ESTIMATION OF JOB-TO-JOB FLOW RATES UNDER PARTIALLY MISSING GEOGRAPHY *

by

Cody Henderson *
University of Texas

and

Henry Hyatt *
U.S. Bureau of the Census

CES 12-29 September, 2012

The research program of the Center for Economic Studies (CES) produces a wide range of economic analyses to improve the statistical programs of the U.S. Census Bureau. Many of these analyses take the form of CES research papers. The papers have not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Republication in whole or part must be cleared with the authors.

To obtain information about the series, see www.census.gov/ces or contact C.J. Krizan, Editor, Discussion Papers, U.S. Census Bureau, Center for Economic Studies 2K130F, 4600 Silver Hill Road, Washington, DC 20233, CES.Papers.List@census.gov.

Abstract

Integration of data from different regions presents challenges for the calculation of entity-level longitudinal statistics with a strong geographic component: for example, movements between employers, migration, business dynamics, and health statistics. In this paper, we consider the estimation of worker-level employment statistics when the geographies (in our application, US states) over which such measures are defined are partially missing. We focus on the recent pilot set of job-to-job flow statistics produced by the US Census Bureau's Longitudinal Employer- Household Dynamics (LEHD) program, which measure the frequency of worker movements between jobs and into and out of nonemployment. LEHD's coverage of the labor force gradually increases during the 1990s and 2000s because some states have a longer time series than others, so employment transitions involving missing states are only partially or not at all observed. We propose and implement a method for estimating national-level job-to-job flow statistics that involves dropping observed states to recover the relationship between missing states and directly tabulated job-to-job flow rates. Using the estimated relationship between the observable characteristics of the missing states and changes in the employment measures, we provide estimates of the rates of job-to-job, and job-to-nonemployment, job-to-nonemployment-to-job flows were all states uniformly available.

*Department of Statistics, University of Texas, San Antonio and Center for Economic Studies, US Census Bureau, respectively. The research program of the Center for Economic Studies (CES) produces a wide range of economic analyses to improve the statistical programs of the US Census Bureau. Many of these analyses take the form of CES research papers. These papers have not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the Census Bureau or its staff. All results have been reviewed to ensure that no confidential information is disclosed. Republication in whole or part must be cleared with the authors. We would like to thank John Abowd, Erika McEntarfer, Liliana Sousa and the participants of the Census Bureau's SEHSD Research Lunch series for helpful comments and suggestions. All mistakes are, of course, our own.

Introduction

Calculation of statistics on entity-level outcomes is simplest when the geographies over which the components of those outcomes occur are completely observed. This paper deals with the more difficult case of incomplete geographic coverage, a common problem in aggregated administrative records sources, and its implications for statistics where the components of an entity's outcome can span across different geographies: such as worker movements from one job to another or person movements from one region to another. For example, in statistics derived from Unemployment Insurance payroll tax records of US states, it is quite common for a worker to exhibit an employment separation and then not be employed in any available geography. These separations can in fact be movements to a job in a state that is not part of the geographic frame, a problem noted in such studies as Jacobson, Lalonde and Sullivan (1993) and von Wachter et al. (2010).

In this study, we propose and implement a method for using observable characteristics of missing states to obtain better estimates of person-level employment statistics. This approach involves calculating a statistic using all available data and re-estimating it with certain geographies excluded. Observable characteristics of the excluded geographies can then be used to explain the difference between the statistics calculated without exclusion and those calculated with excluded geography. Assuming that the observable characteristics of available geographies' relationship with the statistics is similar to that of the unavailable geographies, it is possible to obtain an estimate of the statistic were all geography available.

Our application addresses a particular type of person-level statistic: the job-to-job flow measures considered in Hyatt and McEntarfer (2012a, 2012b). Their statistics employ data from an *almost* national dataset of Unemployment Insurance wage records that the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program maintains as part of the Local Employment Dynamics partnership. We consider the production of job-to-job flow statistics for the year 1998 onwards, a year by which coverage of the employed population was almost national. Nevertheless, Massachusetts has at the time of this writing not submitted data to the LEHD program, and Connecticut, New Hampshire and the District of Columbia were not available in the dataset used for this analysis. Furthermore, ten other states' LEHD time series begin after the first quarter of 1998, and these states continue to enter the data frame until 2003.

Missing data presents problems when calculating job-to-job flow statistics, as the US labor force is quite mobile and workers frequently move across states as they change jobs, see Molloy, Smith, and Wozniak (2011). In this study, we use characteristics of missing states from the decennial Census and the Quarterly Census of Employment and Wages to adjust the raw tabulations by the amount implied by these observable characteristics.

This paper is most closely related to Abowd and Vilhuber (2011), who address the problem of calculating national-level statistics when only a subset of states are available, and who use publicly available LEHD data. Specifically, they construct national-level estimates of a subset of the LEHD program's Quarterly Workforce Indicators: job creation, job destruction, separations, and hires. All the indicators they consider are job-specific, that is, concerning combinations of a single worker with a single establishment. Missing job-level data is a fundamentally different (and much simpler) problem than the one we address in this paper: because all jobs have a specific geography, the establishment's location, they can impute the employment outcomes for each missing geography as a missing data problem as in Rubin (1987).[2]

Job-to-job flow rates must be transformed somewhat before they can be treated as a missing data problem. A job-to-job flow naturally involves the location of both the origin job and the destination job, and those locations are not always near each other. For example, when an employment separation in an observed state is contemporaneous with a hire in an unobserved state, direct tabulation of the underlying microdata will indicate a separation into nonemployment rather than a direct job-to-job flow. To account for partially or fully unobserved events, we exploit our ability to omit certain observed geographies to obtain some understanding of how missing geography affects tabulations.

Our approach involves first calculating job-to-job flow rates directly using all available data, then dropping a state and re-estimating the rates, and calculating the difference between the two. This exercise is repeated until we have calculated the effect of excluding each of thirty-eight states. We then estimate the relationship between these differences and observable

[2] Because this distinction is important to note but is not the focus of this paper, we provide a detailed mathematical treatment of how the employment measures considered in this paper differ from those of Abowd and Vilhuber (2011) in the Appendix.

characteristics available for all states from other data sources, specifically the Quarterly Census of Employment and Wages and the Census 2000.

Our results are as follows. Naturally, when more states are missing from the data, our estimated values are more different from the direct tabulations. We find that the quarterly frequency of direct job-to-job flows (that involve little or no nonemployment), which is on the order of ten percent, is initially lower in our direct tabulations than our estimates by slightly less than one half of a percentage point, and this difference almost disappears as more states enter the data. Flows involving more than one quarter are slightly underestimated, but always within one three hundredths of one percentage point of the direct tabulation. Flows into longer (two or more full quarters) of nonemployment are overestimated by less than two tenths of one percentage point, and the degree of underestimation of direct job-to-job flows is more than twice as large than the overestimation of flows to nonemployment. These results are implications of employment activity in missing states: workers appear to have no employment transitions when changing jobs in missing states, and also separations into and hires from nonemployment are, with some frequency, the product of actual inter-state job-to-job flows.

Although this paper is limited to employment transitions, our approach can be applied to a broader class of statistical problems. In the case of migration, movements within and between missing geographies are unobserved. Estimates of the frequency of migration, or person movements from one location to another, can be structured quite similarly to our employment location transition analysis. Statistics on business dynamics that have a strong entity-level component, such as transfers and within-firm job reallocation, can be addressed using a similar framework. Cumulative patient-level health outcomes are another category of statistics that could be estimated using this method. More broadly, our method should be able to be adapted for the calculation of any entity-level level statistics that where the outcomes have components that span geographies and the underlying data come from different regions.

