IDENTIFYING LABOUR MARKET DYNAMICS USING LABOUR FORCE SURVEY DATA

Concha Artola
and Una-Louise Bell

IDENTIFYING
LABOUR MARKET
DYNAMICS USING
LABOUR FORCE
SURVEY DATA

Concha Artola (*)
and Una-Louise Bell (**)

The Banco de España disseminates its principal reports
and economic indicators via INTERNET and INFOVÍA.PLUS at:
http://www.bde.es

ABSTRACT:

Despite the increasing focus on labour market dynamics, little is known about either the quality of the underlying data or the appropriateness of the standard methodologies frequently used to analyse labour market dynamics in Europe (the retrospective and matched files approaches). This paper is a first attempt to fill these gaps. In section one we highlight the problems associated with the use of the alternative methodologies. In section two, we provide evidence as to the size and nature of the errors arising from the use of these two approaches. The results of our survey validation work suggest that, due to problems of recall error and heterogeneous survey design, the retrospective approach tends to result in a considerable number of spurious transitions being recorded. Whilst the use of quasi-longitudinal LFS data should in theory overcome such problems, issues of sample attrition and, more importantly, misclassification error tend to result once again in significant over-reporting of labour market transitions. For example, it is estimated that in the first quarter of 1994, 12.5% of the Spanish unemployed were incorrectly classified into a different labour market state. These results, whilst preliminary, emphasise the need to allow for the underlying error structure of LFS data in empirical studies of labour market dynamics, and draw into question the findings of existing work which fails to do so. We conclude by providing an example of one of the ways in which the error probability rates estimated in section 2.2 can be used to adjust the underlying data, so as to correct for potential errors arising from the use of the matched files approach. The results obtained suggest, however, that techniques developed in the US to correct the observed labour market flows for misclassification errors may not be readily applicable to Spanish data.

IDENTIFYING LABOUR MARKET DYNAMICS USING LABOUR FORCE SURVEY DATA

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 has been developing along two distinct paths: i) an analysis of gross labour market flows[1]; and ii) a microeconometric analysis of the determinants of individual labour market transitions[2]. Many of these studies have been conducted using the national Labour Force Survey (LFS) data, researchers either adopting: 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 referred to here as the *matched files approach*), in order to identify labour market transitions.

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 (1988), and Chua and Fuller (1987))[3], little attention has to our knowledge been paid to the reliability/appropriateness of LFS data for the identification of labour market transitions in Europe. Yet, the consequences of the use of raw LFS data in an analysis of labour market dynamics can be of paramount importance: the failure to correct for erroneous data caused for example, by misclassification, will in studies based on individual records matched across consecutive waves of the LFS, typically result in a considerable over-estimation of 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 between the two periods. The implications of this for empirical work are self-evident: If the objective of the study is either duration analysis or the

---

[1] See for example, Blanchard and Diamond (1990, 1992), Davis, Haltiwanger and Schuh (1996), Burda and Wyplosz (1990, 1993), Konnings (1993), Antolin (1994, 1995, 1996), Gomez and Dolado (1995).
[2] See Katz-Meyer (1990), Clark-Summers (1979), Narendranathan and Stewart. (1993), Peracchi and Welch (1994), Arulampalan and Stewart (1995), Alba (1996a, 1996b, 1997a, 1997b), Bover et al. (1996) and Portugal and Addison (1997).
[3] 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, strongly rejected this idea. They argued instead that unemployment durations were actually much longer than those implied by simply looking at the raw labour force data. Their reasoning was that the use of national labour force surveys for the measurement of labour market turnover was in itself problematic, in that the use of such surveys, which are originally designed to provide cross-sectional images of the economy at specific points in time, to create quasi-longitudinal databanks tends to result in a considerable amount of noise, due for example, to misclassification errors and mismatching of individual records over time, entering into the longitudinal data. This leads to a distorted picture of the degree of dynamism of the market.

measurement of labour market dynamics, then research using unadjusted data is likely to be subject to considerable error[4].

In this paper therefore, we assess the magnitude and ramifications of the use of LFS data in studies focusing on the dynamic nature of the labour markets in countries which differ considerably from the US. The objective is 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, and the nature, 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 LFS, in conjunction with previously under-utilised information obtained from Spanish re-interview surveys. Finally, we conclude the paper by providing an illustrative example of one of the ways in which the error probability rates estimated in section 2.2 can be used to adjust the underlying data, so as to correct for potential errors arising from the use of the matched files approach. The results obtained suggest however, that the techniques used in the US to correct observed flow data for the presence of misclassification error would appear to be inappropriate for Spain, since in the Spanish case they result in negative gross flows across a number of labour market states[5]. Whilst this over adjustment obviously tends to question the validity of the underlying assumptions of such methods, it is argued that access to more disaggregated re-interview survey data is required before the much needed additional research can be carried out in this area.

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

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

---

[4]Poterba and Summers (1986) for example, in their analysis of the effects of unemployment benefits on unemployment transitions, conclude that "the sensitivity of spell duration to UI (unemployment insurance) 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".

[5]Note however, that Poterba and Summers (1986) also obtain negative flows for prime age males in their (1986) study of the correction of gross labour market flows disaggregated according to both age and sex.

1.1. The Retrospective Information Approach

National labour force surveys typically contain a significant amount of retrospective information regarding an individual's previous labour market activity. Such retrospective information tends to be concentrated into two distinct types of questions: those relating to the duration of specific events, be it employment, unemployment or search related duration questions; and those referring to an individual's previous labour market status[6]. In view of their nature, the latter tend to suffer from problems arising from both heterogeneous sample design and recall error bias, whilst the former tend to be particularly susceptible to various facets of the recall error problem[7].

Recall Error and Previous Labour Market Status: Whilst it may be reasonable to expect a respondent to provide an accurate answer to his *current* labour market status at the time the survey is being carried out, it is more questionable whether he will be able to accurately recall his labour market status at some specified time in the past[8]. This recall error problem is further accentuated 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. Moreover, when it comes to having to recall the previous labour market activity of other household members, it is highly likely that the respondent will only be able to accurately recall events which are "salient" for the household unit, such as the loss of employment of a father, or main income earner (individuals who are not normally subject to periods of unemployment)[9]. It would appear reasonable to assume that this particular aspect of recall error is likely to be a more acute problem in labour markets, such as the Spanish one, where a large percentage of individuals are currently employed under temporary contracts of very short duration, and thus subject to a greater degree of turnover.

Recall Error and Duration Data: Recall Errors are also an important problem in the other type of retrospective questions, i.e. those dealing with spell duration. Poterba and

---

[6] The US Labour Force Survey for example, 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

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

[8] 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 is 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.

[9] Evidence of this being found in the comparison of retrospective and contemporaneous measures of unemployment carried out by Akerlof and Yellen (1986). Their results suggest that the retrospective measure of unemployment is larger for those groups for which unemployment is to be considered as a more serious situation, i.e. prime age males and both men and women over 55 .

