results of the previous section clearly underline the need to account, before such data can be used in an analysis of labour force dynamics, for errors in misclassification in the underlying "raw" flow data obtained from matching consecutive waves of LFS. As discussed, adjustments have in fact been undertaken in a number of studies in the US (see for example, Abowd & Zellner (1985), Chua & Fuller (1985) and Poterba & Summers (1985)) using information from reinterview surveys. In this section therefore, we apply the adjustment methodology developed in the US, in conjunction with the error response probability matrices estimated in section 2.2, in order to obtain estimates of the true Spanish labour market flows across the first and second quarters of 1994.

Recall that the fundamental assumption u*nderlying this adjustment methodology is that of independent classification errors (ICE).* To understand better the role played by the *ICE* assumption, we consider in the first instance the general case in which no particular assumption is made with respect to the stochastic structure of the response errors.

Denote the (9*9) error response matrix of the true gross flows matrix by C. The ij,kl-th element of which is given by:

$C_{ij,kl}$(t-1,t) = P[observed flow *(t-1)- t* is from i to j|true flow *(t-1)-t* is from k to l], where: i,j,k,l= *E, U, NILF* [29].

The element $C_{ij,kl}$. is therefore, the conditional probability of being observed transiting from state i to state j across the period t to *(t-1),* when the individual's true labour market transition is actually from state k to l.

The relationship between the observed and true flows is given by:

---

[29] Assuming lexicographic ordering of rows and columns.

$$VecF_{(t-1,t)} = C_{(t-1,t)} VecF^{*}_{(t-1,t)} . \qquad 7)$$

Given the observed grow flow matrix *F*, and an estimation of the error response matrix of the true gross flows matrix $\hat{C}$, an estimation of the matrix of the true flow matrix, F^{*}, can be readily derived from equation 7:

$$\mathrm{Vec}\hat{\mathrm{F}}^{*}_{(t-1,t)} = \hat{\mathrm{C}}^{-1}_{(t-1,t)} \mathrm{VecF}_{(t-1,t)} . \qquad 8)$$

Denoting, as before, the error response matrix at time *t* by *K(t)*, we have that, under the *ICE* assumption, the following holds:

$$C_{ij,kl(t-1,t)} = K_{ik(t-1)} K_{jl(t)} ,$$

which in matrix notation is equivalent to:

$$C_{(t-1,t)} = K_{t} \otimes K_{t-1} .$$

Thus under the assumption of independent classification errors, the probability of observing a transition from state *i* to state *j* in the time interval (*t-1*) to *t*, when the *true* flow is (*k*,*l*), is the product of the probability of observing an individual in *i* in *t*, when his *true* state is *k*, and the probability of observing an individual in *j* in (*t-1*) when his *true* state is *l*. In this manner, we are therefore, implicitly assuming that the error rate in any period *t* depends only on the true state at *t*, and on neither the true or observed states in time (*t-1*). Whilst it can be argued that misclassifications are likely to be serially correlated over time, in that it is not difficult to imagine that individuals who respond erroneously in one period, are likely to do so again in the next, the *ICE* assumption could still be valid if one considers that misclassification errors arise from a number of very different sources (miscoding, respondent error, interviewer error etc.), which by their very nature they are likely to be serially independent.

In the case of independent classification errors, estimates of $\hat{C}$ can be obtained directly from $\hat{K}$; with the matrix $\hat{K}$ being estimated using available data from the interview-reinterview surveys, according to the methodology outlined in the previous section. The relationship between the *true unobserved* flows and the observed gross flows is then given by:

$$Vec\hat{F}^{*}_{(t-1,t)} = \left[\hat{K}^{-1}_{(t)} \otimes \hat{K}^{-1}_{(t-1)}\right] VecF_{(t-1,t)}, \qquad 9)$$

which is equivalent to equation 1 in section 2.2.

In table 8 we present estimates of the true Spanish labour market flows, F^{*}, computed using the estimations of K obtained in section 2.2 and equation (9).