The paper proceeds as follows. We begin with a description of our data source, including a discussion of the extent and nature of the missing geography. Thereafter, we provide some statistics on the frequency of interstate migration. Then, we introduce our employment concepts, discuss how they can be impacted by missing geographies, and present the frequency with which our job-to-job flow measures are within and across states. Then, we propose and implement a

method for correcting job-to-job flow statistics using the observable characteristics of missing states.

Data

The statistical problem that motivates our analysis is a new prototype database of job-to-job flows, which was developed recently by the LEHD program and which has been analyzed in recent work by Hyatt and McEntarfer (2012a, 2012b). The data used to calculate job-to-job flows come from the LEHD infrastructure files, which are largely derived from Unemployment Insurance wage data and the Quarterly Census of Employment and Wages, which states provide to LEHD through the Local Employment Dynamics federal-state partnership. Included in the LEHD infrastructure files are the wages paid to the employee, employee demographic characteristics, and employer characteristics such as industry and location. For additional information on these data, see Abowd et al. (2009).

The microdata we use in this study is a research snapshot of the LEHD infrastructure files.[3] This data has increasing coverage for the 1990s and 2000s as additional states began submitting data to the LEHD program as part of the Local Employment Dynamics federal-state partnership. Missing data by state and quarter is shown in Table 1 for the years 1998 onwards. The first entrance of an unobserved state into the data occurs in 1998:3 with Delaware, a relatively small state. We then see three more relatively small states enter: Iowa in 1998:4 and Utah and Nebraska both in 1999:1. Ohio, a larger state, enters the data in the first quarter of 2000, along with two small states Oklahoma and Vermont. The last the states to enter are Alabama, which enters in 2000:1, Arkansas in 2002:3, and Mississippi in 2003:3. We also see from Table 1 that Connecticut, Massachusetts, New Hampshire and the District of Columbia are never observed in the snapshot. Of these states, Massachusetts and Connecticut are the largest with 2000 Census populations of 6.3 million and 3.4 million respectively.

[3] Due to data infrastructure constraints, it was not feasible to conduct this analysis on the microdata used to construct the Hyatt and McEntarfer (2012a, 2012b) database of job-to-job flows, which was taken directly from LEHD's production database. Instead, we used the more commonly used s2008 snapshot, a research extract analogous to the s2004 snapshot described in McKinney and Vilhuber (2011). This later snapshot has been used in such papers as Fallick, Haltiwanger and McEntarfer (2011) and Andersson et al. (2011). The principle differences between the data used in this paper and those used by Hyatt and McEntarfer are that they include New Hampshire and the District of Columbia, and use data through 2010.

To make our application as relevant as possible to the LEHD program, we closely follow the method of aggregating data and developing job-to-job flow measures described in Hyatt and McEntarfer (2012a, 2012b). We focus on employees who have worked in the same set of what they term reference states: CA, FL, GA, IL, KS, MI, NV, NC, and ND. Specifically, the frame is any worker who has worked in any of the reference states at any point in the time series 1998:1-2008:4. We then construct national job histories for these workers so that flows from and to out of state jobs are included in their job history. Due to the large size of the data, over 120 million individuals and businesses, and the limited computing space, all calculations were performed on a one percent sample of the workers selected from the nine reference states, which nevertheless contains information on millions of individuals and businesses.

The focus on these nine reference states was done for several reasons. First, these states constitute a large fraction of the population of the United States: in the year 2000, these states had nearly ninety-four million residents, about a third of the population. Secondly, these states contain states that both did and did not experience the housing bust that lead to a precipitous decline in housing prices starting in 2006. Furthermore, these states do not have as many issues with cross-state employment as, for example, Virginia and Maryland, or many of the smaller states in the Northeast. It has been conventional wisdom (although, to the best of our knowledge, never subjected to quantitative analysis) that excluding such states will mitigate effects of interstate migration due to integrated labor markets that cross state boundaries. A few of these reference states, such as North Carolina, California and North Dakota do not share a border with any state that is missing from 1998 onwards, and all other states border at most one state that has missing data (the exception, Kansas, borders both Nebraska and Oklahoma). We also note that limiting the number of states substantially reduced processing time.

Descriptive Statistics

Before we present our job-to-job flow outcomes, in this section we provide information on the underlying employment process and its implications for calculating job-to-job flow statistics under partially missing geography. To begin, we assess the magnitude of the phenomenon of inter-state mobility and tabulate the frequency with which an employee works in

a given number of states over 1998:1-2008:2. As shown in Table 3, more than sixty percent of the individuals worked in only one state. A little over twenty-five percent of the sample worked in two states, almost nine percent worked in three states, and about five percent of the sample worked in four or more states. This indicates that a large percentage of workers work in more than one state over a span of ten years. Some of these workers may also work in states for which data is never available or in a state when its employment is unobserved earlier in the series.

New states provide both potential destinations for workers currently in the data frame as well as potential origins for workers that migrate into reference states. To provide a rough indication of the extent to which the missing data is mitigated as more states enter as part of the time series, Figure 1 shows quarterly measures of the frequency with which individuals are employed in each reference state, in another state, in a reference state and another state, or are nonemployed. This figure shows that as the time series progresses and more states enter the series, the proportion of workers classified as nonemployed decreases, reaching a series low of 41.3% in 2007:3. Note that half of the decline from an initial value of 51.4% occurs by early 2000 and corresponds with the entry of NE, UT and IA at the start of 1999 and OH, OK and VT in 2000. The proportions of workers in the reference states and in other states exhibit contemporaneous increases. Throughout the series, most of the employed workers work in the reference states, although this share declines as more states enter the series. At the beginning of the series (1998:1) when thirty-seven states have observable data, about 41% of observed workers work in the set of reference states only while about 7% work only in non-reference states. In 2001:1, when 44 states are in the data, about 44% work in only the reference states, while 9% work only in other states. At the end of the frame (2008:4), when 47 states are in the data, 45% work in only the reference states, while 10% were only in other states. Throughout the series, about 1% work in both reference and other states. Note that at the end of the series, nearly a fifth of the employed workers in our sample are working only outside the reference states. Given the early drop in nonemployment and the corresponding increase in work outside the reference states, we can say that as much as 3% of workers, or 6% of the employed population, is unobserved early in the time series.

Table 3 provides insight into the effects of adding subsets of states at different population levels to the original set of reference state on the proportions of flows at varying lengths of

nonemployment. This table presents calculations on the set of workers earning a positive wage in the reference states at 2000:1 that were earning no wages in the reference states at 2000:2, suggesting a separation during the first quarter of the year 2000 with at least one full quarter of nonemployment. When only observing the nine reference states, we see that roughly 23% of those separations did not result in a subsequent hire. The table also includes alternative calculations for sets of selected states who entered the LEHD frame as late as 2001:1. When states with a population of at least 8 million[4] are added to the reference states, the percentage of separations without an observed hire drops to about 20%, Adding states with a population of at least 3.5 million[5] population to the previous subset, drops the percentage of separations without an observed hire to roughly 16.4%, and adding the remaining fully observed states,[6] decreases the percent of flows with non-observed hires to 15%. This decline is more than accounted for by an increase in the number of flows that involve no quarters of nonemployment, that is, within-quarter and adjacent-quarter flows. This fraction was initially zero (by construction), but increases to 9.2% once the three sets of states are added. There is also a slight increase in the number of flows that involve exactly one or two full quarters of nonemployment, as well as slight decreases in flows with longer durations of nonemployment.

