Scand. J. of Economics 109(4), 807–836, 2007
DOI: 10.1111/j.1467-9442.2007.00510.x

Occupational and Job Mobility in the US*

Giuseppe Moscarini
Yale University, New Haven, CT 06520-8268, USA
giuseppe.moscarini@yale.edu

Kaj Thomsson
Yale University, New Haven, CT 06520-8268, USA
kaj.thomsson@yale.edu

Abstract

We propose a new methodology to measure worker mobility across occupations and jobs in the US, building on the limited longitudinal dimension of monthly CPS data. For the period 1979–2006, we find that about 3.5% of male workers employed in two consecutive months report different three-digit occupations. This rate is procyclical, mildly rising in the 1980s and falling after 1995. We also revise upward current estimates of aggregate job-to-job mobility since 1994, from 2.7% to 3.2% of employment per month. Despite extreme similarity of average levels and time-series behavior, occupational and job mobility are only weakly correlated.

Keywords: Occupational mobility; job mobility; turnover

JEL classification: *E*32; *J*62; *J*63

I. Introduction

In this paper, we propose a new methodology to measure worker mobility across occupations and jobs in US Census data at the monthly frequency. We build on our findings on occupational mobility to shed new light on employer-to-employer (EE) transitions.

There are at least three reasons why economists are interested in occupational and job mobility. First, the concept of "mobility" (defined broadly) is tightly linked to the notion of "opportunity", which is a central part of the history of ideas and the formation of a national identity in the United States. In particular, the perception of great social mobility, measured among other things by income mobility and occupational mobility (between or within generations), is typically considered one of the key components of what is sometimes referred to as "American exceptionalism", the idea that the US is different in some fundamental ways from most other countries.[1] Second, a

*We are grateful to two anonymous referees, and participants in the NY area Macroeconomics Workshop at the NY FED for extensive and useful comments. The usual disclaimer applies.
[1] For a discussion of the relationship between American exceptionalism and the concept of occupational mobility, see Ferrie (2005).

prominent tradition in macroeconomics, going back at least to Schumpeter (1939), emphasizes the continuous reallocation of resources across heterogeneous production units as the "mode" of aggregate business fluctuations and economic growth. If capital is a quasi-fixed factor, technological progress can only be implemented through the "creative destruction" of installed capital and the reallocation of labor to new production processes. In this paper, we attempt to relate to this tradition by focusing on occupational mobility at a level of disaggregation that, in our view, corresponds well to a change of technology for a worker. The same tradition has partly inspired modern equilibrium job search theory, which is empirically informed by several types of labor market flows. Lately, much attention has been paid to job-to-job transitions. We revisit this flow in light of what we learn about occupational mobility. Third, on the labor supply side, human capital may to a large extent be occupation-specific, and thus lost upon a transition; see Kambourov and Manovskii (2002).

The focus of this paper is on the actual *measurement* of labor market transitions, particularly across occupations. We believe that, when attempting to understand the nature of occupational mobility in the US, the Current Population Survey (CPS) is, for several reasons, a natural source of empirical information. First and foremost, since our focus is largely macroeconomic, we require a large and representative sample. The CPS is designed for this very purpose, and is superior in this sense to longitudinal datasets that are not representative, such as the National Longitudinal Survey of Youth (NLSY), and to others that feature smaller sample size and larger sampling error, such as the Panel Study of Income Dynamics (PSID). Second, the CPS is the primary source of US official labor market statistics. Hence, comparisons with official reports and labor market information is greatly facilitated by the use of the CPS. Third, information accrues at high, monthly frequency.

Measuring labor market transitions is in general a very difficult task, because mobility rates are typically small at the high frequency required to avoid severe time aggregation, and thus are very sensitive to measurement error in the labor market state of interest. This is particularly true in the case of occupations, which, unlike the three employment states (employed/unemployed/out of the labor force), number from dozens to thousands, depending on the level of disaggregation. The variety of tasks that US workers perform makes the description by the survey respondent often problematic, and small errors in stock classification lead to large spurious flows. In this difficult scenario, Murphy and Topel (1987) and Kambourov and Manovskii (2004) point out that there are particular problems with the annual (March) CPS files. Both studies provide convincing evidence of vast classification error in occupations, leading to hugely inflated transition rates, and the latter study notes the lack of a well-defined time period over which the mobility that can be derived from the March CPS files is measured.

We use *monthly* rather than annual CPS data in order to minimize problems of time aggregation and to exploit the longitudinal nature of monthly data to deal with occupational coding error. Our approach builds on two main ideas. First, since 1994 the survey includes a battery of so-called dependent coding questions that are meant to minimize false transitions. Second, both before and after 1994, we inspect each single occupational transition between two consecutive months under the magnifying lens of a "global" view of that worker's employment history over four consecutive months, including one month before and one after the two months spanned by the transition. Records that are (based on dependent coding and other cross-checks) reliable provide guidance on which kinds of transitions and four-month career trajectories are plausible. Sequences of four consecutive occupations that involve unusual mobility patterns, and that do not correspond to changes in industry or class of workers or to active job search in the past month, are suspicious and should be, to a large extent, discarded.

Our main results are readily summarized. While raw data are extremely noisy before 1994, and still somewhat suspicious after 1994, applying our filters makes pre- and post-1994 occupational mobility lower and essentially identical on average over the entire 1979–2006 period. About 3.5% of workers employed in two consecutive months report different three-digit occupations, implying a significant change of career. This flow is procyclical, rising from 1979 to the mid-1990s and then falling quite sharply, especially after the 2001 recession, with a cyclical rebound only after 2004.

Based on these findings, we then turn our attention to job-to-job (EE) transitions after 1994. Here, our contribution is different. Previous studies of this very important flow in the monthly CPS, most notably, Fallick and Fleischman (2004) and Nagypal (2004), discard the 3.4% of the sample where an answer to the question on job change is missing. We are able to assign reliable information on occupational mobility to those records with missing job mobility information. We then exploit the correlation between occupation and job mobility to impute an answer, when missing, to the job mobility question. This leads us to revise upward current estimates of the EE flow, from 2.7% to 3.2% per month. The time-series pattern of the revised EE rate remains roughly in line with results from previous studies: flat in the 1990s, falling in 2001–2004, rising in 2004–2006. But a different interpretation is also possible, where the EE mobility rate mildly declines in 1994–1996, rises in 1997–2000 and drops precipitously from the 2001 recession to 2004, to rebound only and modestly in 2004–2006. This second interpretation suggests that the negative impact of recessions on job-to-job mobility is extremely persistent.

Finally, we document that, in spite of their almost identical average levels and time-series behaviors, occupational and job mobility are far from being perfectly correlated. About 40% of all occupational transitions are internal to an employer, and over a third of all job-to-job transitions involve no change whatsoever in the narrowly defined tasks that the worker performs.

Although our exercise consists essentially of measurement, it is certainly not theory-free. Our cleaning of raw data on labor market transitions exploits the unique high-frequency longitudinal nature of the monthly CPS to assess the plausibility of various kinds of transitions. We exploit this feature to extend backwards, to 1979–1993, the significant improvements introduced in the survey in 1994, but we also inspect and further correct the much more reliable post-1994 observations. All this requires a fair amount of judgment calls. Standard theories of worker turnover underlie some of our choices. Turnover begets turnover, that is, the chance of a separation from a job or occupation is much higher if a separation occurred recently, both across jobs, as in Jovanovic (1979) and Farber (1994), and occupations, as in McCall (1990). At the same time, two consecutive monthly career switches from occupation A to B and back to A appear unlikely, because turnover is in part motivated by dissatisfaction with the initial match in A.

We take confidence in our choices because of the consistency of our results, both across subperiods and with other surveys. After 1994, we rely on detailed questions about the actual change of tasks and activities to isolate the likely miscoded records, where we find an extent of measurement error that is almost exactly the same as before 1994, although we use partly different procedures to address those two subperiods. The final estimates of occupational mobility that we settle upon are essentially the same over the two subperiods, when data were collected differently. Finally, the 3.5% average monthly mobility rate across three-digit occupations that we find is easily compatible, after taking into account correlated transitions and difference in definitions, with the annual 20% rate found by Kambourov and Manovskii (2008) in an extensive revision of PSID information from 1968–1997. The time-series behavior of our and their series is the same in the overlapping 1979–1997 period.