Table 8:
Observed and Adjusted Flows in *EPA* across the First and Second Quarter 1994
(number of people:unweighted)

Second Quarter 1994			
First Quarter	Employed	Unem-ployed	Not in the Labour Force
Total Unadjusted Flows			
Employed	40,908	1535	771
Unemployed	1,902	10,748	1,029
Not in the Labour Force	745	1,328	35,238
Total Adjusted Flows			
Employed	43,624	439	-474
Unemployed	878	13,997	-745
Not in the Labour Force	-554	-384	37,424
Unadjusted Men			
Employed	27,558	1033	354
Unemployed	1264	5395	391
Not in the Labour Force	286	459	9,794
Adjusted Men			
Employed	28,864	425	-136
Unemployed	692	6,651	-153
Not in the Labour Force	-226	-70	10,486
Unadjusted Women			
Employed	13,350	502	417
Unemployed	638	5,353	638
Not in the Labour Force	459	869	25,444
Adjusted Women			
Employed	14,760	45	-346
Unemployed	216	7,433	-669
Not in the Labour Force	-331	-383	26,945

As it is evident, the adjustment procedure, which has been widely used in the US to correct the observed flows, would in the case of Spain appear to be highly problematic. In that a number of the off-diagonal terms in the adjusted aggregate transition matrix are found to be negative. In an attempt to investigate the extent to which these negative flows are in fact caused by the failure to account for significant differences in the incidence of misclassification across the sexes, the analysis was replicated for both men and women separately. The continued attainment of a number of negative off-diagonal elements in the disaggregated adjusted transition matrices would however, appear to raise serious doubts as to the appropriateness of the general adjustment techniques developed in the US, for the Spanish labour force data.

One possible explanation for this over-adjustment is that the assumption of independent misclassification errors underlying the adjustment procedure of equation (9) could in the case of Spain be rejected by the data. Although it has been argued that the assumption of *ICE* can be maintained on the grounds that misclassification errors are themselves the result of very different causes, if the classification errors were largely due to respondent error (in other words if the other possible sources of misclassification, such as miscoding or interviewer error were of relatively minor importance) then it would be more reasonable to assume that the error structure of the underlying data is likely to be a function of true transitions. Should this be the case, the adjustment procedure will need to be modified in order to allow for the presence of serially autocorrelated errors. In other words, equation (8) should be estimated instead of equation (9). The estimation of matrix $\hat{C}$ does however, require access to longitudinal data of matched individuals covered by both the original and reinterview surveys, in order to correctly identify the nature of the stochastic process generating these misclassification errors. Unfortunately, such additional disaggregated data from the Spanish reinterview survey is not currently made available to the public by *INE*.

Finally, it is worth noting that allowing for the presence of a particular form of serially correlated misclassification errors did not, in the case of Poterba and Summers (1986) resolve the problem of the negative off-diagonal elements obtained in the adjusted gross flow matrix for prime age males. Whilst this failure could still be a consequence of an inadequately specified structure of the error process, it is also important to emphasise the fact that in some cases, the attainment of negative adjusted may not necessarily be a consequence of the failure of the *ICE* assumption. This argument can more easily illustrated by considering the following simplistic example in which there are only two labour market states: E (employment) and U (non-employment). If we also assume that the initial state is observed without error, then it can be easily shown (see appendix 3) that in order for the off-diagonal flows to be non-negative one requires that the ratio of movers to stayers is greater than the ratio of errors to non-errors.

4. Conclusions:

This paper evaluates the appropriateness of alternative strategies, the retrospective and matched files approach, frequently used in the literature to measure labour force dynamics. The evidence presented clearly illustrates that the implementation of these conventional methods to unadjusted national labour force survey data are likely to result in considerable over-estimation of the extent of labour market dynamics: tending to result in an increase in the number of transitions and a reduction in the number of continual states. The ramifications of these finding for empirical studies of both labour market transitions and behaviour models of duration would therefore appear somewhat problematic: since failure to allow for the error structure of the underlying data is likely to result in seriously biased results.

More specifically, the survey validation work of section 2 illustrates that the combination of both recall errors and heterogeneous survey design results in the retrospective information of the national labour force surveys being subject to considerable error. This is ultimately reflected in the generation of a not insignificant number of spurious transitions; particularly between the stock unemployment and out of the labour force. Thus producing a considerably distorted image portrayed of the extent of labour market turnover. Whilst the identification of labour market transitions using the matched files approach should in theory, produce a considerably more accurate measure of the underlying dynamics of labour market, the results of section 2.2, which are consistent with those found in the US, clearly indicate that measures of labour market dynamics based on the use of unadjusted matched labour force data is also, due to the problems of sample attrition, but more importantly, the presence of significant errors in classification, subject to considerable error. If unaccounted for, such errors once again result in considerable biases being introduced into the gross flow data.