As an aside, note that Table 3 presents information on how to interpret evidence from studies of the earnings consequences of employment separations, such as Jacobson, Lalonde and Sullivan (1993), von Wachter et al. (2009), and Couch and Placzek (2010) who use data from a single state: Pennsylvania, California and Connecticut, respectively. All these studies cannot track inter-state movers. There are some comparability caveats: these studies generally focus on separators with long job-tenure, which we do not consider, nor do we identify separations associated with substantial swings in the employer's employment, to identify "mass layoffs." Nevertheless, our evidence suggests that the ability to track movers across states substantially shifts the nonemployment distribution away from longer-term nonemployment, and that this decline is accounted for by an increase in job-to-job flows with no nonemployment. Our calculation also indicates that any observations of one, two or three quarters of nonemployment

[4] OH, NJ, NY, PA, and TX had a population of more than 8 million in the Census 2000.
[5] AZ, CO, IN, KY, LA, MD, MN, MO, SC, TN, VA, WA and WI had a population of between 3.5 and 8 million in the Census 2000.
[6] AK, HI, IA, ID, ME, MT, ND, NE, NM, OK, OR, RI, SD, UT, WV, VT and WY had a population of less than 3.5 million in the Census 2000.

are almost always true nonemployment. In short, these and related studies may employ disproportionately few direct job-to-job flows, and relatively more workers that have short spells of nonemployment.[7] As nonemployment is associated with earnings penalties as shown by Fallick, Haltiwanger and McEntarfer (2011), single-state studies of the earnings consequences of employment displacement may overstate the earnings penalty associated with displacement.

These simple tabulations on the underlying microdata and provide some guidance for exploring how we should think about missing states affecting national estimates of job-to-job flow rates. Given the frequency of inter-state migration and the frequency of work outside of the geography that determines whether a worker is in scope, there is some room for missing states to influence our estimates away from direct job-to-job flows and into spells that involve longer-term nonemployment.

Calculating the Frequency of Job-to-Job Flows

We now use the LEHD microdata described above to calculate job-to-job flows. Following Hyatt and McEntarfer (2012a, 2012b), we restrict ourselves to flows between primary jobs, that is, the jobs where workers exhibit maximal earnings. We refer to these flows between primary jobs as job-to-job flows. We then track flows between these primary jobs, distinguishing between flows that occur with little or no nonemployment, and those with nonemployment spells between jobs. This will permit some assessment of our hypothesis that separations to non-employment may be overestimated, while direct job-to-job flows may be underestimated.

In this study, we aggregate ten minor categories of job flows into three major categories, as shown in Table 4.[8] The first major category is job-to-job flows that occur in the same or adjacent quarter as the primary job separation. This aggregate of flows also includes flows that

[7] These and related studies often drop those who separate from employment and never are observed working subsequently.

[8] For exact definitions of our measures, which closely follows Hyatt and McEntarfer (2012b), see our Appendix. There, we also provide an explanation for how missing geographic data affects job-to-job flow estimation differently from the establishment-level concepts such as hires and separations and job creation and job destruction considered in Abowd and Vilhuber (2011).

occur without a separation, a hire, or both. The second major grouping of job flows are those flows which separate into exactly one full quarter of nonemployment. The final major category contains flows that separate into more than one full quarter of nonemployment. Within this category of flows are flows with 2-3 quarters nonemployment, four or more quarters nonemployment, and primary separations to nonemployment without an observed subsequent hire. Note that we limit our focus on the core job-to-job flow measures: the frequency of flows with little or no nonemployment, and flows into longer term nonemployment. We do not incorporate employer characteristics such as industry or geography, nor do we consider flow rates for demographic subgroups. For each flow, the denominator is the number of dominant jobs in that quarter.

In Table 5, present some evidence on how sensitive our job-to-job flow measures are to the inclusion or exclusion of different states. Specifically, we calculate the percentages in which flows occurred in a single reference state, across reference states, from a reference state to a non-reference state (and vice-versa), from a non-reference state to a non-reference state, and within a single non-reference state. The majority of each flow type occurs within the set of reference states. More than 63% of each flow types occur exclusively in the reference states with more than 61% of the flows happening within a single reference state. Flows with one full quarter of nonemployment had the highest proportion (72.1%) of flows within a single reference state.[9] For each measure, between 26% and 36% of flows include non-reference states. Of the flows including non-reference states, most occurred within a single non-reference state. There are also a substantial number of inter-state flows: 17% of all within-quarter flows are across state boundaries, and 20% of all adjacent-quarter flows are across state boundaries.

Table 5 shows the limitations of calculating job-to-job flows based on a set of reference states alone, omitting information for other states, and so provides an indication of the sensitivity of job-to-job flow estimates to the inclusion or exclusion of states. Flows that move from reference states into other states would be viewed as either "last separations" or if later returning to the reference states could be viewed as flows to longer lengths of nonemployment than what actually occurred. Conversely, flows from non-reference states to reference states would either

[9] Some of these gaps may be due to spurious one quarter gaps in employment due to errors in longitudinal linkage, some types of which are considered in Abowd and Vilhuber (2005).

be viewed as "first hires" or again as longer periods of nonemployment. All flows that occur exclusively within non-reference states (13% to 21% of each type) would not be observed at all.

Estimation

In this section, we describe our method for estimating job-to-job flow rates in the presence of partially missing geography, focusing on. Our method involves estimating the marginal effects of removing states with different observable characteristics from the frame in order to estimate the unobserved national time series. To do so, we remove each state (excluding reference states[10]) and then re-calculate the job-to-job flow counts and the number of dominant jobs when the state is removed from the data. Specifically, we drop each of thirty-eight[11] observed non-reference states and re-calculate the measures with forty-six observed states (thirty-seven states plus the nine reference states). We then use weighted least squares to estimate the effect of missing states on the calculated job-to-job flow statistics. Weighted least squares allows larger states as well as those states that are closer to the reference states to have more of an impact on the outcome of interest.

Each state's omission constitutes an observation in our dataset. We employ a linear model for the relationship between each state's observable characteristics, and employ a standard estimation method linear estimation on state-level data, weighted least squares. The response variables are the change in flow rates of each type of job-to-job flow when each state is removed from the data. The regressors are observable characteristics of each state: the state's total employment from the Quarterly Census of Employment and Wages interacted with the state's difference in distance between each reference state.[12]

Our method is as follows: First, we calculate the marginal effects from an OLS estimation of the relationship between the observable characteristics and their marginal effects

[10] Reference states do not need to be included by construction: there are no partially or fully missing state that will affect the list of workers considered in this analysis.
[11] There are 38 non-reference states that are observed at some point. At the start of the time series this number is just 32.
[12] We also experimented with different independent variables, such as the Census Bureau's population estimates as well as additional data from the Quarterly Census of Employment and Wages such as the number of business establishments and total wages, but these additional characteristics were highly collinear with total employment, and did not seem to add to the model's explanatory power.

on the outcome variable of interest. The dependent variable is the difference between a rate when all available states $s \in P$ are used r_P, where P is the set of available states, and the rate when state s is omitted r_{P-s}, which we write as

$$\Delta_s r = r_P - r_{P-S} \, .$$

The equation we estimate is

$$\Delta_s r = X_s \beta + \varepsilon_s,$$

where X_s are the observable characteristics of state s, β are their marginal effects on the dependent variable, and ε_s is an error term which we assume is from a normal distribution but is heteroskedastic, as states are of different sizes, have different populations, and are different distances from each other. In this case, Ordinary Least Squares (OLS) estimation of β will be unbiased but inefficient.