Section II contains a brief introduction to the monthly (basic) CPS, describes the procedure of matching cross-sectional information from different months, and covers the main issues with the occupational information in the CPS. Section III is an overview of our approach to dealing with these issues. In Section IV we apply this approach and produce our estimates of genuine occupational mobility. Section V addresses job-to-job mobility. Section VI concludes.

II. Issues with Matching Files and Occupational Codes in the CPS

The Current Population Survey

The Current Population Survey (CPS) is a monthly survey of about 50,000 households, which has been conducted by the Bureau of the Census for the Bureau of Labor Statistics for more than 50 years. However, because the information required to reliably match individuals over time is only available in later surveys, for our purposes information is available for the past 28 years only.[2]

Until very recently, studies of labor market transitions in the CPS have used the March Income Supplement files, which is a collection of files with data released on an annual basis. Given our interest in labor market transition, in this study we use instead the Basic Monthly files. As we will see shortly, this high frequency allows us to fully exploit the limited, yet still very informative, longitudinal component of the CPS, and to minimize attrition.

In fact, despite not being primarily intended for longitudinal analysis, the Current Population Survey contains a panel component and can be used to follow individuals over short periods of time. In each month the full CPS sample is divided into eight "rotation groups", with each housing unit being interviewed for four consecutive months, then removed from the sample for an eight-month period and finally interviewed for another four months. Hence, in any month one-eighth of the sample households, the first rotation group, are interviewed for the first month, one-eighth are interviewed for the second month, one-eighth for the third month, etc. Since the interviewers follow housing units (i.e., addresses) and not families or individuals, attrition can then occur for one of three main reasons: temporary absence (hospitalization, imprisonment, vacation), migration (to go to college, to enlist in the military, to form a family, to follow or to separate from a spouse, and for work-related reasons, including retirement), and mortality. To minimize attrition and the resulting possible sample selection, we will restrict attention to the first four months in the sample. Details follow.

The CPS has several advantages and disadvantages over panel datasets, such as the PSID and the NLSY, for studying various measures of labor market states (employment/unemployment, even disaggregated by various demographics, occupation and industry membership, etc.) and transitions

[2] Most of the overview information presented in this section is directly based on the official description of the CPS at the Bureau of Labor Statistics website (www.bls.census.gov/cps). We use a particular package—The CPS Utilities—consisting of data, documentation and Windows software developed by the company UNICON Inc. in order to simplify the process of finding and extracting data from the CPS.

(across employment states, occupations, industries, etc.). The first advantage is the very large number of individuals in the sample. This means that the CPS is particularly suitable to eliminate most sampling error when measuring employment distributions across cells that number in the dozens or even many hundreds, such as three-digit occupations, birth cohorts, etc. The second advantage is the high frequency of observations and time-series dimension, as the CPS is conducted monthly, as opposed to panels that conduct yearly interviews about the entire history of the previous 12 months. The monthly frequency minimizes (although does not completely eliminate) time aggregation problems due to multiple within-period undetected transitions and incorrect respondent's recall of past events, both obvious problems with annual data. The third advantage is the wealth of information about demographics, which compares well with that of proper panel data.

The main disadvantage of the CPS for our purposes is its address-based nature, which aligns attrition with geographical mobility, potentially correlated with occupational mobility. In contrast, panel datasets continue to track the same individuals wherever they are, although panels too suffer from significant attrition because of their much lower interview frequency. The Survey of Income Program Participation has the same desirable features of the monthly CPS and does not suffer from geographical attrition, but it covers a short period of time. In the working paper version of this article, Moscarini and Thomsson (2006), we use annual (March) CPS files to derive an estimate of monthly geographical mobility, and we find it to be small relative to the monthly rate of occupational mobility that we estimate for those who stay at the same address. Therefore, this issue turns out to be a minor concern. Another disadvantage of the CPS is the very limited longitudinal dimension, as individuals are followed for eight (non-consecutive) months, as opposed to decades for panel surveys. This is an unavoidable consequence of the much richer information set provided by the CPS: since so many questions are asked, they can be asked only a few times, lest becoming harassment. Our cleaning procedure is designed specifically to take this limitation of the CPS into account; four consecutive observations on the same individual is the least that we need to create our filter for spurious occupational transitions. Finally, coding of occupations and industries in the CPS changes with every decennial census, while panel datasets have a uniform coding maintained throughout the period.[3] We address this issue in a later subsection.

[3] This is a mixed blessing because the Census revises coding every 10 years precisely to avoid carrying outdated classifications. For example, suppose that a three-digit occupational group grows in employment membership and in the variety of tasks it encompasses, so as to lead the Census Bureau to subdivide it into two separate three-digit codes. Then, the old coding system will miss the mobility among these new categories, thus understating true mobility.

Matching of CPS Files and Sample Characteristics

In principle, the re-interviewing process in the monthly CPS should allow us to match three-fourths of the sample in any given month to the next month, while one-fourth of the sample exits due to rotation (though individuals in their fourth month can be linked eight months forward). However, various kinds of attrition and absence reduce the fraction of individuals that can actually be matched. Two recent papers, Madrian and Lefgren (2000) and Feng (2001), evaluate in depth the design of the matching criteria of annual (March) CPS records.[4] In this study, we adopt the somewhat more conservative matching approach, suggested by Madrian and Lefgren (2000), and we apply it to monthly CPS files.[5] That is, we first match datasets based on the formal identifier variables (*HHID, HHNUM, LINENO*), and then we evaluate the validity of potential match-pairs by comparing the sex, age and race variables. In addition, in order to minimize selection to begin with, we focus only on individuals in their first four months of the sample, and study their mobility between months 2 and 3, so as to eliminate the additional attrition that takes place during the eight months between the two sampling periods. Furthermore, we restrict attention to men who are employed in two consecutive months.

Aggregate match statistics for men aged 16–64 during the years 1979–2006 are shown in Figure 1 and Table 1. We show results only for individuals that are in the CPS for their second month, as these are the only

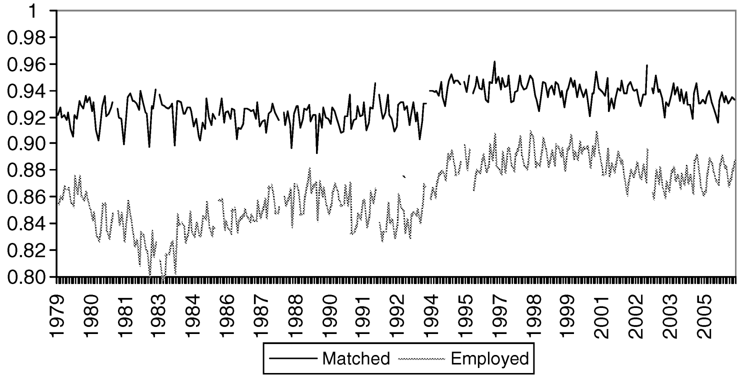

Fig. 1. Proportion of matched records and of employed in consecutive months

[4] Both of these papers are inspired by earlier work in Welch (1993) and Peracchi and Welch (1994).

[5] That is, we choose to be somewhat more conservative in our assessment of the validity of matches than Feng (2001). Simple checks indicate, however, that the difference in the final match rates between the two different proposed criteria would be small.