Summers (1983) and Bowers and Horvarth (1984) analyse unemployment spells as measured in *CPS* data, showing that errors in reported unemployment duration are substantial (3 out of every 4 survey respondents reporting inconsistent unemployment duration over consecutive surveys) with these inaccuracies tending to be an increasing function of unemployment duration itself, i.e. the long-term unemployed tend to understate the duration of their unemployment spell, while the opposite is true for people in short unemployment spells.

Heterogeneous Survey Design: A further problem researchers often have to overcome, when using retrospective questions either to create retrospective measures of the labour market stock or 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[10]. It is not difficult to envisage how such differences in the assignment of labour market status could lead to a significant number of misclassifications, particularly in the case of those individuals with relatively weak levels of labour force attachment. The fact, for example, that the retrospective survey does not involve any degree of detailed questioning with respect to job search activity, would tend to imply that a discouraged worker (who wants to work, but is not actively searching), will tend to be contemporaneously classified as not in the labour force (NILF), but retrospectively as unemployed. In a similar vein, Ureta (1987) illustrates that someone with minimal search effort in the reference week, tends to be classified contemporaneously as being unemployed, but retrospectively as NILF. One reason is that there is relatively more pressure, due for example to issues of benefit eligibility, to answer positively to job search requirement activity questions in the contemporary survey than in the retrospective one.

The implications of such issues for the accurate identification of labour market transitions is evident from the results of Levine (1990), who finds that 35-50% of individuals contemporaneously classified as being unemployed in t fail to report retrospectively, when interviewed in $t+1$, periods of unemployment in t[11]. Whilst some of this difference is undoubtedly due to recall errors, a large proportion of this difference is to be attributed to misclassifications arising from differences in both survey design and the classification procedures used in the contemporaneous and retrospective definitions of unemployment. More specifically, Levine argues that the retrospective measure of unemployment in *CPS* requires, due to the characteristics of the survey, greater labour force attachment than is required to be classified contemporaneously as unemployed.

---

[10]Labour market status, for example, is typically self-reported by individuals in the retrospective part of the questionnaire, whereas current labour market status is normally assigned by the interviewer on the basis of the interviewee's response to a series of questions regarding his contemporaneous labour market activity.

[11]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.

1.2. The Matched Files Approach.

The matched files approach exploits the rotating panel structure of the national labour force surveys, whereby a selected household remains in the sample for a number of consecutive quarters before being replaced by a newly surveyed household[12]. This rotating characteristic allows researchers 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.

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[13]. 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, namely sample attrition and misclassification errors, 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, as it belongs to an ongoing rotation group, does not respond to the survey. This failure to respond may be due either to migration (a member of the household or the entire household unit itself leaving the sampled address)[14] 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 month's respondents cannot be located in the previous month's survey. Thus in the US, on average

---

[12]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 *CPS* 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's records for three quarters of the respondents, and are able to match files for 50% of the current month's survey with the survey in the same month of the following year; and 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, so that 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.

[13]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 gross labour market flows.

[14]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.

15% of all those individuals who in theory should be matched across consecutive months cannot be matched at all. The available Spanish evidence (INE (1996)) also suggests that the fraction of attriters in Spain is both 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.

It is important to note however, that if this attrition process is not random. In other words, if the behaviour of those in the outgoing rotation segment 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[15]. Available evidence for Spain also suggests that sample attrition in the Spanish LFS data is a non-random process. The results of the Moreno & Toharia (1998) study indicate that attrition is more frequent among women, individuals aged between 25 and 34 years, those with higher levels of education and household heads; and it is less frequent for unemployed individuals without previous work experience. Naturally, the bias produced by non-random attrition will 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 errors, miscoding or interviewer errors [16], are however the major pitfall that researchers have to overcome when using the matched files approach to obtain a measure of labour market 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

---

[15]It should be noted however, that the study by Pitts (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.

[16]In the majority of national LFSs the unique identification code is only allocated at the household level, since these surveys are by their very nature household based surveys. Researchers are therefore forced 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 also occur, therefore, when incorrect matches are made.

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. However, as the majority of individuals remain in the same labour market state (in other words the diagonal cell counts far outweigh those of the off-diagonals), one will, 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-diagonals to the diagonals. Misclassification errors which are independent across individual surveys will therefore tend to increase the number of reported changes. It is only when misclassification errors are perfectly positively correlated across time that the calculated gross flows will be unbiased.

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 therefore, the Spanish labour force survey, the *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, Antolín (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. The choice of Spain was, however, first and foremost dictated by both data availability and accessibility, and preliminary investigations indicating that the quality control information disseminated by INE readily lent itself to the analysis in hand[17]. Spain is, however, a somewhat interesting case study, given both the changes in employment legislation that have occurred over the past 10-15 years, and the continual emergence of an impressive array of statistics and empirical studies, which if to be believed, support/document the transformation of the Spanish labour market from a typical sclerotic one, to one which exhibits levels of turnover not that dissimilar to those of the US.

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

---

[17] Information on the quality control process implemented in other European LFS's would appear to be somewhat more difficult to obtain. In Italy for example, such information is only available for internal use within the national statistical office, ISTAT. Unfortunately, this lack of accessibility also prevents us from being able to ascertain the extent to which the evaluation procedures adopted in this section can be readily applied elsewhere in the EU.

market status 12 months prior to the current survey[18]. It is this retrospective information which has been used to identify transitions between labour market states across two different time periods. The accurate identification of a transition is, however, somewhat hampered by the considerable differences which exist with respect to both survey design and classification procedures, between the contemporaneous and retrospective Spanish surveys. For example, in contrast to the regular *EPA* questionnaire, where responses to a much wider set of questions with respect to current labour market activity are used 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 part of 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. 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[19]. Moreover, the Spanish 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 information sources is 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. There are i) the aforementioned retrospective supplement to the labour force questionnaire, and ii) data obtained by matching individual records across various quarters of the labour force survey

This type of validation procedure is feasible, since the longitudinal characteristics of the *EPA* ensure that: i) in the second quarter of each year the *EPA*, contains information on the labour market status 12 months prior to the current interview 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

---

[18]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.

[19]The *EPA* classifies men doing military service in a special category denoted *Poblacion Contada Aparte*.

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.

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 $\Pi_0 = \sum \Pi_{ii}$, where Π_{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 Π_0 will be 1. Using the data in Table 2, a value of 0.94 is obtained for this index of agreement. This implies that, for the period second quarter 1992 - second quarter 1993, *recalled* labour market status would in fact appear to be a reasonable indicator of true labour market status[20]. A better measurement of agreement is however, the *kappa* index 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:

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

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

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 (as reflected by Π_e).