We proceed using a standard method of implementing weighted least squares where the exact form of the heteroskedasticity is unknown. We first obtain an OLS estimate of β, which we call $\hat{\beta}_{OLS}$ and calculate the fitted residuals from this regression

$$\hat{\varepsilon}_s = \Delta_s r - X_s \hat{\beta}_{OLS} \, .$$

Then, we estimate the marginal effects of the observable characteristics on the absolute value of those residuals,

$$|\hat{\varepsilon}_s| = X_s \gamma + \eta.$$

The inverse of the squared predicted residuals from this specification,

$$1/(X_s \hat{\gamma})^2 \, ,$$

is used to weight each observation, in a final weighted linear regression that yields the weighted least squares estimate of β, which we call $\hat{\beta}_{WLS}$.

Given this framework, it is straightforward to calculate an estimate of the rate that would be observed were all states available r_A where A is the set of all states, both $s \in P$ and $s \notin P$. We can now write the estimate of r_A as

$$\widehat{r_A} = r_P + \sum_{s \notin P} X_{s_1} \hat{\beta}_{WLS} \, .$$

The estimated error distribution associated with each state is a function of its observable characteristics and the estimated effects of those relationships on the rate of interest, so the total error will be distributed according to

$$\mathcal{N}\left(0, \sum_{s \notin P} (X_s \hat{\gamma})^2\right)$$

which we use to formulate confidence intervals to test whether our estimates $\widehat{r_A}$ are significantly different from r_P.

Note that this method is not without its limitations, and readers should exercise some caution when interpreting our results. Our estimate $\widehat{r_A}$ will only be a valid estimate of r_A if the linear model applies beyond the set of available states r_P. That assumption seems less plausible when more states are unavailable. In the results that follow, earlier in the time series, ten states are ,ossomg, while later in the time series, only three are. We do not attempt to test this linearity assumption, but we at least find it plausible that the number of missing states could be sufficiently small for our method to yield valid estimates.

Results

In this section, we present the results of our estimation method on the three major categories of job-to-job flows described above: direct job-to-job flows that involve no full-quarter nonemployment, in which the flow occurs in the same or adjacent quarters, flows that involve a single quarter of nonemployment, and flows that involve more than one quarter of nonemployment. We present three Figures, each of which contains the rate of job-to-job flows measured in two ways: one is a direct tabulation of the microdata using all available states, and the other is the rate implied by the estimation method above. For our estimated line, we provide a 95% confidence interval for reference on whether the differences are statistically different. Our analysis indicates that missing states do have a statistically significant impact on job-to-job

flow rates, and that as the number of missing states decreases, the direct tabulations and the rates that result from our estimation method converge.

Our estimation indicates that the rate of job-to-job flows that involve no full-quarter nonemployment is underestimated in a direct tabulation of the microdata, as shown in Figure 2. The rates from both direct tabulation and our estimation method exhibit very similar trends. Both rates show a very strong downward trend, decreasing from over 10% in the late 1990s, and with both exhibiting a series high in 2000:1. During and after the 2001 recession, the rates decline, and each hits a local minimum of just over 8% in 2003:2, and then both increase to more than 9% in 2005:3, and subsequently decline again, exhibiting a series low in 2008:2, toward the middle of the recession that began in 2007. Throughout the series, the estimate is higher than the direct tabulation. The gap is slightly less than half a percentage point at the start of the series, and declines to less than four tenths of a percentage point in the year 1999, is less than 2 tenths of a percentage point in 2000, and is less than one tenth of a percentage point in 2002:3 and afterward. The differences are statistically different, even when the gaps are small.

The declines in the gap between the rate of job-to-job flows that follow from direct tabulation and the rate implied by our estimation method, yields some understanding of how missing states affect tabulation. As an initial set of relatively small states (Delaware, Iowa, Nebraska, and Utah) become available from 1998:3 to 1999:4 we see a rather modest change in the difference in rates. However, when a large state like OH, enters the time series in 2000:1, along with the rather small Oklahoma and Vermont, the gap between observed and estimated rates exhibits more of a decrease. As the last few states (Alabama, Arkansas and Mississippi), enter the series, the difference continues to decline. After 2003:3, no other states enter the time series, and only Massachusetts, New Hampshire, Connecticut, and the District of Columbia are missing. Throughout this period, the difference between observed and estimated rates is relatively consistent. Their omission leaves a small but statistically significant difference between direct tabulation and the result of our estimation method.

Job-to-job flows that involve exactly one quarter of nonemployment are shown in Figure 3. The rate of a flow involving exactly one quarter of non-employment declines between 1998 and 2008 from around three percent to just over 2 percent. The gap between the rate obtained through direct tabulation and the result of our method is quite small, never more than three

hundredths of a percentage point from each other. Although this gap is small, direct tabulation indicates a rate that is within the 95% confidence interval around our estimate in only eleven out of forty-two quarters in the series. The estimated rates are slightly above those of direct tabulation. For flows involving one quarter of nonemployment, there is less of a clear pattern between the particular states that are missing at different points in the series and the gap between direct tabulation and our estimates. This small difference is the sum of many mechanisms by which missing states may affect tabulation, with no clear overall direction for the error in direct tabulation: some apparent movements into longer-term nonemployment are in fact movements into exactly one quarter of nonemployment, some within-state flows in missing states involve exactly one full quarter of nonemployment, and some apparent quarters of nonemployment are flows correspond with jobs held in missing states.

Separations that involve more than one quarter of nonemployment are shown in Figure 4. Again, the rates that result from direct tabulation and those implied by our method tend to track each other. Both rates are around 4.5% until the 2001 recession, when they both jump to more than 4.9% in 2001:2, then decline and are consistently under 4% from 2006:3 onwards. The gap between direct tabulation and our estimates is initially more than one tenth of a percentage point, drops below one tenth in 1999:4, and starting in 2002:3 are different by less than five hundredths of a percentage point. The direct tabulation is always above the maximum of the 95% confidence interval for our estimate.

Taken together, the results may seem to contradict the finding in Abowd and Vilhuber (2011), that missing geography has no impact on aggregate separation rates. In fact, the results are not inconsistent. This is because of how the sample is defined, based on a list of workers who were ever employed in a number of reference states. Workers appear to have little or no dynamics when working in a non-included state, which means that missing states will basically by construction[13] lead to lower rates of separations and accessions. This does not imply that separation rates calculated at the establishment level will be affected by missing geography: by definition, a business establishment has a single geography at any point in time, so all of its separations and accessions are observed if the establishment is in the available geography at all.

[13] Except in the unlikely case that the missing geographies involve extremely stable employment.

For precise definitions of the measures considered in this paper and in Abowd and Vilhuber (2011), please see our Appendix.

Conclusion

We have demonstrated what is, to the best of our knowledge, the only method available to date for estimating job-to-job flow rates when certain geographies over which those flows are defined are missing. This new method was necessitated by the problem of attempting to define worker-level employment measures when data for only a subset of US states are available. The method we propose involves first calculating a direct tabulation of a job-to-job flow rate, then dropping one state at a time to estimate the marginal effect of an omitted state with particular observable characteristics on the job-to-job flow rate. We then use the observable characteristics of the omitted geographies and the estimated marginal effects of their omission to create an estimate of what job-to-job flow rates would be tabulated were no states missing.

Our analysis shows that the inter-state market for labor has a substantial impact on the job-to-job flow rates estimated by Hyatt and McEntarfer (2012a, 2012b). From a one percent sample of list of workers who ever worked in any of nine states, and within a subsequent frame of up to forty-seven states, over one-third of the workers in dataset have been employed in more than one state. The number of missing states declines for later years, there is a clear positive relationship between the number of missing states have entered the frame, and the percent of workers who are nonemployed in a given quarter. About 17% of within-quarter flows and about 20% of other flows are across states rather than within a single state.