Table 1. *Sample characteristics*

	Labor force			Labor force, matched		
Characteristic	All	Unmatched	Unmatched and unemployed	All	Employed at t	Employed at t and $t+1$
Age	37.77	33.73	30.33	38.09	38.41	38.52
Race						
White (%)	87.66	82.25	73.32	88.11	88.65	88.77
Black (%)	7.94	11.67	19.62	7.63	7.15	7.05
Other (%)	4.39	6.08	7.06	4.25	4.19	4.18
Family status						
Head of household	69.58	53.87	38.83	70.83	71.99	72.36
Married	63.53	44.50	28.36	65.14	66.51	66.94
Education						
< High school (%)	17.51	26.34	42.97	16.80	15.82	15.53
High school (%)	57.82	55.50	49.46	57.96	58.05	58.03
College (%)	24.67	18.16	7.57	25.23	26.13	26.44
Number of observations	1,505,884	112,270	17,964	1,392,917	1,314,943	1,289,822

ones used in our measurement of mobility. In light of the 1994 overhaul of CPS interviewing methodology, we expect and indeed find the post-1994 data to be of higher quality than the pre-1994 data.

Figure 1 displays the time series of the match rate. The average rate is 92.20% before 1994 and 93.88% afterwards.[6] These match rates are higher than the rates that have been achieved with the March CPS files, for instance by Madrian and Lefgren (2000). Still, quite a few individuals in their second month are lost when we attempt to match them with their own observation one month ahead. The implications for our interest in occupational mobility are not perfectly clear, since we do not know what happens to the individuals who fall out of the sample. However, a simple check that can be done is to compare the matched sample with the full sample. Table 1 displays the observable characteristics of individuals in their second month. There seems to be only one truly significant difference: the individuals in the matched sample have more education than the average for the full sample. Of course, there might also exist a selection on unobservable variables that is significant for the case of occupational mobility.[7]

We also note a new fact from Figure 1. It is clear that the match rate is mildly countercyclical. That is, it is easier to match files in a recession. This is suggestive of the role played by migration for job search reasons, because labor market mobility is typically procyclical, as is occupational mobility. The cyclical volatility is, however, small, as is that of geographic attrition.

Occupational Coding Issues

In this paper, an "occupation" is a code defined by the US Census three-digit occupational classification system. Examples are bartenders, accountants, computer programmers, chemistry teachers and automobile

[6] Note that some months are missing in the time series. The reason is that between some months the individual identifier was, for various reasons such as the confidentiality of the respondents, changed and/or non-unique. Hence, these months cannot be matched. This is the case for pre-1978 data, and for some months in 1985, 1994 and 2004. In order to correct for this, we simply removed observations from the months that cannot be matched, so all the results presented below are derived only for the months in which the individual identifier variables at least theoretically should uniquely identify individuals during their (maximally) eight months in the sample.

[7] For further analyses of the consequences of attrition, see Moscarini and Thomsson (2006), and for more in-depth treatments, see Welch (1993) and Peracchi and Welch (1994). Somewhat surprisingly, Peracchi and Welch (1994) conclude that while selection on the matched records can introduce various biases, "no major bias appears in the estimates of transitions between labor force states after controlling for sex, age and labor force status at the time of the first survey". Note that this is a conclusion regarding the potential problems in the measurement of a flow variable (transitions in and out of the labor force), and not just about the measurement of the underlying stock variable.

mechanics. Measuring occupations at this level corresponds most closely to the notion of "labor technology", with labor input differentiated by the tasks involved and the kind of training required.[8]

Although the three-digit Census codes correspond nicely to concepts of theoretical interests, using the occupational classification in the CPS to measure aggregate worker mobility is not without problems. In particular, the system of classification has changed several times. This complicates any analysis using time-series or longitudinal data. First of all, during the period we consider (1979–2006), the occupational codes were modified on three occasions. Second, the interviewing technique changed dramatically in 1994. Third, occupational data for some individuals are missing and have been imputed, with the obvious implication that occupational mobility might be overestimated. We now discuss briefly the first two issues, and address the third issue after presenting the results.

Two changes in occupational codes, in 1992 and 2003, were negligible, involving a handful of categories eliminated or added, and a very small fraction of employment. However, the 1983 change was significant, as the number of three-digit occupations increased by about 17%, from 375 to 438, which *per se* would tend to increase measured mobility. This is a straightforward consequence of the methods of classification, and we have made no attempt to correct for these changes.[9]

An overhaul of the interviewing technique took place in 1994. In order to understand the fundamental change in the reliability of the occupational mobility data before and after 1994, we briefly describe the respective interviewing procedures.[10] Before 1994, occupations were coded into numbers independently ("independent coding"). That is, *every* month, respondents were asked anew: (i) for whom they worked, (ii) what kind of business that was, (iii) what kind of work they were doing, (iv) what their most important activities were, and (v) what sector they were working in. This information was later used by CPS staff to assign occupation (and industry)

[8] The alternative to using three-digit level would be the two-digit occupation coding system, which we consider too coarse. For instance, both "Chefs" and "Waiters" (three-digit occupations) fall under the two-digit code of "Food Preparation and Related", and both "Architects" and "Biomedical Engineers" fall under the same two-digit code of "Architecture and Engineering Occupations". Clearly, the three-digit codes here identify what most people think of as an occupation, whereas the two-digit codes are groupings of related but distinct occupations.

[9] However, by removing all the matched month-pairs that span two different classification systems, we at least make sure that no occupational mobility observation is due to the fact that the same job, for the same individual, is classified differently in two different months.

[10] This description is based on Polivka and Rothgeb (1993) and the UNICON documentation reporting the original questions from the Census protocols. For more information about the coding procedures, we refer to these sources, as well as to an extended discussion in Moscarini and Thomsson (2006).

codes to each individual, a procedure that had at least two serious problems. First, asking these questions was very cumbersome for the interviewer, and respondents typically complained about answering the same questions repeatedly. Second, and more importantly for this paper, asking these questions independently from month to month introduced a significant amount of spurious shifts in occupation and industry. Indeed, in a small validation study of occupational coding based on company records and employees' descriptions of their own tasks, Mathiowetz (1992) finds a roughly 50% error when occupations are coded without telling the coders that the two records concern the same individual. More remarkably, when told that the two records did come from the same individual, expert coders still found a 12% disagreement rate between the company record and the employee's description of the employee's task, although each coder knows that they should agree.

To reduce the interview burden and the possibility of misclassification, a number of changes were introduced in 1994. Most importantly, "dependent coding" (sometimes referred to as "dependent interviewing") was introduced. Individuals interviewed in successive months were asked the following questions:

1. "*SAMEJOB*". Last month, it was reported that (name/you) worked for (input company name). (Do/Does) (you/he/she) still work for (input company name) at (your/his/her) **main** job?
   - Yes (ask next question)
   - No, Don't know, Refused (skip to independent industry/occupation questions)
2. "*CHDUTY*". Have the usual activities and duties of (your/his/her) job changed since last month?
   - Yes (skip to independent industry/occupation questions)
   - No, Don't know, Refused (ask next question)
3. "*SAMEACT*". Last month (name/you) (was/were) reported as (a/an) (input occupation) and (your/his/her) usual activities were (input duties 1) (input duties 2). Is this an accurate description of (your/his/her) current job?
   - Yes, Don't know, Refused (end series and use dependent coding)
   - No (skip to independent industry/occupation questions)

and if dependent coding was used, then the same occupation as in the previous month was automatically assigned. So the pre-1994 "independent coding" questions were asked after 1994 only when the interviewer could be confident that there had been a genuine change of activity. In addition, when independent coding was used after 1994, a direct question—"That is, what

Table 2. *Occupational transitions, by validity status, 1994–2006*

		All records	Dependent coding answers		
				Suspicious	
Occupational		Raw	Valid	Raw	*ANY3*-adjusted
Stayers	Frequency	511,807	499,425	12,382	16,375
	Percent	95.0	97.2	49.8	65.9
Movers	Frequency	27,068	14,611	12,457	8,464
	Percent	5.0	2.8	50.2	34.1
Total	Frequency	538,875	514,036	24,839	24,839
	Percent	100	100	100	100

is your occupation?"—was added. For this reason, we refer to pre-1994 as "unconditional independent coding", and to post-1994 as "conditional independent coding" of different occupations, where in the latter case the coding is conditioned on received information on an ascertained change of employer and/or activity and task.