---

[20]*INE* (1996) states that the proportion of classification agreement is high, so that the results obtained from retrospective questions can be considered reliable, and that a comparison of the marginal classsifications, which are extremely similar, strongly supports this claim.

In the denominator, instead of $\prod_e$, we have 1, which is the maximum possible value of $\prod_o$. Thus the *kappa* index 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. $(1-\prod_e)$. 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[21].

Whilst 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 distributions of observations across the three states would seem to suggest that *recalled* LMS is actually a relatively good proxy of the *true* LMS at any point in time, the fact that 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 is often overlooked. For, as is evident from the data presented in Table 2, there are in fact in a number of cases 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 become 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 clearly evident that recall errors are, particularly in the case of transitions from unemployment, relatively important. For, whilst the observed rate of labour force withdrawal from unemployment from panel A, is equal to 8.5%, the data in panel B indicates 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 error rates. The observed probability of an individual moving from unemployment to employment between 1992 and 1993 is 27% according to the retrospective survey. Note however, that according to panel B, almost 10% of the people initially classified as unemployed were in fact working.

---

[21]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).

The use of the retrospective part of the labour force survey for both the construction of gross labour market flows and for analysing labour market dynamics would, in the 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, researchers also have to overcome the problem of missing data, when using the retrospective information to construct a measure of labour market dynamics. 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. If, as previously discussed, these observations are not missing at random, a further bias will be introduced into the gross flows data.

2.2. The Matched Files Approach:

As previously noted 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 household eventually drops out of 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[22].

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*[23].

In common with the re-interview system of the *CPS*, the Spanish re-interview 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 re-interviewing a sub-sample of the original *EPA* respondents approximately 15 days after the original interview takes place[24]. Respondents of the re-interview survey are once again asked to respond to the same set of questions used in the

---

[22]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 representativeness of the sample, one should limit oneself to five quarters only.

[23]Here we do not therefore, address the problem of sample attrition briefly discussed in section 1.2. The reasons for this are : first and foremost, in contrast to the case of misclassification errors, some evidence as to the size and nature of sample attrition already exists for Spain; secondly, it is widely accepted that in these types of rotating panel surveys misclassification is a significantly more important problem than sample attrition.

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

regular labour force survey, although the responses refer to labour force activity in the original survey period, and not to the re-interview period. The re-interview 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 re-interview procedure does not involve any *reconciliation process* in order to identify the correct response when inconsistencies occur between the original and re-interview surveys[25]. Subsequently, one is unable to assume that the Spanish re-interview process correctly identifies the individual's true labour market state. This inability, as discussed in due course, does tend to complicate somewhat attempts to calculate the magnitude of errors present in the original survey. As the re-interview is carried out by more experienced interviewers, one can however, assume that the data are of superior quality to those collected in the original survey and thus that the extent of error is smaller in the follow-up survey.

Information from the re-interview survey enables one, therefore, to cross check 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[26]. Cross-tabulating the distribution of recorded labour market status for the subset of the individuals who respond to both the *EPA* interview and the re-interview surveys for example, provides one with an idea of the incidence and magnitude of the misclassification errors present in the original survey[27]. In the absence of misclassification errors, the respondents' labour market status in the original *EPA* and the re-interview survey should be identical: i.e. the off-diagonal elements of the interview-re-interview table, which correspond to inconsistent responses, should be equal to zero.

In the first instance, it is evident, that the use of raw LFS data does not pose a significant problem for the estimation of populations stocks; as illustrated in Panel 1 of

---

[25]In the US, the re-interview procedure is divided into 2 sub-samples: the reconciled and unreconciled components. In the reconciled case, which involves approximately 75% of the re-interview survey, interviewers are provided with the original interview. Thus after having conducted the second interview, the re-interviewer compares the responses given in both the original and re-interview surveys. Any discrepancies uncovered are then reconciled *with the respondents* before they leave 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.

[26]It is important to note however, that the percentage of *true* misclassifications is likely to be higher than that detected by the re-interview 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.

[27]Since exactly the same questionnaire is used in both the original and re-interview surveys, it is in fact in principle possible to check for the consistency of a very large number of variables. *INE* indeed publishes 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, the value of this index in 1994 ranged 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 error). Note, that the values of the Global index of agreement are in fact very similar from one year to another.

Table 4, the differences in the marginal distributions across the two surveys are minimal. In contrast, the off-diagonal values of Panel 1, 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 evident from the data presented in Panel 2, in which a comparison of individual responses for labour market status in the original and re-interview surveys clearly illustrates the magnitude of the impact of misclassification errors for the accurate measurement of labour market transitions. These figures are particularly alarming, when one compares them to the observed quarterly transitions between two consecutive *EPA*, since the transitions in Panel 2, all of which are spurious, are very large in size relative to the observed quarterly transitions reported in Panel 3, e.g. 7.5% of those observed as unemployed in the first quarter are recorded as having withdrawn from the labour force in the second quarter. According to the re-interview survey however, 15% of those classified as unemployed in the original interview were later reclassified as not being in the labour force in the re-interview.

It is interesting to note, however, that the magnitude of inconsistencies with respect to labour market status differs 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 associated with classifications between unemployment and not in the labour force is particularly high for women, a group known to have a lower level of labour force attachment.

These findings clearly underline the need to control for the underlying error structure of the LFS data, before the "raw" flow data obtained from matching consecutive waves of LFS, can be used in any analysis of labour force dynamics. Quality control information of the type obtained from the re-interview survey procedure has in fact been used in the US in order to obtain estimates of the probability of a misclassification occurring in the LFS data[28]. Underlying the general approach taken in these studies is the key assumption that misclassification errors are independent across time. In other words the probability of an individual being misclassified at time t is independent of the probability of being missclassified at time $(t - 1)$. To clearly illustrate the implications of the ICE assumption for the estimation of misclassification error rates, we consider in the first instance, the more general case in which no particular assumption is made with respect to the stochastic nature of the response errors.

Let $F(t - 1, t)$ denote the observed gross flows matrix between $(t-1)$ and t and $F^*(t - 1, t)$ the true unobserved gross flow matrix.

---

[28]See for example, Abowd & Zellner (1985), Chua & Fuller (1987), Fuller & Chua (1985) and Poterba & Summers (1985).

There are two types of error probabilities; the probability of incorrectly observing the state at time t, and the probability of wrongly observing the flow between $(t-1)$ and t. let $K(t)$ denote the *(3*3) error response matrix of the true stocks,* whose *ijth* entry is given by

$$K_{ij}(t) = P(\text{observed state in } t \text{ is } i \mid \text{true state in } t \text{ is } j),$$

whilst the *(9*9) error response matrix of the true gross flows* is given by C, whose *ij,kl-th* element is:

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

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

In this framework, the relationship between the observed and true labour market flows is given by:

$$VecF(t-1,t) = C(t-1,t) \cdot VecF^{*}(t-1,t), \qquad\qquad 1)$$

where: $VecF(t-1,t)$ denotes the (9*1) vector obtained from stacking the rows of matrix $F(t-1,t)$ one on top of the other; the same applies to $F^{*}(t-1,t)$.