Whilst in principle estimates of the error probability matrix based on information regarding the quality of the underlining data available from quality control procedures, such as a reinterview surveys, can be used to adjust the observed gross flows for misclassification error, the work of section 3 suggests that techniques currently available for the adjustment of the observed gross flows data would not seem to be universally appropriate. In the Spanish case, for example, they result in an over-adjustment of the gross flows, particularly between the ambiguous states of unemployment and not in the labour force, ultimately resulting in these cases in negative adjusted flows. It has been suggested that one possible reason for this over-adjustment could be that the assumption of independent error classification underlying these adjustment techniques may not in fact be valid. Further work is therefore required in order to acquire some knowledge as to the nature of the stochastic process generating the error responses. This requires longitudinal data in which respondents of both interview and reinterview surveys are matched across surveys. Unfortunately, such data is not released for public use at this moment in time.

References:

Abowd, J.M. and Zellner, A. (1985): "Estimating Gross Labor-Force Flows", *Journal of Business and Economic Statistics*, 3(3), pp. 254-83.

Agresti, A. (1990): "The Analysis of Categorical Data", *John Wiley & Sons*.

Akerlof, G. and Main, B. (1980): "Unemployment Spells and Unemployment Experience", *American Economic Review*, 70, pp. 885-893.

Akerlof, G. and Yellen, J. (1986): "Unemployment Through the Filter of Memory", *Quarterly Journal of Economics*, pp. 745-773.

Alba, A. (1996a): “Employment transition of young workers in Spain”, Universidad Carlos III, working paper 96.

Alba, A. (1996b): “Explaining the transition out of unemployment in Spain: The effect on unemployment insurance” Universidad Carlos III, WP 71.

Alba, A. (1997a): “Labour force participation and transition of older workers in Spain”. Universidad Carlos III, Madrid, WP 39.

Alba, A. (1997 b): “How temporary is temporary employment in Spain”, Universidad Carlos III, WP 14.

Antolìn, P. (1994): "Unemployment Flows and Vacancies in Spain". Instituto Valenciano de Investigaciones Económicas (IVIE), WP-EC 94-05.

Antolìn, P.(1995): "Transition Probabilities to Employment and Non-Participation", IVIE. WP-EC 95-20.

Antolìn, P. (1996). “Gross Worker Flows: How Does the Spanish Evidence Fit the Stylized Facts? the Flows of the Unemployed”; Centre for Economic Policy Research. Discussion Paper No. 1398, May 1996.

Arulampalam, W. and Stewart, M. (1995): “Determinants of Individual Unemployment Duration in an era of High Unemployment”, *The Economic Journal*, 105, pp. 321-332.

Blanchard, O.J. and Diamond, P.A. (1990): “ The Cyclical Behavior of Gross Flows of Workers in the United States”, *Brookings papers on Economic Activity*, 2, pp. 85-155.

Blanchard, O.J. and Diamond, P.A. (1992): “The Flow Approach to Labour Markets”, *American Economic Review*, Papers and Proceedings of the American Economic Associations Congress, May 1992.

Bover, O. et al. (1996) “Unemployment Duration, Benefit Duration, and the Business Cycle. Banco de España WP 57.

Burda, M. and Wyplosz, C. (1990): “Gross Labor Market Flows in Europe: some stylised facts”, CEPR Discussion Paper 439

Burda, M. and Wyplosz, C. (1993): “Gross Worker and Job Flows in Europe”, CEPR discussion paper 868.

Chua, T.C. & Fuller, W. A (1987): " A model for Multinomial Response Errors Applied to Labor Flows". Journal of the American Statistical Association, Vol. 82, No. 397.

Clark, K.B., and Summers, L. (1979): " Labor Market Dynamics and Unemployment Spells: A reconsideration, *Brookings Papers on Economic Activity*, 2, pp.13-60.

Feldstein, M.S. (1973?): "Lowering the Permanent Rate of Unemployment". Washington D.C.:US Government Printing Office.

