

Abraham, Katharine G.; Haltiwanger, John C.; Sandusky, L. Kristin; Spletzer, James R.

**Working Paper**

## The Consequences of Long Term Unemployment: Evidence from Matched Employer-Employee Data

IZA Discussion Papers, No. 10223

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Abraham, Katharine G.; Haltiwanger, John C.; Sandusky, L. Kristin; Spletzer, James R. (2016) : The Consequences of Long Term Unemployment: Evidence from Matched Employer-Employee Data, IZA Discussion Papers, No. 10223, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<http://hdl.handle.net/10419/147909>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 10223

**The Consequences of Long Term Unemployment:  
Evidence from Matched Employer-Employee Data**

Katharine G. Abraham  
John Haltiwanger  
Kristin Sandusky  
James R. Spletzer

September 2016

# **The Consequences of Long Term Unemployment: Evidence from Matched Employer-Employee Data**

**Katharine G. Abraham**

*University of Maryland, IZA and NBER*

**John Haltiwanger**

*University of Maryland, IZA and NBER*

**Kristin Sandusky**

*U.S. Census Bureau*

**James R. Spletzer**

*U.S. Census Bureau and IZA*

Discussion Paper No. 10223

September 2016

IZA

P.O. Box 7240

53072 Bonn

Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **The Consequences of Long Term Unemployment: Evidence from Matched Employer-Employee Data\***

It is well known that the long-term unemployed fare worse in the labor market than the short-term unemployed, but less clear why this is so. One potential explanation is that the long-term unemployed are “bad apples” who had poorer prospects from the outset of their spells (heterogeneity). Another is that their bad outcomes are a consequence of the extended unemployment they have experienced (state dependence). We use Current Population Survey (CPS) data on unemployed individuals linked to wage records for the same people to distinguish between these competing explanations. For each person in our sample, we have wage record data that cover the period from 20 quarters before to 11 quarters after the quarter in which the person is observed in the CPS. This gives us rich information about prior and subsequent work histories not available to previous researchers that we use to control for individual heterogeneity that might be affecting subsequent labor market outcomes. Even with these controls in place, we find that unemployment duration has a strongly negative effect on the likelihood of subsequent employment. This finding is inconsistent with the heterogeneity (“bad apple”) explanation for why the long-term unemployed fare worse than the short-term unemployed. We also find that longer unemployment durations are associated with lower subsequent earnings, though this is mainly attributable to the long-term unemployed having a lower likelihood of subsequent employment rather than to their having lower earnings once a job is found.

JEL Classification: J64

Keywords: unemployment duration, heterogeneity, state dependence

Corresponding author:

James R. Spletzer  
Center for Economic Studies  
U.S. Census Bureau  
4600 Silver Hill Road  
Washington, DC 20233  
USA  
E-mail: [James.R.Spletzer@census.gov](mailto:James.R.Spletzer@census.gov)

---

\* We thank Lucia Foster, Henry Hyatt, and Till von Wachter for helpful comments and suggestions. We also have benefited from the opportunity to present earlier versions of the paper at the U.S. Census Bureau, the 2015 Society of Labor Economists meetings and the Henry S. Farber Festschrift held at Princeton University. Any opinions and conclusions expressed herein are those of the authors 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.

## I. Introduction

Long-term unemployment soared during the 2007-2009 recession and remained stubbornly high even as the economy recovered. By December 2009, the number of people unemployed for 27 weeks or more had risen above 6 million and it did not fall below that mark until October of 2011. The long-term share of total unemployment exceeded 40.0% from December 2009 through November 2012 and as of August 2016, more than seven years after the recession's trough, was still at approximately the same level as was observed at the previous series peak (26.0%) attained in June 1983. This paper provides new evidence on how experiencing a spell of long-term unemployment affects subsequent employment and earnings.

Concern that the long-term unemployed might have become permanently detached from the labor market rose in the wake of the Great Recession and the dramatic increase in long-term unemployment that followed. In an analysis using Current Population Survey (CPS) data, Krueger, Cramer and Cho (2014) (hereafter KCC 2014) found that, among those who had been unemployed 27 weeks or more as first-wave CPS respondents during 2008-2012, only 35.9% were employed 15 months later and just 10.8% were in full-time jobs that had lasted at least four months. While those who had been unemployed for a shorter period (26 weeks or less) also appear to have experienced significant labor market challenges—in the same data, only 49.5% of the short-term unemployed were employed 15 months later and just 14.4% were in full-time stable jobs—outcomes for the long-term unemployed were notably poorer. These findings are robust to controlling for observable worker characteristics such as age, gender, and education.