Given the observed gross flows matrix F, and provided an estimation $\hat{C}$ of the error response matrix of the true gross flows is available, an estimation of the true flow matrix, F^{*}, can be readily derived from equation 1:

$$Vec\hat{F}^{*}(t-1,t) = \hat{C}^{-1}(t-1,t) \cdot VecF(t-1,t). \qquad\qquad 2)$$

In general, however, one does not have access to the information required for the direct estimation of C. Under the ICE assumption the informational requirements to estimate F^{*} are milder. In this case, 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. i.e:

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

---

[29] Assuming lexicographic ordering of rows and columns.

which in matrix notation is equivalent to:

$$C(t-1,t) = K(t) \otimes K(t-1) \, .$$

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 $(t-1)$. Notice that, insofar as we believe that misclassification errors arise from a number of very different sources (miscoding, respondent error, interviewer error etc.) which by their very nature are likely to be serially independent, the ICE assumption will be a valid one.[30]

Under the assumption of independent classification errors, estimates of $\hat{C}$ can therefore be obtained directly from $\hat{K}$, with the matrix $\hat{K}$ being estimated using available data from the interview-re-interview surveys, according to the methodology outlined in the following section. In such a framework the relationship between the true unobserved flows and the observed gross flows is given by:

$$Vec\hat{F}^*(t-1,t) = \left[\hat{K}^{-1}(t) \otimes \hat{K}^{-1}(t-1) \right] \cdot VecF(t-1,t) \, . \qquad 3)$$

The existing studies seeking to allow for the underlying error structure of the LFS data, differ substantially in their estimation of the probability of being misclassified, K. More specifically, in using only the information from the reconciled sample of the *CPS* re-interview survey, Abowd and Zellner (1985) are implicitly assuming in their adjustment procedure that the correct labour market status is revealed by the re-interview survey. Thus their matrix K is simply the interview-re-interview 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. For, if the procedure specified for the reconciled re-interview survey is correctly carried out, one should observe that the incidence of discrepancies in the unreconciled sub-sample is similar to the incidence of discrepancies in the reconciled one (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 re-interview surveys is much greater in the unreconciled than in the reconciled sample (before reconciliation), an anomaly which indicates that re-interviewers "cheat" by using information from the original survey to complete the reconciliation process. This suggests 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. In order to allow for such differences in the incidence of discrepancies between the reconciled and unreconciled samples, and thus overcome the

---

[30] If on the contrary, most misclassification is due to respondent errors, it may appear more reasonable to have serially correlated errors; i.e. a person who makes a mistake in *(t-1)* is more likely to make a mistake in *t*, either because he is more error-prone or because his situation is ambiguous

potential shortcoming of the Abowd-Zellner approach, in addition to the *CPS* error classification rates obtained from the reconciled sample, Poterba and Summers (1986) also use 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 re-interview survey and the incidence of errors estimated from the unreconciled sub-sample.

As *INE* does not carry out a reconciliation procedure, we are unable to assume that the re-interview process correctly identifies an individual's true labour market status. Estimates of the error response probabilities are therefore obtained using an alternative methodology 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 re-interview data, Fuller and Chua developed a parametric procedure to estimate the response matrix, K, based on the unreconciled sub-sample only[31].

Estimating the matrix K with unreconciled data.

The method proposed by Fuller and Chua to estimate matrix K relies on a set of assumptions that we underscore in the course of the following description.

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). This classification process is subject to error.

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.

Assumption 1: Unbiased Response Error (URE). The misclassification process only affects the distribution of the interior cells but not the marginal proportions. The marginal proportions are then 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. In other words, the failure of the unbiased response error assumption would 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:

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

---

[31] In the following sub-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.

Under the URE assumption the matrix of error response probabilities, K, can be estimated in the following manner. Let $\hat{A}$ be the observed interview-re-interview table (e.g. the first panel in Table 4). $\hat{A}$ 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 re-interview.

Assumption 2: The misclassification process in the interview is independent of the classification process in the re-interview.

Assumption 3: The error response probabilities are identical in the two trials. Thus in both the original interview and in the re-interview, the probability of missclasifying in state i somebody who is really in state j is given by K .

In this set-up, the *population* interview-re-interview table, the matrix A, can be expressed as:

$$A = E\{\hat{A}\} = K \left[diag\, P \right] K' . \qquad\qquad 4)$$

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 URE assumption ($KP = P$), 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 4 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 4. 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. Thus 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, the matrix A, given P, 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.

Assumption 4: Matrix K has the following parametric form:

$$K_{ij} = \left[1 - \sum_{l=1}^{3} \alpha_{ij} P_l \left(P_l + P_i \right)^{-l} \right] \delta_{ij} + \alpha_{ij} P_i \left(P_i + P_j \right)^{-l}, \quad i, j = 1,2,3 \qquad 5)$$

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

The probability of being misclassified is therefore a function of the parameters α and the marginal proportions P. The assumption of symmetry, i.e. $\alpha_{ij} = \alpha_{ji}$, 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 α_{ij} can therefore be considered to represent an index of the probability of making a misclassification error.

The intuition behind assumption 4) can be better seen in an example. The probability of being incorrectly classified in state 1, when the individual's true labour market state is 2, is given by:

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

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:

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

Taken together, these two expressions, mean that the greater the number of individuals classified in state 1, relative to state 2: (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 matrix K, as defined in 5), is a   probability matrix, in that the elements of each column add up to one.

Before proceeding to the estimation of the model itself, it is, however, 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-re-interview table (table 4, panel 1)[32]. This rejection thus implies that in our case the A matrix will not be symmetric and that equation 4 no longer holds.

To accommodate the presence of a non-symmetric interview-re-interview 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, assumption 3 is replaced by:

Assumption 3': The error response probabilities are given by matrix K in the original interview and by matrix G in the re-interview.

Once again, due to the lack of a sufficient number of free parameters one needs to assume a particular functional form for matrix G

Assumption 5: The error response matrix G in the re-interview survey has the functional form[33]:

$$G_{ij} = \vartheta_{ji} K_{ij} \qquad \text{if } A_{ij} < A_{ji} \text{ and } i \neq j$$
$$G_{ij} = K_{ij} \qquad \text{if } A_{ij} \geq A_{ji} \text{ and } i \neq j \qquad \qquad 6)$$
$$G_{ij} = 1 - \sum_{i=1}^{r} G_{ri} \qquad \text{if } i = j \, ,$$

where $\vartheta_{ij} \geq 1$.