The estimation method yields rates of job-to-job flows that generally track the rates implied by direct tabulation. Estimates are more different from the direct tabulation the more states are omitted. For the three categories of job-to-job flows we analyzed, the rate of job-to-job flows that involve no full-quarter nonemployment are higher in our estimation method, separations to more than one quarter of nonemployment are lower in our estimation method, and flows that involve exactly one quarter of nonemployment are very close between our estimate and that implied by direct tabulation. This is consistent with our intuition that certain flows from non-missing to missing states are sometimes miscoded as flows to nonemployment when they

are, in fact, direct job-to-job flows. The rate of underestimation of job-to-job flows that involve no nonemployment is more than twice the rate of overestimation of separations that involve more than one quarter of nonemployment, which is at least in part due to the fact that direct job-to-job flows that occur within missing states are not observed.

As LEHD's job-to-job flows project moves from the pilot database of Hyatt and McEntarfer (2012a, 2012b), which involves workers in a number of reference states, toward a frame that is the LEHD universe, the relationships between the observable characteristics and the employment flow measures may be somewhat different. In particular, there will be no "reference states" (or, rather, every state will be a reference state) and so omitting a state would necessarily involve modifying the list of individuals who work in the subsample versus all nonmissing states. Nevertheless, the basic approach should continue to be a valid method for correcting the output statistics with respect to missing geography. We also predict that the rate of job-to-job flows

This analysis has some bearing on how researchers should approach constructing a sample from a multi-state aggregation such as the LEHD infrastructure files. There is a certain strain of thought that suggests that researchers should not use all available states: analyses should be structured to avoid the use of states that have potential boundary issues for part of their time series, such as population centers that are either on or just across state boundaries. A frequent example is a state like Maryland - because the District of Columbia and Delaware have availability limitations. Despite the exclusion of such potentially problematic states, we still found that a substantial number of cross-state flows were not included in the dataset we constructed. This indicates that excluding states with potential boundary issues has a limited role in reducing the impact of cross-state movements on longitudinal analysis of employment. We recommend that researchers utilize cross-state flows in order to adjust estimates for the inclusion or exclusion of given states, rather than ignoring those states entirely.

Finally, this analysis highlights some of the difficulties that arise when defining a population based on whether someone has worked in a geography over a particular time interval, and how it may not be a representative cross-section of this geography. Inter-state mobility is quite common and those movements can lead a population defined as ever worked in a particular time interval to be different from a cross section, because a population defined on an interval

includes both movements into and movements out from the geography. This consequence of sample definition may account for some of the differences in studies of earnings loss from employment displacement in administrative records sources relative to survey data that is constructed to be representative, such as the Current Population Survey and the Panel Study of Income Dynamics.

Appendix

Following Hyatt and McEntarfer (2012b), we can define our employment measures. We define Flow Employment, where w_{ijt} is the total earnings of individual *i* at employer *j* in quarter *t*, as

$$m_{ijt} = \begin{cases} 1, \text{ if } w_{ijt} > 0 \\ 0, \text{otherwise} \end{cases}$$

and we define individual *i*'s dominant job as

$$d_{ijt} = \begin{cases} 1, \text{if } w_{ijt} > w_{imt} \ \forall j \neq m \\ 0, \text{otherwise} \end{cases}.$$

We consider dominant job flows from "origin" employer *j* to "destination" employer *k*. Two events indicate what we call a "within-quarter" job-to-job flow. Where the separation from the origin employer occurs in quarter *t*, we define

$$dd_{ijkt0_A} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 0 \text{ and } m_{ikt-1} = 0 \text{ and } m_{ikt} = 1 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

and where the separation from the origin employer occurs in quarter *t*+1, we define

$$dd_{ijkt0_B} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 1 \text{ and } m_{ijt+2} = 0 \text{ and } m_{ikt} = 0 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

When there is no coincidental hire (and so a continuing job becomes a main job), we define

$$dd_{ijkt0_C} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 0 \text{ and } m_{ikt-1} = 1 \text{ and } m_{ikt} = 1 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

and

$$dd_{ijkt0_D} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 1 \text{ and } m_{ijt+2} = 0 \text{ and } m_{ikt} = 1 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

When there is a hire but no coincidental separation, we define

$$dd_{ijkt0_E} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 1 \text{ and } m_{ikt-1} = 0 \text{ and } m_{ikt} = 1 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

and

$$dd_{ijkt0_F} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 1 \text{ and } m_{ijt+2} = 1 \text{ and } m_{ikt} = 0 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

For transitions where there is no separation from the origin job, we define dominant job-to-dominant job flows as follows.

$$dd_{ijkt0_G} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = m_{ijt+2} = m_{ikt-1} = m_{ikt} = 1 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

For cases in which the separation and subsequent hire occur in adjacent quarters, we define

$$dd_{ijkt1} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } m_{ijt+1} = 0 \text{ and } m_{ikt} = 0 \text{ and } d_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}$$

which would not include any spell of nonemployment for cases in which the separation and hire occur immediately before and after, respectively, the date on which a quarter starts, but may frequently imply some small duration of nonemployment.

For job-to-job flows with a nonemployment spell, for any $p \geq 2$

$$dd_{ijktp} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } w_{i \cdot t+1} = \cdots = w_{i \cdot t+p-1} = 0 \text{ and } d_{ikt+p} = 1 \\ 0, \text{otherwise} \end{cases}$$

where $w_{i \cdot t}$ is the total earnings of individual *i* in quarter *t*.

We also allow flows into nonemployment that never has an observed hire, defined as

$$dn_{ijtp} = \begin{cases} 1, \text{if } d_{ijt} = 1 \text{ and } w_{i \cdot \tau} = 0 \ \forall \tau > t \\ 0, \text{otherwise} \end{cases}.$$

We group these employment flow measures into rates via the following formulas, with respect to time *t*. Direct job-to-job flows are

$$\frac{\sum_i \sum_j \sum_k \left(dd_{ijkt0_A} + dd_{ijkt-10_B} + dd_{ijkt0_C} + dd_{ijkt-10_D} + dd_{ijkt0_E} + dd_{ijkt-10_F} + dd_{ijkt0_G} \right)}{\sum_i \sum_j d_{ijt}},$$

adjacent-quarter job-to-job flows are

$$\frac{\sum_i \sum_j \sum_k dd_{ijkt1}}{\sum_i \sum_j d_{ijt}},$$

single-quarter nonemployment is

$$\frac{\sum_i \sum_j \sum_k dd_{ijkt2}}{\sum_i \sum_j d_{ijt}},$$

and separation to nonemployment of more than one quarter is

$$\frac{\sum_i \sum_j \sum_k \left(\left(\sum_{p>2} dd_{ijktp}\right) + dn_{ijt}\right)}{\sum_i \sum_j d_{ijt}},$$

This paper is most closely related to Abowd and Vilhuber (2011), who address the problem of calculating national-level statistics when only a subset of states are available. Specifically, they construct national-level estimates of a subset of the LEHD program's Quarterly Workforce Indicators: job creation and destruction as well as separations and hires. All the indicators they consider are job-specific, that is, concerning worker-employer combinations. As all jobs are tied to specific geographies, they can impute the employment in each missing geography under as a conventional missing data problem.