Fuller, W. A. & Chua, T.C. (1985): "Gross Change Estimation in the Presence of Response Error", in Proceedings of the Conference on Gross Flows in Labor Force Statistics, Washington, D.C.: US Bureau of the Census and US Bureau of Labor Statistics, pp.65-77.

Garcia Serrano, C. (1996): El mercado de trabajo español: una perspectiva desde los flujos.

Gomez-Salvador, R. and Dolado, J.J. (1995): “Creacion de Empleo en España: Un Analisis Descriptivo con Datos de la CBBE”, Banco de España WP .

Hall, R.E. (1972: "Turnover in the Labor Force", *Brookings Papers on Economic Activity*, 3, pp.709-756.

Instituto Nacional De Estadistica (1995): Evaluación de la calidad de los datos de la Encuesta de Población Activa

Instituto Nacional De Estadistica (1996): Estadistica de Flujos - Metodologia.

Katz, L. and Meyer, B. (1990): “The impact of the potential duration of unemployment benefits on the duration of unemployment” Journal of Public Economics.

Konnings, J. (1993): “Job Creation and Job Destruction in the UK Manufacturing Industries”, Centre for Economic Performance Discussion Paper, No. 176.

Levine, P.B. (1990): Contemporary vs. Retrospective Unemployment: Through the filter of memory or the muddle of the current population survey. Working Paper, No. 276, Industrial Relations Section, Department of Economics, Princeton University.

Mathiowetz, N.A. & Duncan, G. J. (1988): "Out of Work, Out of Mind: Response Errors in Retrospective Reports of Unemployment." *Journal of business and Economic Statistics*, pp. 21-229.

Meyer, B.D (1988): "Classification-Error Models and Labor-Market Dynamics", *Journal of Business and Economic Statistics*, July 1988, Vol. 6, No.3.

Narendranathan, W. and Stewart, M. (1993): “How Does the Benefit Effect Vary as the Unemployment Spell Lengthens?" *Journal of Applied Econometrics (1993)*.

Moreno, G. and Toharia, L. (1998): "Mujer, Familia y Movilidad En Espana", *Papeles De Economia Espanola*, No. 77.

Peracchi, F. and Welch, F. (1995): "Trends in Labor Force Transitions of Older Men and Women". *Journal of Labour Economics*.

Peracchi, F. and Welch, F. (1995): How representative are matched cross-sections? evidence from the CPS. *Journal of Econometrics*.

Pitt, A. (1988) "Matching adjacent years of the Current Population Survey", Mimeo. Unicon Research Corporation, Santa Monica, CA.

Poterba, J.M. and Summers, L.H. (1983): "Survey Response Variation in the Current Population Survey", NBER Working paper, no. 1109.

Poterba, J.M. and Summers, L.H. (1986):"Reporting Errors and Labor Market Dynamics", *Econometrica*, 54 (6), pp. 1319-38.

Poterba, J.M. and Summers, L.H. (1995): "Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model With Errors in Classification" *The Review of Economic and Statistics*.

Stevens, L.K, & Register, C.A. & Grimes, P.W. (1987): "The Labour Force Attachment of Discouraged Workers", *Work and Occupations*, February , pp. 3-21.

Trivellato, U. & Torelli, N. (1989): "Analysis of Labour Force Dynamics from Rotating Panel Survey Data", I.S.I-47th Session, Paris, August 29-September 8, 1989.

Ureta, M. (1987): "Are you Really Out of the Labour Force? or the Effects of Recoding on Unemployment Statistics.", unpublished paper, Department of Economics, The Claremont Graduate School, May 1987.

APPENDIX 1

The Epa Retrospective Supplement

The *EPA* retrospective supplement, undertaken in the second quarter of each year as a supplement to the regular labour force survey, contains the following restricted set of questions regarding the respondent's situation one year prior to the current survey:

1. Which was your city of residence exactly one year ago?

    - The same as my present one
    - Different

2. Where were you living one year ago?

    - If in Spain, indicate which city.
    - If in a foreign country, indicate country.