The idea behind this formulation is simply that the symmetry of the interview-re-interview table breaks down because of a pure shift in the proportion of individuals 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 an individual classified, for example, as unemployed in the interview and as employed in the re-interview is truly either employed or unemployed. Thus it is assumed that an individual is always correctly

---

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

[33]Whilst the functional form chosen for the error probabilities of the re-interview survey, G, is not the only possible one, it does produce smaller estimates for a_{ij} than the other forms considered. One obvious criticism of such a formulation is, as the authors themselves point out, that it results in the second interview having a larger response variance than the initial one. When, as in our case, the second interview under consideration forms part of a quality control process, the appropriateness of this characteristic is somewhat questionable. In Appendix 2 the results obtained from other parameterisations of the G matrix are therefore reported.

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

$$G_{12} = \vartheta_{12} \alpha_{12} \frac{P_1}{P_1 + P_2},$$

and

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

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-re-interview table, the probability that an individual in true category j reports category i on the re-interview 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 4 becomes:

$$E\left\{\hat{A}\right\} = K\left[diag.P\right]G'. \tag{7}$$

The parameters α, ϑ, P, embedded in equation 7 are then estimated using a generalised non-linear least squares procedure (described in greater detail in Appendix 3), assuming that the proportions A_{ij} follow a multinomial sampling.

Substituting the estimated parameters reported in table 6 into equation 5, 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 of an individual who is actually unemployed, being incorrectly classified as not in the labour force is 8.1%. Somewhat surprisingly, however, the incidence of misclassification of those individuals truly unemployed, as being employed, is also significant: an estimated 4.4% of the truly unemployed are found to be classified as employed. Errors of this magnitude would tend to question the findings of empirical studies, be it behaviour models of transitions or analyses of labour market flows, since the results obtained, if based on unadjusted data, are likely to be subject to considerable bias.

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. These include most notably, 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 force 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 need to use disaggregated information on the error structure of the underlying data when working with unadjusted gross flows data.

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) papers. 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[34]. Secondly, there do 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 classified as not in the labour force. The size of this error is 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 thus tend to suggest that, in analysing labour market flows, one should be particularly careful when dealing with transitions from and to unemployment.

3. Using the Estimated Error Probabilities: an application to the accurate measurement of gross flows data.

The results of the previous section clearly underline the need to account for errors in misclassification in 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. In this section, therefore, we conclude by providing an example of one of the ways in which the error probability rates estimated in section 2.2 can be used to adjust the underlying data, so as to correct for potential errors arising from the use of the matched files approach when trying to measure gross labour market flows [35]. As previously mentioned, adjustments of this nature 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

---

[34]Note however, that the Abowd-Zellner results are in general smaller than the rest, due to the underlying assumptions adopted for the computation of their K matrix that we previously outlined.

[35]An example of the way in which the estimated error probabilities can be used to adapt behavioural models in order to allow for the underlying error structure of the data can be found in Poterba and Summers (1995).

information from re-interview surveys. In this section, therefore, we apply the adjustment methodology developed in the US and briefly outlined at the beginning of section 2.2, 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 underlying this adjustment methodology is that of independent classification errors (ICE), and that under this assumption, estimates of the error response matrix $\hat{C}$, can be obtained directly from $\hat{K}$; the matrix $\hat{K}$ being estimated according to the procedure outlined in the previous section using the data from the re-interview survey. Recall also that in such a set-up the relationship between the true unobserved flows and the observed gross flows will be given by equation (3)

$$Vec\hat{F}^*(t-1,t) = \left[\hat{K}^{-1}(t) \otimes \hat{K}^{-1}(t-1)\right] \cdot VecF(t-1,t) \qquad 3)$$

where, to recap: $F(t-1,t)$ denotes the observed gross flows matrix between period $(t-1)$ and t; and $F^*(t-1,t)$ the true unobserved gross flows matrix.

In Table 9 we present estimates of the adjusted Spanish labour market flows, F^*, obtained when plugging into equation 3) the estimations of K obtained in section 2.2.

Evidently, 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. Whilst the extent of over-adjustment is considerably lower in the case of male flows, 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[36].

One possible explanation for this over-adjustment is that the assumption of independent misclassification errors underlying the adjustment procedure of equation (3) could, in some cases, and in particular 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

---

[36] The reduction in the degree of over-adjustment could, however, be indicative of the need to disaggregate the analysis for both the sex and the age structure of the sample, in order to obtain estimated error probabilities rates that are representative of the entire sample population and not only for specific levels of disaggregation, such as males and females.

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 (2) should be estimated instead of equation (3). Direct estimation of matrix $\hat{C}$ does, however, require access to longitudinal data of matched individuals covered by both the original and re-interview surveys, in order to correctly identify the nature of the stochastic process generating these misclassification errors. Unfortunately, although essential for further research in this area, such additional disaggregated data from the Spanish re-interview survey are 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 flows may not necessarily be a consequence of the failure of the ICE assumption. This argument can more easily be 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 4) that in order for the off-diagonal flows to be non-negative, the ratio of individuals changing state to those remaining in the same state must be greater than the ratio of errors to non-errors. This argument is coherent with the negative adjusted flows obtained both in the Spanish and in the Poterba & Summers estimations, in that, in both of these cases, the response error probabilities are high relative to the observed flows.

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 is likely to result in a considerable over-estimation of the extent of labour market dynamics, tending to give rise to an increase in the number of transitions and a reduction in the number of continual states. The ramifications of these findings 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 distorted image 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 the 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 are 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.

Our results, whilst preliminary, would tend to have serious implications for existing empirical Spanish studies, calling into question the findings of both duration and labour market turnover analyses conducted on raw LFS data. The extent to which our findings can be applied to other EU studies of a similar nature is somewhat difficult to gauge without conducting survey validation work of a similar nature to that carried out in this paper. On the one hand, the fact that, since 1992, the national LFS in all member states has adhered to the Eurostat guidelines with respect to both survey design and methodology, would suggest that the problems highlighted as being inherent in both the retrospective and matched files approach in the Spanish EPA are also very likely to be present in other national LFS data. Whether the magnitudes involved are of a level comparable to that of Spain is another matter, which obviously requires empirical verification. Whilst undoubtedly, these problems will tend to be more acute in an economy such as Spain's, characterised by a not insignificant level of informal economic activity and the dominance of fixed-term employment contracts, the trend towards an increasing use of more flexible employment arrangements throughout the EU would suggest that such issues could become even more relevant in the not too distant future in other member states.

It is true in principle that one of the ways in which estimates of the error probability matrix based on information regarding the quality of the underlying data available from quality control procedures, such as a re-interview surveys, can be used is in the adjustment of the observed gross flows for misclassification error. Yet 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 re-interview 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. (1985): "Unemployment Through the Filter of Memory", *Quarterly Journal of Economics*, pp. 747-773.

Alba, A. (1996a): "Employment transitions of young workers in Spain", Universidad Carlos III de Madrid, WP 96-61.