Of all men matched and employed in months in samples 2 and 3, about 5% on average change occupation each month after 1994 (see Table 2). This is the result of raw data. Before 1994, raw occupational mobility is 33.6% per month, clearly an unreasonable value. Hence, we rely on post-1994 data to provide a benchmark measure of occupational transitions, that we believe to be reasonably accurate for two main reasons. First, the main goal of dependent coding is precisely to make sure that occupations are coded independently *if and only if* the occupation has in fact changed with very high probability. On the other hand, inaccurate responses are very unlikely to produce a false negative (no mobility, when in fact the occupation did change). A validation study by Polivka and Rothgeb (1993) provides hard evidence for these arguments: (Conditionally) independent coding implied a true mobility in the acceptable range. The worrisome validation studies, such as Mathiowetz (1992), concern unconditional independent coding of occupations, naturally much noisier but also not a concern for us after 1994.

These considerations still leave the door open for some miscoding post-1994, that we take seriously. The main purpose of this paper is to design a "cleaning" algorithm of spurious occupational transitions, that builds on the limited longitudinal dimension of the data to replicate, and refines the logic of, dependent coding. The next section illustrates our two procedures and shows that, once applied to post-1994 records, they eliminate most suspicious conditionally independently coded transitions, leaving us with a fairly homogeneous and clean sample.

III. Identifying Occupational Mobility with Matched CPS Data

In addition to exploiting dependent coding after 1994, our approach to distinguish valid from spurious changes in occupational status has two parts. First, we use other variables whose difference over the two months is likely to be correlated with changes in occupation, in order to validate possible occupational mobility. Second, we use the structure of each individual's sequence of (maximally four) consecutively reported occupations to identify sequences that appear "implausible" based on the most reliable post-1994 data. From now on, *MOB* refers to a dummy variable indicating whether two occupation codes for the same individual differ in months in samples 2 and 3, which would be a perfect measure of occupational shifts in the case of no misclassifications.

A First Pass: The ANY3 Filter

Our first check exploits variables that are likely to be correlated with the occupation (*OCC*) variable. If these variables change between two months for which the *OCC* codes change, the change in *OCC* codes is likely to represent true occupational mobility. We have identified the following variables, that are available both before and after 1994 and possibly correlated with changes in occupation: class of worker (*CLASS*: private firm, federal, state or local government, or self-employed), three-digit industry code (*IND*), looked for work in past four weeks (*LK*/*LKWK*), hours of work (*HOURS*), full-time or part-time work status (*WKSTAT*), family income (*FAMINC*).

The correlation between these variables and *MOB* reveals that in the post-1994 period (in which we believe the *OCC*/*MOB* data to be closer to the truth) changes in *CLASS* and *IND* happen almost only when $MOB = 1$, and the correlation between *LK*/*LKWK* and *MOB* is also very strong (though of course not perfect). Conversely, changes in *HOURS*, *WKSTAT* and *FAMINC* correlate poorly with *MOB*. In the pre-1994 data these correlations maintain the same ranking but are all lower, as we would expect since the pre-1994 occupation data is of lower quality. In the text below, "*ANY3*" refers to the criterion which preserves an observed transition with a suspicious flag if *any* of *CLASS*, *IND* and *LK*/*LKWK* has also changed. Conversely, we consider spurious any change of occupation which satisfies simultaneously three criteria: no change of industry, no change of class of worker, no active job search in the preceding month. In that case, we reset *MOB* to zero. Since three mobility criteria have to be simultaneously satisfied for this resetting, ours appears to be a conservative approach. Yet, it has a significant bite on the raw data. As we can immediately infer from our tables, the *ANY3* filter alone reduces mobility by two-thirds before 1994 and by one-third after 1994. These facts confirm both that conditional independent coding is vastly more reliable than the pre-1994 unconditional independent coding, and that applying this additional filter also to post-1994 data is appropriate.

Exploiting the Longitudinal Dimension: The FLAG Filter

Our main filter exploits the limited but still rich longitudinal dimension of the monthly CPS. For this filter to apply, we focus on a transition between two months t and $t+1$ such that potentially we also have information at least about months $t-1$ and $t+2$. This allows us to identify the "trajectory" of three consecutive occupational transitions, i.e., to analyze them jointly rather than in isolation. This implies that we need information about at least four consecutive months, which is exactly what the CPS contains, and focus on the transition between months 2 and 3. Although we could also use information on months in samples 5–8, we focus on the first four months to minimize sample attrition and the resulting selection. In order to use each individual's sequence of (four) consecutive observations on occupations, we have to identify all different possible cases. We are mainly after misclassified mobility, since the problem due to miscoded non-mobility is likely to be minor.

For records with an occupational transition, $MOB = 1$, we now introduce a variable *FLAG*, which takes on different values to indicate the type of transition. To illustrate, assume that A, B, C and D represent different three-digit occupations, and N stands for the case where occupation information is not available. Consider the possible sequences of codes in the first four months in the sample, for an individual who has valid and different codes in months 2 and 3 (a condition that defines $MOB = 1$). The first sequence, A-A-B-B, is the "ideal case". As another example, the sequence A-A-B-C is plausible, based on the typical decline in the hazard rate of separation with tenure, as in Farber (1994), and predicted by the canonical theory of worker turnover in Jovanovic (1979). The first transition from A to B in the third month reveals a likely mismatch in occupation A and triggers a job-shopping process, which may not be immediately successful, and may lead the worker to keep trying again the next month with occupation C. For the same reason, the sequence A-A-B-A is unlikely, although of course not impossible, as a worker may realize soon that his old A job is preferred to the new career B he tried.

The *FLAG* variable takes on values from 1 to 13, as detailed in Table 3. Although other sequences are logically possible, these 13 sequences are exhaustive in our sample. The fact that some of these occupation sequences are more likely than others to represent true occupational mobility forms the basis of our approach for cleaning the mobility data. The

Table 3. *FLAG distribution, 1994–2006*

				Suspicious				All	
		Valid		Not cleaned		Cleaned by *ANY3* only		Valid, cleaned by *ANY3* and *FLAG*	
FLAG number	Type	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	A-A-B-B	8,307	56.85	814	6.53	570	6.73	8,877	47.70
2	N-A-B-B	1,277	8.74	858	6.89	581	6.86	1,858	9.98
3	A-A-B-N	833	5.70	4,140	33.23	*2,663*	31.46	833	4.48
4	A-A-B-A	405	2.77	413	3.32	338	3.99	743	3.99
5	A-B-A-A	436	2.98	739	5.93	545	6.44	981	5.27
6	A-A-B-C	942	6.45	471	3.78	363	4.29	1,305	7.01
7	A-B-C-C	1,252	8.57	1,219	9.79	836	9.88	2,088	11.22
8	A-B-C-D	395	2.70	893	7.17	539	6.37	934	5.02
9	N-A-B-C	193	1.32	336	2.70	226	2.67	419	2.25
10	A-B-C-N	204	1.40	840	6.74	*610*	7.21	204	1.10
11	N-A-B-N	260	1.78	1,356	10.89	*933*	11.02	260	1.40
12	N-A-B-A	61	0.42	136	1.09	*93*	1.10	61	0.33
13	A-B-A-N	46	0.31	242	1.94	*167*	1.97	46	0.25
Total occup. movers		14,611	100	12,457	100	8,464	100	18,609	100

distribution of *FLAG* for the post-1994 transitions, cleaned according to the *ANY3* (described above) and dependent coding filters, provides us with a benchmark to evaluate the plausibility of a sequence, thus to decide whether to eliminate that kind of transition or not before 1994 or even after 1994 when the dependent coding questions have missing or suspicious answers.