Given our concept of Flow Employment, we can write down the employment measures that are in Abowd and Vilhuber (2011). For example, consider their definition of turnover. This is derived after defining beginning-of-quarter employment as

$$b_{ijt} = \begin{cases} 1, m_{ijt-1} = 1 \text{ and } m_{ijt} = 1 \\ 0, \text{otherwise} \end{cases},$$

end of quarter employment as

$$e_{ijt} = \begin{cases} 1, m_{ijt} = 1 \text{ and } m_{ijt+1} = 1 \\ 0, \text{otherwise} \end{cases},$$

separations as

$$s_{ijt} = \begin{cases} 1, m_{ijt} = 1 \text{ and } m_{ikt+1} = 0 \\ 0, \text{otherwise} \end{cases},$$

and hires (abbreviated from the equivalent term accession) as

$$a_{ijt} = \begin{cases} 1, m_{ijt} = 0 \text{ and } m_{ikt+1} = 1 \\ 0, \text{otherwise} \end{cases}.$$

Their turnover measure is then

$$\frac{\sum_j s_{ijt} + \sum_j a_{ijt}}{\left(\sum_j b_{ijt} + \sum_j e_{ijt}\right)/2}$$

This measure does not require looking across multiple jobs for the same person, nor does it require comparing employers across geographies. Each employer *j* is associated with exactly one geography at any point in time. All of its separations and accessions are observed if the establishment is in the available geography at all. This makes their measures different from the ones considered in this paper, which calculate dominant job flows from origin employer *j* to destination employer *k* for a list of workers *i* who ever worked in a number of reference states.

References

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon D. Woodcock. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In Production Dynamics: New Evidence from Micro Data, Vol. 68, Studies in Income Wealth, ed. Timothy Dune, J. Bradford Jensen and Mark J. Roberts, 149-230. Chicago: University of Chicago Press.

Abowd, John M. and Lars Vilhuber. 2011. "National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail." *Journal of Econometrics*, 161(1): 82-99.

Abowd, John M. and Lars Vilhuber. 2005. "The Sensitivity of Economic Statistics to Coding Errors in Personal Identifiers." *Journal of Business and Economic Statistics*, 23(2): 133-152.

Andersson, Fredrik, John Haltiwanger, Mark Kutzbach, Henry Pollakowski and Daniel Weinberg. 2011. "Job Displacement and the Duration of Joblessness: The Role of Spatial Mismatch." US Census Bureau Center for Economic Studies Discussion Paper No. CES-WP-11-30.

Fallick, Bruce, John Haltiwanger and Erika McEntarfer. 2011. "Nonemployment Duration and the Consequences of Job Separations." Paper presented at the 2011 Society of Labor Economists meeting in Chicago, IL. Available at http://www.sole-jole.org/11156.pdf (accessed: August 2nd, 2012).

Hyatt, Henry, and Erika McEntarfer. 2012a. "Job-to-Job Flows in the Great Recession." *American Economic Review*, 102(3): 580–83.

Hyatt, Henry and Erika McEntarfer. 2012b. "Job-to-Job Flows and the Business Cycle." Center for Economic Studies Discussion Paper CES-12-04.

Jacobson, Louis, Robert LaLonde and Daniel Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review*, 83(4): 685-709.

Molloy, Raven, Christopher Smith and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, 25(3), 173-196.

Rubin, Donald. 1987. *Multiple Imputation for Nonresponise in Surveys*. New York: Wiley.

McKinney, Kevin and Lars Vilhuber. "LEHD Infrastructure Files in the Census RDC: Overview of s2004 Snapshot." US Census Bureau Center for Economic Studies Discussion Paper No. CES-WP-11-13.

Von Wachter, Till M., Elizabeth W. Handwerker and Andrew K. G. Hildreth. 2009. "Estimating the 'True' Cost of Job Loss: Evidence Using Matched Data From California 1991-2000." US Census Bureau Center for Economic Studies Working Paper 09-14.

Table 1: Missing Data, by State

	Entrance Dates								
State	1998:2	1998:3	1998:4	1999:1	2000:1	2001:1	2002:3	2003:3	2008:3
AL						A	A	A	A
AR							A	A	A
CT									
DC									
DE		A	A	A	A	A	A	A	A
IA			A	A	A	A	A	A	A
MA									
MS								A	A
NE				A	A	A	A	A	A
NH									
OH					A	A	A	A	A
OK					A	A	A	A	A
UT				A	A	A	A	A	A
VT					A	A	A	A	A

Notes: The letter A indicates that a state has available data in this and any subsequent quarter. Blank indicates that data is not available for the state at that time. Any state with data in both 2003:3 and 2008:3 has data for the intermediate quarters.

Table 2: Frequency of Ever Working in a Given Number of States, 1998-2008 (in thousands)

Number of States	Frequency	Percent
1	563.1	60.6%
2	236.0	25.4%
3	82.9	8.9%
4	28.8	3.1%
5	10.3	1.1%
6 or more	8.7	0.9%

Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Employment histories of workers are taken from a set of 48 states from 1998-2008. The start year can vary by state, see text for additional details.

Table 3: Effects of States' Inclusion on Estimates of Nonemployment

		Nonemployment Type					
Calculation Set	Cumulative Number of States	No Full Quarters of Nonemployment	One Full Quarter of Nonemployment	Two Full Quarters of Nonemployment	Three Full Quarters of Nonemployment	Four or More Quarters Nonemployment	No Accession Observed by 2008:4
Reference States Only	9 States	-	37.0%	12.0%	5.7%	21.8%	23.5%
Add Large States	14 States	2.6%	37.2%	12.1%	5.7%	22.0%	20.1%
Add Medium States	27 States	7.8%	37.2%	12.1%	5.6%	20.9%	16.4%
Add Small States	44 States	9.2%	37.2%	12.1%	5.6%	20.6%	15.3%

Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Tabulations are done on the subset of individuals who separate from a job in a reference state in 2000:1 who have no earnings in any reference state in 2000:2. Employment histories of workers are taken from a set of 44 states with continuous available data for the year 2000 onwards. Large states are non-reference states with a population of at least 8 million in the Census 2000: OH, NJ, NY, PA, and TX. Medium states are non-reference states with a population at least 3.5 million but less than 8 million in the Census 2000: AZ, CO, IN, KY, LA, MD, MN, MO, SC, TN, VA, WA and WI. Small states are non-reference states with a population less than 3.5 million in the Census 2000: AK, HI, IA, ID, ME, MT, ND, NE, NM, OK, OR, RI, SD, UT, WV, VT and WY.

Table 4: Job-to-Job Flow Concepts

Major Category	Minor Category	Analogue in Appendix
Direct Job-to-Job Flows	Within-Quarter Flows	$dd_{ijkt0_A} + dd_{ijkt-10_B}$
	Adjacent-Quarter Flows	dd_{ijkt1}
	No Coincident Hire	$dd_{ijkt-10_C} + dd_{ijkt0_D}$
	No Coincident Separation	$dd_{ijkt-10_E} + dd_{ijkt0_F}$
	No Hire or Separation	dd_{ijkt0_G}
One Quarter of Nonemployment	One Quarter of Nonemployment	dd_{ijkt2}
More than One Quarter of Nonemployment	2-3 Quarters of Nonemployment	$dd_{ijkt3} + dd_{ijkt4}$
	Four or More Quarters of Nonemployment	$\sum_{p>4} dd_{ijktp}$
	No Observed Hire	dn_{ijt}