Alba, A. (1996b): "Explaining the transitions out of unemployment in Spain: The effect on unemployment insurance" Universidad Carlos III de Madrid, WP 96-71.

Alba, A. (1997a): "Labour force participation and transitions of older workers in Spain". Universidad Carlos III de Madrid., WP 97-39.

Alba, A. (1997b): "How temporary is temporary employment in Spain?", Universidad Carlos III de Madrid, WP 97-14.

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

Antolin, P.(1995): "Transition Probabilities to Employment and Non-Participation", Instituto Valenciano de Investigaciones Económicas, WP-EC 95-20.

Antolin, P. (1996). "Gross Worker Flows: How Does the Spanish Evidence Fit the Stylized Facts?, Centre for Economic Policy Research. Discussion Paper No. 1398, May 1996.

Arulampalam, W. and Stewart, M. (1995): "The Determinants of Individual Unemployment Durations 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 the Gross Flows of U.S. Workers ", *Brookings Papers on Economic Activity*, 2, pp. 85-155.

Blanchard, O.J. and Diamond, P.A. (1992): "The Flow Approach to Labor 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. Servicio de Estudios. Documento de trabajo Nº 9657.

Bowers, N. and Horvarth, W. (1984): "Keeping Time: An Analysis of Errors in the Measurement of Unemployment Duration". *Journal of Business & Economic Statistics,* Vol 2, No 2, April.

Burda, M. and Wyplosz, C. (1990): "Gross Labor Market Flows in Europe: Some Stylised Facts", CEPR Discussion Paper Nº 439.

Burda, M. and Wyplosz, C. (1993): "Gross Worker and Job Flows in Europe", CEPR Discussion paper Nº 868.

Chua, T.C. and Fuller, W. A (1987): " A Model for Multinomial Response Error 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 : A Reconsideration". *Brookings Papers on Economic Activity*, 1, pp.13-60.

Davis, S.J., Haltiwanger, J.C. and Schuh, S. (1996) *Job Creation and Destruction.* The MIT Press.

Feldstein, M.S. (1973): "Lowering the Permanent Rate of Unemployment". U.S. Congress Joint Economic Committee, Washington D.C.

Fuller, W. A. and 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". *Revista Española de Economía*, Vol 13. Nº 2.

Gómez Salvador, R. and Dolado, J.J. (1995): "Creacion y Destrucción de Empleo en España: Un Análisis Descriptivo con Datos de la CBBE". Banco de España. Servicio de Estudios. Documento de trabajo Nº 9526 .

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

Instituto Nacional De Estadlstica (1995): Evaluación de la calidad de los datos de la Encuesta de Población Activa 1994.

Instituto Nacional De Estadística (1996): Encuesta de Población Activa. 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,* 41, pp 45-72.

Konnings, J. (1993): "Job Creation and Job Destruction in the UK Manufacturing Industries", *Oxford Bulletin of Economics and Statistics*, 57 (1), pp 5-24.

Levine, P.B. (1990): Contemporaneous vs. Retrospective Unemployment: Through the Filter of Memory or the Muddle of the Current Population Survey. Industrial Relations Section, Department of Economics, Princeton University. WP Nº 276.

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 Spells Lengthen?" *Journal of Applied Econometrics*, Vol 8, pp. 361-381.

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

Peracchi, F. and Welch, F. (1994): "Trends in Labor Force Transitions of Older Men and Women". *Journal of Labour Economics*, 12(2) , pp. 210-242.

Peracchi, F. and Welch, F. (1995): "How Representative are Matched Cross-Sections? Evidence from the *CPS*. *Journal of Econometrics*, Nº 68, pp. 153-179.

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

Portugal, P. and Addison, J. (1997) "Unemployment insurance and joblessness: a discrete duration model with multiple destinations". Banco de Portugal WP.

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*, Vol 54 (6), pp. 1319-1338.

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*., Vol 77(2), pp. 207-216.

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

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.

Table 1: Average Monthly Gross Labour Market Flows 1977-1982
(figures in 000's)

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

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

Table 2: Actual and *Recalled* Labour Market Status of *EPA* Individuals in the Second Quarter 1992.
(figures in 000's.)

Actual LMS	Recalled LMS				
	Working	Searching	NILF	Military Service(*)	Total
Employed	3350.4	70.4	77	5.1	3502.9
Unemployed	69	549.4	122.8	2.5	743.7
NILF	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

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

Panel A: Observed Transition Probabilities (Calculated from the Retrospective Questionnaire)			
		Actual LMS 2/93	
LMS in 2/92 as **Recalled in 2/93**	Employed	Unemployed	NILF
Employed	0.875	0.082	0.043
Unemployed	0.272	0.643	0.085
NILF	0.028	0.034	0.938

Panel B: Recall Errors		**Actual-LMS**	
Recalled LMS	Employed	Unemployed	NILF
Employed	0.965	0.020	0.015
Unemployed	0.094	0.734	0.172
NILF	0.017	0.027	0.955

Source: calculations based on data presented in Tables II.1 and II.4 *INE* (1996)[37].

---

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

Table 4: Survey Response Inconsistencies in the *EPA*.

1. Survey Response Inconsistencies: Re-interview Survey

	Re-interview			
Initial Interview	Employed	Unemployed	NILF	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

2. Spurious Transition Probabilities from the Interview and Re-interview Surveys

	Re-interview			
Initial Interview	Employed	Unemployed	NILF	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. Observed Transition Probabilities First to Second Quarter 1994

	Second Quarter			
First Quarter	Employed	Unemployed	NILF	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

Source: calculations based on matched files of individuals and on data presented in table C.4 *INE* (1995).

Table 5 : Labour Market Status in Original and Re-interview Survey (First quarter 1994)[38]

	Re-interview			
Original Interview	**Employed**	**Unemployed**	**NILF**	**Total**
Employed	2007	65	75	2147
	(0.327)	(0.011)	(0.012)	(0.350)
Unemployed	60	562	108	730
	(0.010)	(0.092)	(0.018)	(0.119)
NILF	93	159	3008	3260
	(0.015)	(0.026)	(0.490)	(0.531)
Total				6137
				(1.000)
Men	**Employed**	**Unemployed**	**NILF**	**Total**
Employed	1387	41	31	1439
	(0.469)	(0.014)	(0.011)	(0.494)
Unemployed	36	322	39	397
	(0.012)	(0.111)	(0.013)	(0.136)
NILF	39	50	989	1078
	(0.013)	(0.017)	(0.339)	(0.370)
Total				2914
				(1.000)
Women	**Employed**	**Unemployed**	**NILF**	**Total**
Employed	624	24	44	692
	(0.196)	(0.008)	(0.014)	(0.218)
Unemployed	22	237	68	327
	(0.007)	(0.075)	(0.021)	(0.103)
NILF	54	108	1996	2158
	(0.017)	(0.034)	(0.628)	(0.679)
Total				3177
				(1.000)

Source: *INE* (1995).