IV. Results

Occupational Mobility after 1994

We first tackle the potential miscoding problem that might remain after 1994 due to conditionally independent coding of some occupations. We define a record "Suspicious" if any of the following three events is true: either the answer to the first dependent coding question "same job?" is Blank, or the answer to any of the subsequent two questions (conditional on the series arriving there with previous valid answers), "change of duty?" and "same activity?", is Blank.[11] All other records are "Valid". Our logic is as follows. A Blank answer to the first question triggers independent coding for unknown reasons, i.e., it is effectively unconditional independent coding as pre-1994, so it is suspicious. A No answer signals a different job, so a change of occupation is plausible. Of those who reply to have the same job (Yes to the first question), we are suspicious only of the Blank answers to the second question, because those who say that they have kept the same job but changed duty should be expected to have changed occupation, so the independently coded occupations should be different with very high probability. Same for the third question: suspicious are those who have a Blank answer to the "same activity?" question after having reported the same job and no change in duties.

In Table 2 we break down the sample in Valid and Suspicious records. Although Suspicious records are only $24{,}839/538{,}875 = 4.6\%$ of the total, their occupational mobility is 50.2%, relative to 2.8% of the Valid records. This evidence corroborates our prior beliefs that Suspicious records are truly suspicious and Valid records are more likely to be valid from the point of view of measuring occupational mobility. The mobility of Valid records is evidently due to independent coding of occupations that are likely to be different because of the valid answers to the dependent coding questions. This is the whole point of dependent coding, and as such it must be exploited and cannot be ignored. The Suspicious records are few in

[11] Here Blank also means "Don't know" or "Refusal", but the latter two categories are quantitatively negligible.

Table 4. *Results after cleaning, 1994–2006*

Occupational		All cleaned	By dependent coding answer Valid	 Suspicious cleaned
Stayers	Frequency	520,266	499,425	20,841
	Percent	96.55	97.16	83.90
Movers	Frequency	18,609	14,611	3,998
	Percent	3.45	2.84	16.10
Total	Frequency	538,875	514,036	24,839
	Percent	100	100	100

number but contribute a large amount to mobility, so we still have to deal with them.

We first pass post-1994 Suspicious transitions through our *ANY3* filter. As shown in Table 2, this cuts the proportion of Suspicious movers by a third, from 50.2% to 34.1%, already a significant improvement. Next, we pass the remaining Suspicious transitions through the *FLAG* filter. The benchmark is the flag distribution of the Valid records, which we assume to be immune from miscoding. This is clearly a strong assumption, but the best we can do given data limitations.Table 3 reports the flag distribution for Valid and Suspicious occupational movers. We italicize the frequencies of Suspicious flags that show a great discrepancy from the frequencies in the Valid case. Based on this evidence, all Suspicious transitions with flags 3 and 10, 11, 12 and 13 are eliminated and treated as spurious, that is, *MOB* is reset from 1 to 0 in those cases. Our decision is based on the fact that we do expect Suspicious records to be somewhat different from the Valid ones (there must be some reason why they are Suspicious), but we want to discard the cases where the discrepancy is very large. So we keep flags 8 and 9, although different frequencies appear, and discard 3 and 10–13, because the discrepancy is quite large. While discarding all of flag-3 Suspicious transitions may seem excessive (still 5.7% of Valid transitions have flag 3), so is keeping all of flags 8 and 9 Suspicious transitions, which have a frequency well over twice that among Valid transitions. While this filter calls for a subjective judgment, it is clear from the table that one would arrive at roughly the same number of final transitions with different reasonable criteria.

The result of applying the *ANY3* and *FLAG* filters after 1994, shown in Table 4, is that we retain only 3,998 of the original 12,457 Suspicious transitions, and correspondingly cut the mobility rate of the 24,839 Suspicious records from 50.2% (Table 2) to 16.1% (Table 4). Importantly, this reduction of roughly 34% in the measured mobility rate of Suspicious records is very close to our estimate of pre-1994 independent coding error for monthly CPS data, illustrated in the next subsection.

The final tabulation of cleaned transitions post-1994 is in Table 4. Overall mobility is 3.45% per month. This is a weighted average of 2.84% of Valid transitions and 16.10% of Suspicious transitions cleaned by our *ANY3* and *FLAG* filters. Given the relatively small number of initial and, even more so, surviving Spurious transitions, we are confident that true monthly occupational mobility after 1994 is close to our preferred point estimate of 3.45%.

Occupational Mobility before 1994

Before 1994, all occupations were coded independently anew every month. In our jargon, all of the records that before cleaning show up as occupational movers, a total number of 237,669 (out of 707,182 observations), are to be treated as Suspicious. Table 5 illustrates the distribution of flags for pre-1994 data. As mentioned earlier, raw occupational mobility is 33.6%, which is evidently too high at a monthly frequency and confirms our skepticism about pre-1994 observations. However, this immediately tells us that the miscoding error in the CPS cannot be as high as 50% as reported in Mathiowetz (1992), but is likely close to a relatively more manageable 30%.

Our *ANY3* filter eliminates almost two-thirds of all transitions, leaving us with a mobility rate of 12%. Next, we apply our *FLAG* filter by comparing the distribution of flags for the *ANY3*-cleaned pre-1994 data (Table 5) to our benchmark, which is now the distribution of post-1994 cleaned records (last column of Table 3). This leads us to *retain* before 1994 *only* transitions associated to flags 1, 2, 3, 6 and 7. Although this selection may appear drastic, notice that the retained flags account for three-fourths of all transitions in the post-1994 cleaned sample, and nonetheless too many transitions of types (flags) 6 and 7 survive before 1994. We also tried to either retain flags 4, 5, 8, 9 or 10 transitions as valid if and only if they satisfied the *ANY3* requirement, or to use the *ANY3* criterion directly, without using the sequence/*FLAG* structure at all. In both cases, final mobility was above 15% per month, unrealistically high.

The final occupational mobility rate in 1979–1993 is 3.59% per month, a drop of about 30 percentage points from 33.8% in the raw (uncleaned) data. Final mobility is very close to the 3.45% we found in post-1994 data, and the effect of our filters is virtually identical when applied to pre-1994 transitions and to post-1994 spurious transitions: in both cases, mobility falls by roughly 30%. Since the new interview method was introduced in January 1994, we cannot compute mobility in December 1993 (between that month and January 1994). But, if much residual error remained pre-1994, we would expect final (cleaned) mobility to drop significantly between November 1993 and January 1994. We find the opposite: mobility rises from 4.2% to 4.3% after seasonal adjustment. The same two numbers for

Table 5. *FLAG distribution, 1979–1993*

FLAG number	Type	Raw Frequency	Raw Percent	Final: Cleaned by *ANY3* only Frequency	Final: Cleaned by *ANY3* only Percent	Final: Cleaned by *ANY3* and *FLAG* Frequency	Final: Cleaned by *ANY3* and *FLAG* Percent
1	A-A-B-B	18,400	7.9	6,345	7.6	6,345	25.0
2	N-A-B-B	6,551	2.8	2,567	3.1	2,567	10.1
3	A-A-B-N	7,051	3.0	3,106	3.7	3,106	12.2
4	A-A-B-A	31,918	13.6	*9,820*	11.7		
5	A-B-A-A	33,279	14.2	*10,240*	12.3		
6	A-A-B-C	16,380	7.0	6,293	7.5	6,293	24.8
7	A-B-C-C	18,844	8.1	7,048	8.4	7,048	27.8
8	A-B-C-D	72,628	31.0	*25,696*	30.7		
9	N-A-B-C	6,358	2.7	*2,850*	3.4		
10	A-B-C-N	8,039	3.4	*3,875*	4.6		
11	N-A-B-N	4,252	1.8	*2,136*	2.6		
12	N-A-B-A	4,939	2.1	*1,663*	2.0		
13	A-B-A-N	5,297	2.3	*1,942*	2.3		
Total occup. movers		233,936	100	83,581	100	25,359	100

November 1992 and January 1993 are 3.6% and 3.4%, respectively, and for November 1994 and January 1995 they are 3.9% and 4.0%, respectively. The latter two are time periods which are very close to each other and feature no within-period change in interviewing techniques. So, the 0.1% increase in occupational mobility in 1993:11–1994:1 is nothing special or suspicious. Finally, there is no reason to believe that any sources of pre-1994 residual measurement error have any cyclical or trend component. We now turn to the time-series properties.