3. In which of the following situations were you in exactly one year ago?

    - Working
    - Searching
    - Available, but not searching
    - Military service
    - Studying
    - Other situations

4. What was your professional situation one year ago?

    - Employer

With a franchise
Other

- Member of a co-operative
- Working as a family member in a family owned company
- Public sector employee
- Private sector employee
- Other

5. What was the type of activity of the establishment in which you were working one year ago?

APPENDIX 2.

The Estimation Of The Parameters Of The Model

Let:

$$Y = \left(\hat{A}_{11}, \hat{A}_{12}, \hat{A}_{13}, \hat{A}_{21,} \hat{A}_{22}, \hat{A}_{23}, \hat{A}_{31,} \hat{A}_{32} \right),$$

$$Y = f(y) + e,$$

where Y is an (8*1) vector whose elements are the sample proportions of individuals that are placed in category i on the first interview and in category j on the reinterview. In other words, the elements of vector Y are simply the first 8 elements of the interview-reinterview matrix. $f(y)$ is the vector of the expected values of the sample proportions and e is the vector of deviations of the observed proportions from the expected proportions. The vector of parameters to be estimated is:

$$y = \left(a_{21}, a_{31}, a_{32}, J_{12}, J_{13}, J_{32}, P_1, P_2 \right).$$

Assuming that (n, A_{ij}) follows a multinomial distribution, the covariance matrix of e will be given by:

$$V = n^{-1} \left\{ diag\left[f(y) \right] - f(y) f(y)' \right\},$$

where n is the sample size.

The vector y is estimated using the Gauss-Newton procedure, in which $\hat{V}$ is included as the estimated covariance matrix of e. Where:

$$\hat{V} = n^{-1}\left\{diag[Y] - YY^{'}\right\}.$$

Note however, that the multivariate Central Limit Theorem implies that (Agresti p. 424)

$$\sqrt{n}(\hat{\psi} - \psi) \xrightarrow[L]{} \{[F(\psi)'V^{-1}F(\psi)]\}^{-1},$$

this allows us therefore to obtain the standard errors of the estimates.

APPENDIX 3

Under Which Conditions Is A Corrected Flow Matrix F^* With Negative Terms Obtained?

Consider a somewhat simplistic case in which the following simplifying assumptions are made:

(A1) The initial state is observed without error; and

(A2) There are only two states: E (employment) and U (unemployment and not in the labour force).

In this particular setting the error response matrix K can be expressed as:

$$K = \begin{bmatrix} (1-e) & d \\ e & (1-d) \end{bmatrix} ,$$

where: ε = Probability (observed in state U|truly in state E) and δ = Probability (observed in state E| true state is U). Equation (9) in the main text in this framework reduces to:

$$Vec\hat{F}^{*}_{(t-1,t)} = \hat{K}^{-1}_{(t)} VecF_{(t-1,t)} .$$

This in itself can be expressed as:

$$F^{*} = \frac{1}{(1-e-d)} \begin{bmatrix} (1-d)EE - dEU & -e.EE + (1-e)EU \\ (1-d)UE - dUU & -e.UE + (1-e)UU \end{bmatrix} = \begin{bmatrix} EE* & EU* \\ UE* & UU* \end{bmatrix}.$$

At this point, the following additional "very mild" assumptions are required:

(A3) e < 0.5 and d < 0.5. [This guarantees that |K| is always positive]

(A4) The number of observed stayers in state *i* > the number of observed movers from *i*. In other words:

EU < EE and UE < UU .

We are now in a position to analyse the sign of the elements in the *F** matrix. Given the assumptions A3 and A4:

1) EE* is always positive:

$$EE^{*} = (1-d)EE - dEU = (1-d)\left[EE - EU\right] + (1-2d)EU \ ,$$

2) but EU* can be negative;

$$EU^{*} = -e.EE + (1-e).EU = e.\left[EU - EE\right] + EU\left[1-2e\right] =$$

$$= EU\left[\quad + 1 - \quad\right] - \quad .EE =$$

$$= \ EU\left(1 - e - EE.e\right).$$

This final expression will be positive *iff* $\dfrac{EU}{EE} > \dfrac{e}{(1-e)}$.

In short, the off diagonal term EU* is positive if and only if the ratio of movers over stayers, (EU/EE), is greater than the ratio of errors over non errors $e/(1-e)$. Note that the same type of analysis follows for the other off diagonal element UE*.