---

[38]Figures in parenthesis refer to sample proportions.

Table 6: Estimated Parameters.

Parameter	Estimated Value	Standard Error
Males and Females		
P_1	0.3498	0.0044
P_2	0.1189	0.0030
α_{21}	0.0588	0.0053
α_{31}	0.0292	0.0024
α_{32}	0.0987	0.0064
ϑ_{12}	1.1502	0.2598
ϑ_{13}	1.4715	0.2743
ϑ_{32}	1.8693	0.2397
Males		
P_1	0.4938	0.0069
P_2	0.1363	0.0046
α_{21}	0.0612	0.0070
α_{31}	0.0252	0.0033
α_{32}	0.0724	0.0080
ϑ_{12}	1.2628	0.3497
ϑ_{13}	1.5290	0.4354
ϑ_{32}	1.5240	0.3627
Females		
P_1	0.2178	0.0053
P_2	0.1029	0.0039
α_{21}	0.0536	0.0081
α_{31}	0.0427	0.0046
α_{32}	0.1324	0.0106
ϑ_{12}	1.1722	0.4384
ϑ_{13}	1.4468	0.3523
ϑ_{32}	2.0691	0.3169

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	1.0000	1.0000	1.0000
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	1.0000	1.0000	1.0000
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	1.0000	1.0000	1.0000

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
(First Quarter 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

Table 9: Observed and Adjusted Flows in *EPA* across the First and Second Quarter 1994. (number of people unweighted)

First Quarter	Second Quarter 1994		
	Employed	Unemployed	NILF
Total Unadjusted Flows			
Employed	40,908	1535	771
Unemployed	1,902	10,748	1,029
NILF	745	1,328	35,238
Total Adjusted Flows			
Employed	43,624	439	-474
Unemployed	878	13,997	-745
NILF	-554	-384	37,424
Unadjusted Men			
Employed	27,558	1033	354
Unemployed	1264	5395	391
NILF	286	459	9,794
Adjusted Men			
Employed	28,864	425	-136
Unemployed	692	6,651	-153
NILF	-226	-70	10,486
Unadjusted Women			
Employed	13,350	502	417
Unemployed	638	5,353	638
NILF	459	869	25,444
Adjusted Women			
Employed	14,760	45	-346
Unemployed	216	7,433	-669
NILF	-331	-383	26,945

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.

ALTERNATIVE PARAMETERISATIONS OF THE MATRIX G:

In order to test the robustness of the estimates of the error response probabilities, a number of alternative specifications, to that given by equation 6, of the matrix G were estimated. In particular, we report here the following alternative assumptions.

Assumption 4.1. In the first instance, it was assumed that the error response probabilities in the re-interview were given by:

$$G_{ij} = \left(1 - \sum_{l=1}^{3} \alpha_{lj}\gamma_{il}\right)\delta_{ij} + \alpha_{ij}\gamma_{ij},$$

where: $\gamma_{ij} = 1 - \gamma_{ji}$; δ_{ij} denotes Kronecker's delta and the same restrictions prevailing in assumption 4) apply to the parameters α_{ij}, namely: $\alpha_{ii} = 0$, $\alpha_{ij} = \alpha_{ji}$. Note that this specification is simply a more general case of the K formulation used in the main text, in which γ_{ij} replaces the ratio of marginal proportions $p_i/(p_i + p_j)$. As noted by Fuller- Chua (1985, p. 67) this alternative parameterisation results in estimates of the $\alpha's$, which are close to those one would obtain when using a symmetric, instead of the observed interview-re-interview, table in which both the ij-th and ji-th entries are equal to the average values of the A_{ij} and the A_{ji} cells.

Re-specifying the error probabilities rates for the re-interview survey in this manner results in the following matrix of estimated error probability rates:

	0.966	0.045	0.013
K=	0.015	0.868	0.019
	0.019	0.087	0.968
	1.000	1.000	1.000

Assumption 4.2. As noted in footnote 34, one criticism of the Fuller-Chua specification for the G matrix used in the main text, is that it implicitly assumes that the probability of an error occurring in the re-interview is at least as big as in the original interview. Such an assumption would, however, appear to be inconsistent with the fact that the re-interview is in fact carried our more carefully by, for example, senior interviewers, than the original LFS survey. In this light, a more logical parameterisation would therefore appear to be one in which a higher probability of error is assigned to the original survey. In

our third specification, therefore, the inequality conditions of Assumption 4) are simply reversed in order to allow for this possibility. In other words, it is assumed that:

$$G_{ij} = \vartheta_{ji} K_{ij} \qquad \text{if } A_{ij} \geq A_{ji} \text{ and } i \neq j$$

$$G_{ij} = K_{ij} \qquad \text{if } A_{ij} < A_{ji} \text{ and } i \neq j$$

$$G_{ij} = 1 - \sum_{i=1}^{r} G_{ri} \qquad \text{if } i = j,$$

where $\vartheta_{ij} \geq 1$.

This reformulation results in the following matrix of response error probabilities.

K=	0.963	0.046	0.014
	0.016	0.838	0.026
	0.021	0.116	0.960
	1.000	1.000	1.000

As is evident, both of these re-specifications of the re-interview error response matrix, G, result in estimated error probabilities of a greater magnitude that those reported in Table 7. These results would tend to suggest that the error response matrix. K, reported in the main text should be viewed as representing a lower bound for the extent of misclassification errors in the Spanish LFS.

APPENDIX 3.

THE ESTIMATION OF THE PARAMETERS OF THE MODEL

Let:

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

$$Y = f(\psi) + e,$$

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 re-interview. In other words, the elements of vector Y are simply the first 8 elements of the interview-re-interview matrix. $f(\psi)$ 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:

$$\psi = \left(\alpha_{21}, \alpha_{31}, \alpha_{32}, \vartheta_{12}, \vartheta_{13}, \vartheta_{32}, P_1, P_2 \right).$$

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

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

where n is the sample size.