Time-series Patterns of Occupational Mobility

We can now plot the time series of aggregate occupational mobility one month forward for all male individuals aged 16–64 who were in their second month of the interview sequence, matched and employed both in the second and third months in sample.

Some caveats are in order. As mentioned earlier, data throughout the entire 1979–2006 period are not fully comparable for two main reasons. In 1983 there was a significant change in the Census three-digit occupational classification system, and in 1994 dependent coding was introduced. We cannot do much about the former, and we expect this change to raise measured mobility after 1983. We have tried to make pre-1994 data as similar to post-1994 observations as much as we can. At a minimum, we feel safe in analyzing the time-series behavior of mobility in the following three subperiods separately: 1979–1982, 1983–1993, 1994–2006.

Figure 2 illustrates the time series for the whole period and each subperiod, also after controlling for seasonality using X12A's Census methodology, with variable orthogonal seasonal factors. Over the entire 1979–2006

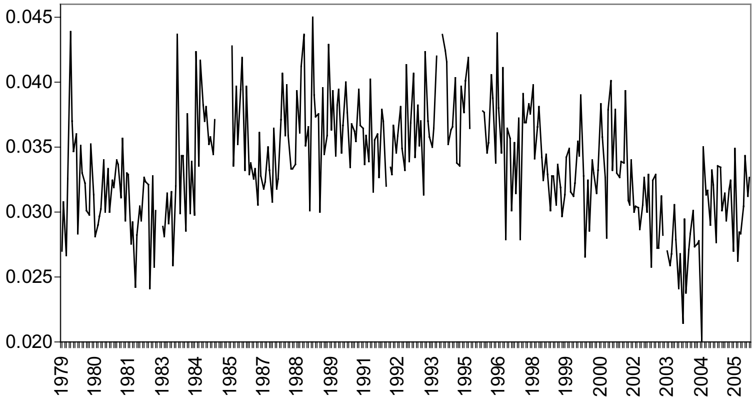

Fig. 2. Occupational mobility rate, cleaned and seasonally adjusted

period, occupational mobility tends to be procyclical. In the first subperiod, 1979–1982, there is a significant decline, which is easy to interpret as a cyclical effect. In the second subperiod, 1983–1993, there is a definite initial rebound from the 1982 recession, followed by a long period of relatively constant mobility, with a dip in the 1989–1991 recessionary period. Perhaps most strikingly, after 1994, the period for which data are cleanest, we observe a definite and large decline in the pace of occupational mobility, with a clear dip in the 2001 recession that triggers a significant continued fall in 2002–2003. The cyclical rebound only materializes in 2004, three years after the end of the recession.

Kambourov and Manovskii (2008) report similar time series of three-digit occupational mobility at annual frequency from the PSID in 1968–1997 and for a slightly different sample: men aged 23–61, heads of household, private sector, not self-employed. That is, they exclude (due to PSID data restrictions) non-heads and very young workers, as well as workers aged 62–65, and also exclude government employees because they found that the occupational mobility of those workers plummeted, masking an overall positive trend. The average annual occupational mobility that they report from the PSID is about 20% at the three-digit level. This is consistent with our 3.5% at the monthly frequency when allowing for repeated and correlated within-year transitions. Indeed, if every individual changed occupation independently every month with chance 3.5%, then annual mobility would be around 34%. The well-known correlation in transitions (the hazard rate of separations from jobs is sharply decreasing in tenure) can easily explain part of the discrepancy from the annual mobility of 20% in the PSID as the result of time aggregation. In addition, Kambourov and Manovskii (2008) consider all employed in each year t, and go back in time for each individual until they find employment and an occupation in years $t-1$, $t-2$, etc., whichever is latest. This allows for more time than the standard one-year interval for a transition to occur. We make a different, equally subjective choice by focusing on workers who are employed in consecutive months. The ideal but not objectively feasible procedure would be to impute a shadow occupation to the unemployed. The relative impact on measured mobility of these two respective approaches is ambiguous. Moreover, in the PSID, occupations are coded according to a unique, coherent, and by the end of the period somewhat obsolete, Census coding system, so it is likely to underestimate mobility as time goes by because of the lack of distinction between new occupations, such as software developers, that emerged in the last three decades of the twentieth century.

In terms of time-series patterns, the PSID shows a secular increase in three-digit occupational mobility, more pronounced in the first half of the sample, before 1986, than in 1987–1997, although the annual data are fairly noisy, with frequent year-to-year variations of over ±2%; see Kambourov and Manovskii (2008). The procyclical pattern is mild but visible, confirming our earlier finding. We rederived our time series using the same sample disposition as Kambourov and Manovskii (2008), again with the proviso of the three subperiods. The results with this sample were qualitatively unchanged from the findings derived with our baseline sample, as presented in Figure 2. Hence, our results agree with Kambourov and Manovskii's for the common few years in our sample period: we too find an increase in mobility from the late 1970s to the mid-1990s. However, after the end of their sample in 1997, occupational mobility reversed course and declined dramatically, falling below 1970s levels, with just a modest cyclical rebound in 2004–2006.

Imputed Occupations

Some values of the occupational codes are imputed, rather than assigned following either dependent coding or the current description, and are therefore likely to contain errors. The CPS Monthly files contain a variable, *AOCC*, that indicates whether the occupation code was imputed or not.[12] Imputation is likely to lead to miscoding and to overestimate occupational mobility. But discarding imputed records could lead to a non-representative sample, if some groups are overrepresented among the persons with imputed values. Clearly, this could be a potential issue with the results presented above.

Table 6. *Records with imputed occupations and cleaned mobility status*

1979–1983		1984–1988		1989–1993		1994–2006	
Imputed	Not imputed	Imputed	Not imputed	Imputed	Not imputed	Imputed	Not imputed
N/A	230,479	1,618	214,119	1,953	223,327	6,990	513,276
	100.00	0.75	99.25	0.87	99.13	1.34	98.66
N/A	7,979	198	10,984	198	8,755	2,236	16,373
	100.00	1.77	98.23	2.21	97.79	12.02	87.98
N/A	238,458	1,816	225,103	2,151	232,082	9,226	529,649

[12] The CPS does allocations for two reasons: missing information, or inconsistent information. With regards to occupation data, the Census has certain requirements for consistency of the occupation, industry and class of worker as a whole. Prior to 1984, no information about allocation is available. In 1984–1988, a binary allocation variable is set to one if either of these operations is used. After 1988, the Census expanded the range of allocation variable values to capture the many nuances of the allocation process. In general, values of 2–5 in the 1989–1993 time frame and values of 10–43 for 1994 to current are the same as the 1984–1988 variable value of one. We thank Gregory Weyland of the US Census Bureau for providing this information.

Table 7. *Impact of cleaning procedure on one-digit and three-digit occupational mobility*

	Raw	Final (cleaned)	Raw minus final
1983–1986			
One-digit	22.18	4.18	18.00
Three-digit	34.49	3.66	30.83
2004–2005			
One-digit	3.42	2.18	1.25
Three-digit	4.98	3.18	1.83

Fortunately, this turns out not to be the case. In Table 6 we tabulate the imputed and non-imputed records in our sample "cleaned" by the *ANY3* and *FLAG* procedures. The number of imputed occupations constitutes a very small share of all observations with occupational changes, only rising to 24% after 1994. Although this last proportion is significant, it would not be wise to treat all of these transitions as spurious, and in fact our cleaning procedure should leave mostly imputed records that correspond to a genuine change of occupation, especially after 1994. Furthermore, few occupations are imputed in the first place. Even counting all occupational transitions associated with an imputed record as spurious, an extremely conservative solution, overall mobility would not change before 1994, and would drop from 3.45% to 3.10% after 1994. Therefore, we feel much safer in keeping the remaining records with allocated values after 1994 in the sample as they are.