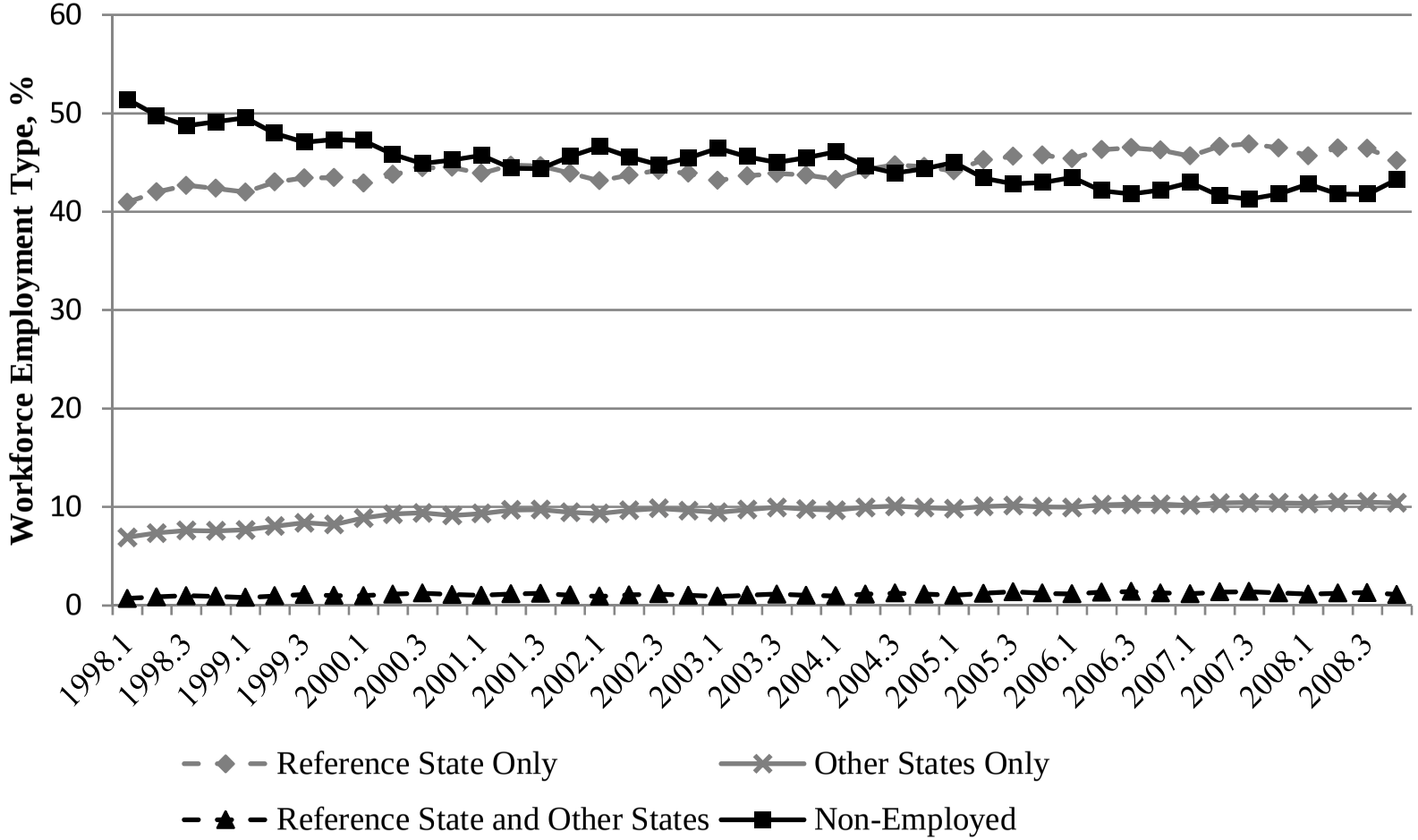
Notes: Categories of job-to-job flows. See text for additional details.

Table 5: Job-to-Job Flows Within and Across States

	Job-to-Job Flow Type (Column Percents)						
	Within-Quarter	Adjacent Quarter	One Full Quarter Nonemplyoment	Two or Three Full Quarters of Nonemplyoment	Four or More Full Quarters of Nonemplyoment	Only Separation	Only Accession
Within Single Reference State	64.2%	61.1%	72.1%	67.0%	61.3%	65.0%	65.8%
Reference to Other Reference	2.2%	2.8%	1.6%	2.3%	4.3%	3.0%	2.6%
ALL REFERENCE ONLY	*66.5%*	*63.9%*	*73.7%*	*69.3%*	*65.6%*	*68.0%*	*68.4%*
Reference to Non-Reference	5.9%	6.9%	4.0%	6.1%	10.6%	6.9%	6.3%
Non-Reference to Reference	5.6%	7.0%	3.9%	5.9%	10.0%	6.8%	6.0%
Within Single Non-Reference State	18.7%	18.6%	16.6%	16.2%	10.7%	15.3%	16.4%
Non-Reference to Other Non-Reference	3.2%	3.7%	1.8%	2.7%	3.1%	3.1%	3.0%
ALL NON-REFERENCE	*33.6%*	*36.1%*	*26.3%*	*30.7%*	*34.4%*	*32.0%*	*31.6%*

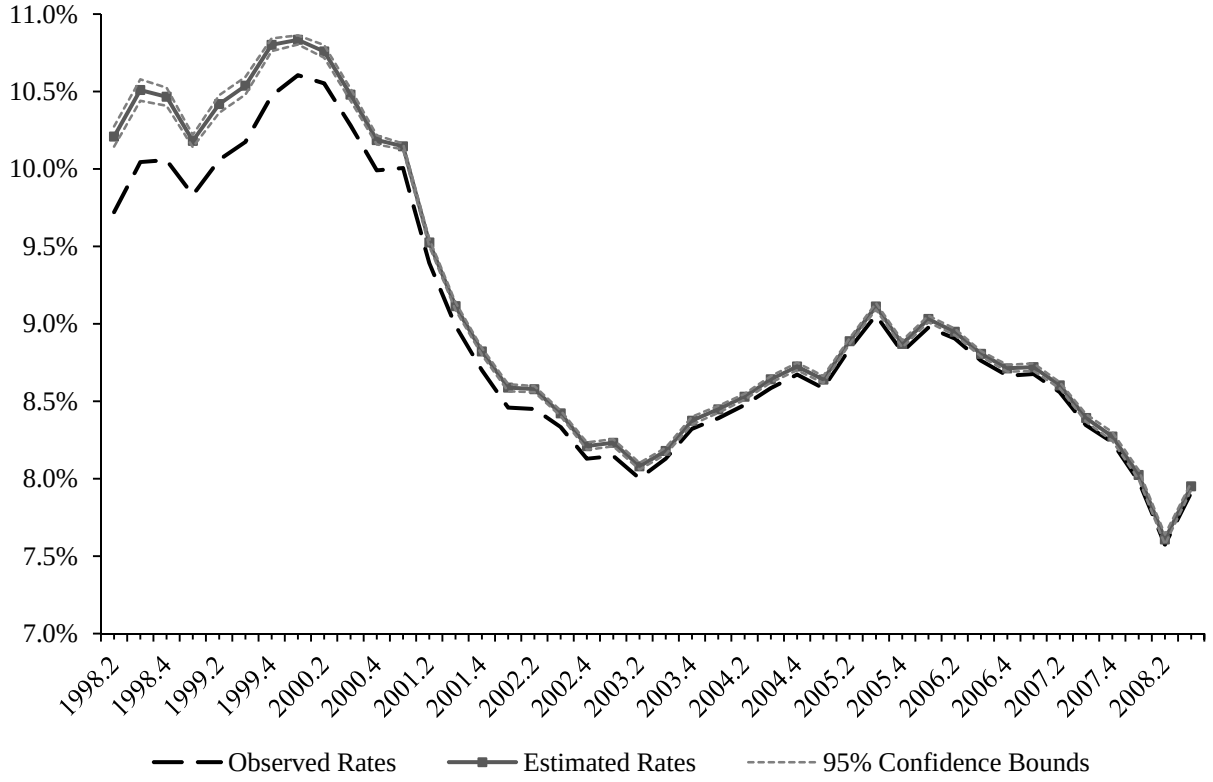
Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Employment histories of workers are taken from a set of 48 states from 1998-2008. The start year can vary by state, see text for additional details.