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

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

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

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

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

APPENDIX 4

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

(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 can be expressed as:

$$K = \begin{bmatrix} (1-\varepsilon) & \delta \\ \varepsilon & (1-\delta) \end{bmatrix} \quad ,$$

where: ε = Probability (observed in state U|truly in state E) and δ = Probability (observed in state E| true state is U). Equation (3) in the main text in this framework reduces to:

$$VecF^*(t-1,t) = \hat{K}^{-1}(t)VecF(t-1,t).$$

This in itself can be expressed as:

$$P^* = \frac{1}{(1-\varepsilon-\delta)}\begin{bmatrix} (1-\delta)EE - \delta EU & -\varepsilon EE + (1-\varepsilon)EU \\ (1-\delta)UE - \delta UU & -\varepsilon UE + (1-\varepsilon)UU \end{bmatrix} = \begin{bmatrix} EE* & EU* \\ UE* & UU* \end{bmatrix}$$

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

(A3) ε < 0.5 and δ < 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:

$$EU < EE \quad \text{and} \quad UE < UU.$$

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:

$$EE* = (1-\delta)EE - \delta EU = (1-\delta)[EE - EU] + (1-2\delta)EU,$$

2) but EU* can be negative;

$$EU^* = -\varepsilon.EE + (1-\varepsilon).EU$$

$$= EU[\varepsilon + 1 - \varepsilon] - \varepsilon.EE$$

$$= EU(1 - \varepsilon - EE.\varepsilon)_.$$

This final expression will be positive if $\dfrac{EU}{EE} > \dfrac{\varepsilon}{(1-\varepsilon)}$.

In short, the off diagonal term EU^* is positive if and only if the ratio of those who change state to those who remain in the same one, (EU / EE), is greater than the ratio of errors to non-errors $\varepsilon/(1-\varepsilon)$. Note that the same type of analysis follows for the other off diagonal element UE^*.

WORKING PAPERS (1)

9802 **Pilar Álvarez Canal:** Evolución de la banca extranjera en el período 1992-1996.

9803 **Ángel Estrada y Alberto Urtasun:** Cuantificación de expectativas a partir de las encuestas de opinión.

9804 **Soyoung Kim:** Monetary Policy Rules and Business Cycles.

9805 **Víctor Gómez and Agustín Maravall:** Guide for using the programs TRAMO and SEATS.

9806 **Javier Andrés, Ignacio Hernando and J. David López-Salido:** Disinflation, output and unemployment: the case of Spain.

9807 **Olympia Bover, Pilar García-Perea and Pedro Portugal:** A comparative study of the Portuguese and Spanish labour markets.

9808 **Víctor Gómez and Agustín Maravall:** Automatic modeling methods for univariate series.

9809 **Víctor Gómez and Agustín Maravall:** Seasonal adjustment and signal extraction in economic time series.

9810 **Pablo Hernández de Cos e Ignacio Hernando:** El crédito comercial en las empresas manufactureras españolas.

9811 **Soyoung Kim:** Identifying European Monetary Policy Interactions: French and Spanish System with German Variables.

9812 **Juan Ayuso, Roberto Blanco y Alicia Sanchís:** Una clasificación por riesgo de los fondos de inversión españoles.

9813 **José Viñals:** The retreat of inflation and the making of monetary policy: where do we stand?

9814 **Juan Ayuso, Graciela L. Kaminsky and David López-Salido:** A switching-regime model for the Spanish inflation: 1962-1997.

9815 **Roberto Blanco:** Transmisión de información y volatilidad entre el mercado de futuros sobre el índice Ibex 35 y el mercado al contado.

9816 **M.ª Cruz Manzano e Isabel Sánchez:** Indicadores de expectativas sobre los tipos de interés a corto plazo. La información contenida en el mercado de opciones. (Publicada una edición en inglés con el mismo número.)

9817 **Alberto Cabrero, José Luis Escrivá, Emilio Muñoz and Juan Peñalosa:** The controllability of a monetary aggregate in EMU.

9818 **José M. González Mínguez y Javier Santillán Fraile:** El papel del euro en el Sistema Monetario Internacional.

9819 **Eva Ortega:** The Spanish business cycle and its relationship to Europe.

9820 **Eva Ortega:** Comparing Evaluation Methodologies for Stochastic Dynamic General Equilibrium Models.

9821 **Eva Ortega:** Assessing the fit of simulated multivariate dynamic models.

9822 **Coral García y Esther Gordo:** Funciones trimestrales de exportación e importación para la economía española.

9823 **Enrique Alberola-Ila and Timo Tyrväinen:** Is there scope for inflation differentials in EMU? An empirical evaluation of the Balassa-Samuelson model in EMU countries.

9824 **Concha Artola e Isabel Argimón:** Titularidad y eficiencia relativa en las manufacturas españolas.

WORKING PAPERS (1)

9901 **José Ramón Martínez Resano:** Instrumentos derivados de los tipos *Overnight: call money swaps* y futuros sobre fondos federales.

9902 **J. Andrés, J. D. López-Salido and J. Vallés:** The liquidity effect in a small open economy model.

9903 **Olympia Bover and Ramón Gómez:** Another look at unemployment duration: long-term unemployment and exit to a permanent job. (The Spanish original of this publication has the same number.)

9904 **Ignacio Hernando y Josep A. Tribó:** Relación entre contratos laborales y financieros: Un estudio teórico para el caso español.

9905 **Cristina Mazón and Soledad Núñez:** On the optimality of treasury bond auctions: the Spanish case.

9906 **Nadine Watson:** Bank Lending Channel Evidence at the Firm Level.

9907 **José Viñals:** El marco general de la política monetaria única: racionalidad, consecuencias y cuestiones pendientes.

9908 **Olympia Bover and Manuel Arellano:** Learning about migration decisions from the migrants: an exercise in endogenous samplig and complementary datasets.

9909 **Olympia Bover and Pilar Velilla:** Migrations in Spain: Historical background and current trends.

9910 **Fernando Restoy:** Los mercados financieros españoles ante la Unión Monetaria.

9911 **Luis J. Álvarez and M.ª de los Llanos Matea:** Underlying inflation measures in Spain.

9912 **Regina Kaiser and Agustín Maravall:** Estimation of the business cycle: a modified Hodrick-Prescott filter.

9913 **Enrique Alberola and José Manuel Marqués:** On the relevance and nature of regional inflation differentials: The case of Spain.

9914 **Agustín Maravall:** An application of TRAMO and SEATS. Report for the «Seasonal Adjustment Research Appraisal» project.

9915 **Regina Kaiser and Agustín Maravall:** Seasonal outliers in time series.

9916 **Enrique Alberola and Humberto López:** Internal and external exchange rate equilibrium in a cointegration framework. An application to the Spanish Peseta.

9917 **José Viñals and Javier Vallés:** On the real effects of monetary policy.

9918 **Regina Kaiser and Agustín Maravall:** Short-term and long-term trends, seasonal adjustment, and the business cycle.

9919 **J. Andrés, J. D. López-Salido and J. Vallés:** Intertemporal substitution and the liquidity effect in a sticky price model.

9920 **J. Andrés, I. Hernando and J. D. López-Salido:** The role of the financial system in the growth-inflation link: The OECD experience.

9921 **Ángel Estrada and Ignacio Hernando:** Microeconomic price adjustments and inflation: Evidence from Spanish sectoral data.

9922 **Concha Artola and Una-Louíse Bell:** Identifying Labour Market Dynamics using Labour Force Survey Data.

---

(1) Previously published Working Papers are listed in the Banco de España publications catalogue.