A Final Check: One-digit Occupational Mobility

In order to further validate our cleaning procedure of occupational mobility, we perform one final test. We apply the same procedures to mobility across one-digit occupations. This is a classification of occupations in about (depending on the time period) 15 to 25 broad categories, such as: Sales; Farming, forestry, and fishing; etc. The one-digit classification system is derived by the Census from the original survey questions, which are meant and coded to capture three-digit occupations. Intuitively, both the level of mobility and miscoding should be smaller at this coarser level of occupational classification. Thus, we should expect our filters to have a larger effect on three-digit than on one-digit mobility, as the former is more finicky and likely to be miscoded.[13]

In Table 7 we report mobility, raw and adjusted by *ANY3* and *FLAG*, as well as the total amount of adjustment, for two subperiods, 1983–1986 and

[13] We thank an anonymous referee for suggesting this check.

Table 8. *Answers to SAMEJOB question by final occupational mobility, 1994–2006*

Occupation/*SAMEJOB*	Refuse	Don't know	Blank	Yes	No	Total
Stayers	364	350	15,566	498,879	5,107	520,266
Percent	0.07	0.07	2.99	95.89	0.98	100
Movers	1	0	2,940	6,547	9,121	18,609
Percent	0.01	0.00	15.80	35.18	49.01	100
Total	365	350	18,506	505,426	14,228	538,875
Percent	0.08	0.06	3.43	93.79	2.64	100

2004–2005. We include more years before 1994 because the raw data are noisier. The results are encouraging, or at least not a cause for concern, as the amount of spurious mobility intercepted by our filters at the three-digit level is over 1.5 times that at the one-digit level.

V. Job-to-Job Mobility, 1994–2006

Another important type of labor market transition is that from job to job. This is a flow of central interest to the calibration and estimation of equilibrium search models of the labor market. The monthly CPS after 1994 contains a *SAMEJOB* question which makes it ideally suited to measure employer-to-employer (EE) flows, although we still face time aggregation of multiple within-month EE transitions and short unemployment spells. This new source of information has been exploited recently by Fallick and Fleischman (2004) and Nagypal (2004). An EE transition occurs when the question *SAMEJOB* the following month receives a negative answer. They report a monthly mobility rate in 1994–2004 of about 2.7% on average. Neither of these studies report their treatment of Blank answers to the *SAMEJOB* question. This is a potentially serious issue because, as reported in Table 8, these records amount to 3.43% of our sample. If we simply ignore Blanks answers, in 1994–2006 we obtain the same 2.7% mobility rate as the other authors, although they focus on all rotation groups that are selected by attrition and on slightly different time periods. Imputing all Blanks to either EE non-movers (Blank → Yes) or movers (Blank → No) makes the total EE rate range from, resp., 2.6% to 6.0%. Without additional information, however, it would be difficult to interpret this very large number of Blank answers to the key *SAMEJOB* question.

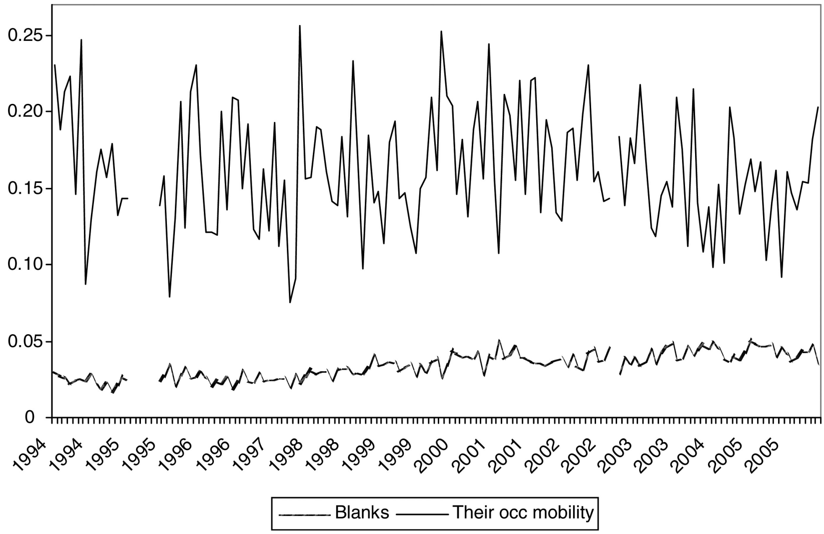

Fig. 3. Proportions of Blank answers to *SAMEJOB* question, and cleaned occupational mobility of those records

Interpretation of Missing Answers

Our analysis of occupational mobility provides some guidance to better interpret the meaning of a Blank answer to the first dependent coding question (*SAMEJOB*) in month 3, whether the individual works for the same company as last month or not. From Table 8, occupational transitions follow 2,940/18,609 = 15.8% of the time a Blank answer to the *SAMEJOB* question, half the time a No answer (employer/ company definitely changed), and the rest of the time a Yes answer (same employer). The 2,940 occupational movers with a Blank *SAMEJOB* answer are a subset of the Suspicious occupational transitions, those that survived our *ANY3* and *FLAG* filters.

Since Blank answers to the *SAMEJOB* question are so numerous and important among occupational movers, they must be investigated further. From Table 8, cleaned occupational mobility conditional on a Blank answer to *SAMEJOB* in month 3 is about 16% (2,940/18,596). Figure 3 shows that this rate has been fairly stable over the period, while the fraction of Blank answers to the *SAMEJOB* question has been increasing in 1994–2006, for reasons that we cannot fathom. The average occupational mobility of the ascertained job movers (*SAMEJOB* = No) is 64%, while that of ascertained job stayers (*SAMEJOB* = Yes) is 1.3%. The "Refusal" and "Don't know" answers to the *SAMEJOB* question are negligible in number. This leads us to interpret Blank answers as having relatively high job mobility. Our *ANY3*

Table 9. *Joint distribution of job-to-job and occupational mobility, 1994–2006*

Occupation/Job	Stayers	Movers	Total
Stayers	498,879	5,107	503,986
Percent	96.00	0.98	96.22
Movers	6,547	9,121	15,668
Percent	1.26	1.76	3.78
Total	505,426	14,228	519,654
Percent	97.26	2.74	100.00

and *FLAG* filters preserve many Blank *SAMEJOB* records as valid for occupational transition purposes, so excluding all of these records from the computation of job-to-job transitions is inappropriate. Since Blank answers to the *SAMEJOB* question are associated to occupational mobility that exceeds both the sample average and that of job stayers, we conclude that the 2.7% job-to-job mobility rate reported in the literature by ignoring Blank answers is likely to underestimate somewhat the true value.

Results

First, in Table 9, we address the correlation between job and occupational mobility. We focus on records with valid *SAMEJOB* answers. Interestingly, job and occupational mobility are only partially correlated. Almost 40% of the occupational movers stay with the same employer, while over 33% of job-to-job movers keep the same occupation. Neal (1999) shows, although in the severely limited (in size, time and representativeness) NLSY data, that young workers follow a two-stage search strategy: first they search for a career (occupation), and then shop for jobs within the chosen career.

Next, we turn to the impact of Blank answers to the *SAMEJOB* question on the measured EE rate. Clearly, the high occupational mobility and the increasing incidence of Blank answers suggests that ignoring them would lead to underestimating the EE rate, especially late in the sample. Figure 4 reports the time series of the EE rate. First, focus on the unadjusted rate, which excludes Blank records altogether, the common practice in the literature. This rate is procyclical, but with very persistent negative effects of a recession. The rate is falling in 1994–1996, rising in 1997–2000, falling sharply after the 2001 recession, and rebounding only after 2004. The average is, as mentioned, 2.7%.