Figure 1: Employment Status of LEHD Workers



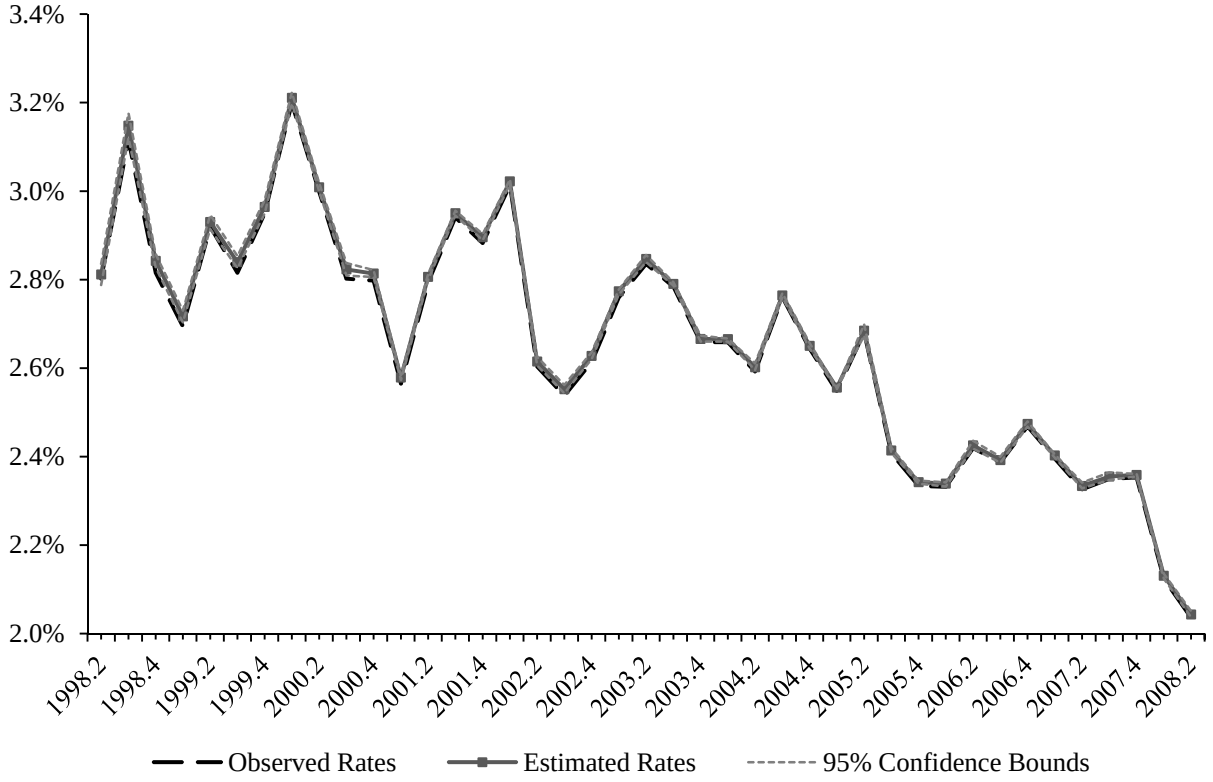
Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Employment histories of workers are taken from a set of 48 states from 1998-2008. The start year can vary by state, see text for additional details.

Figure 2: Job-to-Job Flow Rates (Within- and Adjacent Quarter Flows), 1998:2 – 2008:3



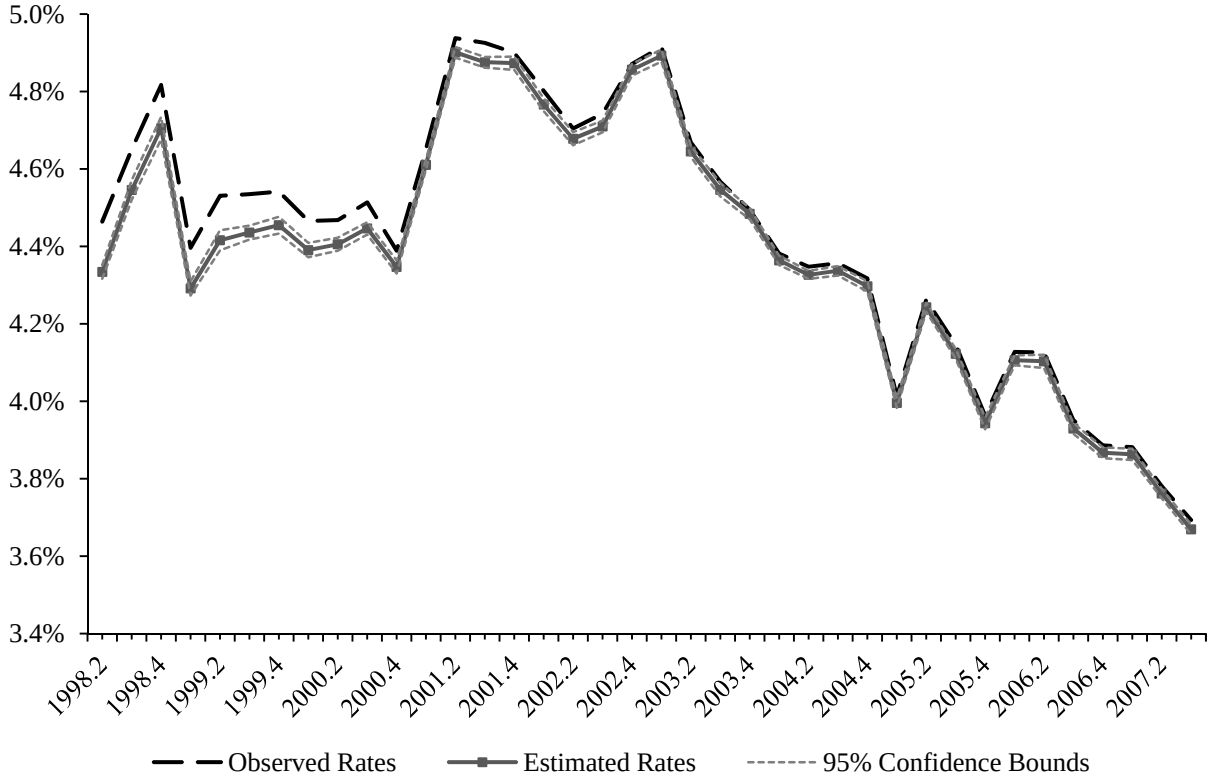
Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Employment histories of workers are taken from a set of 48 states from 1998-2008. All series are seasonally adjusted. The start year can vary by state, see text for additional details.

Figure 3: Job-to-Job Flows with Exactly One Quarter of Nonemployment, 1998:2 - 2008:2



Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Employment histories of workers are taken from a set of 48 states from 1998-2008. All series are seasonally adjusted. The start year can vary by state, see text for additional details.

Figure 4: Separation to More than One Quarter of Nonemployment, 1998:2 - 2007:3



Notes: Calculations on a 1% sample from the s2008 LEHD snapshot, limited to workers ever employed in state unemployment-insurance taxable employment in CA, FL, GA, IL, KS, MI, NV, NC and ND from 1998-2008. Employment histories of workers are taken from a set of 48 states from 1998-2008. All series are seasonally adjusted. The start year can vary by state, see text for additional details.