Following the discussion above, we now allocate Blank answers to Yes and No according to their relative frequencies of cleaned occupational mobility, month by month. If Blanks are a random sample of the population

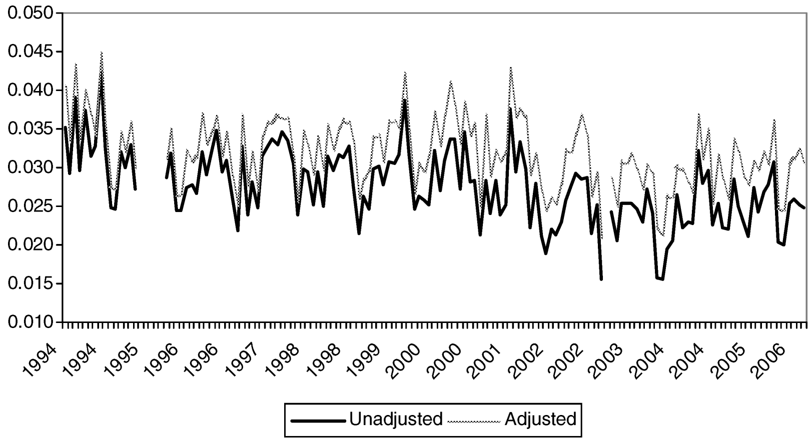

Fig. 4. Employer-to-employer (EE) monthly transition rate, unadjusted and adjusted for missing answers to *SAMEJOB* question

for job mobility purposes, then their occupational mobility rate should be the weighted average of those of the Yes and No, in which case the overall job mobility rate should not change. The closer the occupational mobility rate of the Blank to that of the No, the higher the adjusted job mobility rate should be. Indeed, the occupational mobility of the Blank and No are essentially the same. Therefore, each month we compute the following job-to-job mobility rate:

$$\text{Adjusted EE rate} = \frac{No + BlankOccmovers \cdot \delta}{Yes + No + BlankOccmovers \cdot \delta},$$

where

$$\delta \equiv \frac{\Pr(occmob \mid SAMEJOB = No) - \Pr(occmob \mid SAMEJOB = Yes)}{\Pr(occmob \mid SAMEJOB = No) + \Pr(occmob \mid SAMEJOB = Yes)},$$

and *No*, *Yes* and *BlankOccmovers* are the numbers of answers to the *SAMEJOB* question. All magnitudes are evaluated anew each month. If the occupational mobility rate was the same for all three groups, then Blanks would be a random sample of the population, $\delta = 0$, and the EE rate would be the same counting Blanks or not. If $\Pr(occmob \mid SAMEJOB = No) > \Pr(occmob \mid SAMEJOB = Yes) = 0$, then $\delta = 1$ and all Blanks who change occupation would be counted as job movers.

Figure 4 also shows the adjusted EE rate. The average is nearly 3.2%, significantly larger than the 2.7% average unadjusted EE rate. The discrepancy rises over time with the proportion of Blank records. The cyclical behavior is similar.

VI. Conclusions

In this paper, we outline a new approach to measure and to study worker mobility across occupations and jobs in US Census data at the monthly frequency. The main focus of the paper is on occupational mobility. In addition, we build on our findings on occupational mobility to shed light on employer-to-employer transitions. We exploit the limited longitudinal dimension of monthly CPS data, combined with the post-1994 dependent coding technique, to cleanse the very large spurious labor market flows that appear in the raw data. We find that occupational transitions at the three-digit level (what most people think of as an "occupation") average at about 3.5% of employment per month. This is about the same as, and not so closely correlated with, our own revised measure of employer-to-employer transitions. Occupational mobility is procyclical, and shows a secular rise from the early 1980s to the early 1990s, and a secular decline from the mid-1990s onward, accelerating towards the end of our sample in 2001–2004, with a cyclical rebound only after 2004, a full three years after the end of the last recession. Job-to-job mobility also shows a similar pattern since 1994, when we can measure it.

At this stage, we can only speculate on the sources of this recent prolonged decline in career mobility. Given the well-known negative correlation between age and any measure of labor market mobility, the aging of the US labor force is likely to have contributed to the decline. However, this aging of the population was a factor also in the 1980s, hence it cannot by itself account for the decline in mobility after the mid-1990s. An alternative explanation focuses on the effects of outsourcing of specific tasks of production to low-income countries. On the labor demand side, a reduction in the number of viable careers and the resulting concentration of the labor force into the service sector and more skilled occupations mechanically lead to a lower flow. On the labor supply side, an increase in *perceived* uncertainty might make individuals more defensive in their career-shopping behavior, leading to a decrease in mobility. At this point, we are unable to assess the relative importance of these and other possible hypotheses, but we think it is an interesting direction for future research.

Throughout the paper we have discussed the limitations of our approach. Here is another, somewhat more fundamental limitation. Because we are entirely focused on eliminating false transitions, we are not really interested in whether occupations have the correct code to begin with, but only that the numerical code assigned to an occupation changes if and only if the true underlying occupation does. That is, our cleaning procedure yields what we believe to be reliable estimates of transition rates, but not necessarily of the *stocks* of employed individuals in the different occupations. This prevents us, for example, from presenting reliable estimates of net employment reallocation across occupations. This is a topic of future research, and a goal still on our agenda.

Despite their limitations, we believe that the long, complete and very high-frequency estimates of labor market flows that we obtain from the premier and richest survey of the US labor market may be useful to a variety of goals. Natural examples are the calibration of equilibrium search models to explain aggregate employment fluctuations, the structural estimation of search, sorting and turnover models, and the estimation of the wage/mobility link in a regression context. In addition, from a microeconometric viewpoint, the wealth of information available in the CPS opens a wide range of possibilities.

References

Fallick, B. and Fleischman, C. A. (2004), Employer-to-Employer Flows in the US Labor Market: The Complete Picture of Gross Worker Flows, Working Paper no. 2004-34, Federal Reserve Board, Washington, DC.

Farber, H. (1994), The Analysis of Interfirm Worker Mobility, *Journal of Labor Economics 12*, 554–593.

Feng, S. (2001), The Longitudinal Matching of Current Population Surveys: A Proposed Algorithm, *Journal of Economic and Social Measurement 21*, 71–91.

Ferrie, J. (2005), History Lessons: The End of American Exceptionalism, *Journal of Economic Perspectives 19*, 199–215.

Jovanovic, B. (1979), Job Matching and the Theory of Turnover, *Journal of Political Economy 87*, 972–990.

Kambourov, G. and Manovskii, I. (2002), Occupational Specificity of Human Capital, mimeo, University of Western Ontario, London, ONT.

Kambourov, G. and Manovskii, I. (2004), A Cautionary Note on Using (March) CPS Data to Study Worker Mobility, manuscript, University of Pennsylvania, Philadelphia.

Kambourov, G. and Manovskii, I. (2008), Rising Occupational and Industry Mobility in the United States: 1968–1997, *International Economic Review 49*, 41–79.

Madrian, C. and Lefgren, L. J. (2000), An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents, *Journal of Economic and Social Measurement 26*, 31–62.

Mathiowetz, N. A. (1992), Errors in Reports of Occupations, *Public Opinion Quarterly 56*, 352–355.

McCall, B. (1990), Occupational Matching: A Test of Sorts, *Journal of Political Economy 98*, 45–69.

Moscarini, G. and Thomsson, K. (2006), Occupational and Job Mobility in the US, Yale Working Papers on Economic Applications and Policy, Discussion Paper no. 19, New Haven, CT.

Murphy, K. and Topel, R. (1987), Unemployment, Risk and Earnings: Testing for Equalizing Wage Differentials in the Labor Market, in K. Lang and J. Leonard (eds.), *Unemployment and the Structure of the Labor Market*, Basil Blackwell, New York.

Nagypal, E. (2004), Worker Reallocation over the Business Cycle: The Importance of Job-to-Job Transition, manuscript, Northwestern University, Evanston, IL.

Neal, D. (1999), The Complexity of Job Mobility among Young Men, *Journal of Labor Economics 17*, 237–261.

Peracchi, F. and Welch, F. (1994), Trends in the Labor Force Transitions of Older Men and Women, *Journal of Labor Economics 12*, 210–242.

Polivka, A. E. and Rothgeb, J. M. (1993), Redesigning the CPS Questionnaire, *Monthly Labor Review 116*, 10–28.

Schumpeter, J. (1939), *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process,* McGraw-Hill, New York.

Welch, F. (1993), Matching the Current Population Surveys, *Stata Technical Bulletin 12*, 7–11.