The finding that the long-term unemployed have fared less well than the short-term unemployed during the recent recovery period echoes earlier findings in the literature on how the job-finding hazard evolves over a spell of unemployment. A common finding for the United States, first documented by Kaitz (1970), is that longer unemployment durations are associated with lower exit or job-finding rates. It is less clear *why* the long-term unemployed experience poorer job-finding prospects than those who are more recently unemployed. One set of explanations rests on possible sources of state-dependence. Long-term unemployment may have a causal effect on the job finding rate because of human capital depreciation during spells away from work (Mincer and Ofek 1982, Stratton 1995, Acemoglu 1995, Albrecht et al 1999 and Gorlich and de Grip 2009), decreases in search intensity over the spell of unemployment (Krueger and Mueller 2011, Faberman and Kudlyak 2014), or employer discrimination against the long-term unemployed in hiring (Eriksson and Rooth 2011, Kroft, Lange and Notowidigdo 2013, and Ghayad 2013, but see Farber, Silverman and von Wachter 2015). Alternatively, what appears as a decline in the job finding rate could instead be the result of heterogeneity among the unemployed such that, at longer durations, more of the unemployed are “bad apples” with personal attributes that lead to their having poorer job-finding prospects (Cripps and Tarling 1974, Heckman and Singer 1984). The interpretation of the findings of KCC 2014 depends critically on whether the different outcomes for the long term unemployed are a result of the time they have spent in that state (i.e., are causal) or reflect underlying unobserved heterogeneity.

In data consisting of information on a single spell of unemployment for each included sample member, there is no straightforward way to fully distinguish between these competing explanations. After reviewing a number of studies that have attempted this task, Machin and

Manning (1999) conclude that “it does not really seem possible in practice to identify separately the effect of heterogeneity from that of duration dependence without making some very strong assumptions about functional form which have no foundation in any economic theory” (p. 3111).

One possible approach to distinguishing between the contributions of state dependence versus underlying heterogeneity as explanations for the decline in the job-finding hazard at longer unemployment durations is to exploit the information contained in multiple spells of unemployment for the same individuals, as in Alvarez, Borovickova and Shimer (2015). In this paper, we have adopted a different but related approach. Borrowing from methods used in the job displacement literature, we use information on individuals’ employment experiences prior to their becoming unemployed to control for heterogeneity in employment propensities as well as heterogeneity in the trajectory of those employment propensities that might explain subsequent employment outcomes. To the extent that the long-term unemployed are less likely to become employed even after controlling in this way for underlying heterogeneity, we can feel more confident about attaching a causal interpretation to the relationship between unemployment duration and job finding success.

Related to the question of how longer versus shorter spells of unemployment affect subsequent employment is the question of how they affect subsequent earnings. On the one hand, setting a higher reservation wage and searching longer for a better match could lead to higher earnings on the jobs eventually accepted by the long-term unemployed (Ehrenberg and Oaxaca 1976, Acemoglu and Shimer 2000). On the other hand, available evidence suggests that unemployed individuals turn down relatively few job offers (Schmieder, von Wachter and Bender 2014). To the extent that job searchers target their best opportunities first, this might lead one to expect a negative association between unemployment duration and subsequent earnings. Heterogeneity among the unemployed such that those with lower labor market productivity also take longer on average to find a job could produce the same pattern. We address the impact of differing durations of unemployment on subsequent earnings using essentially the same methods as just described for studying impacts on subsequent employment.

In addition to the literature on the consequences of experiencing unemployment for subsequent labor market outcomes, our paper is closely related to the job displacement literature. The central question in this literature concerns the adverse impact of job displacement on subsequent labor market outcomes, especially subsequent earnings. In addressing this question, it is important to take into account both the prior level and the prior trajectory of earnings. Our formal analysis borrows heavily from the approach taken in the job displacement literature (e.g., Jacobsen, LaLonde and Sullivan 1993; Farber 1993, 1997, 2005, 2013, 2015; Couch and Placzek 2010; von Wachter, Song and Manchester 2011; and Davis and von Wachter 2011). One difference between the population of displaced workers—typically, long-tenure workers who have lost their jobs during a mass layoff—and the population of unemployed individuals we study is that the unemployed include not only job losers but also new entrants to the labor market, labor market re-entrants, and people who voluntarily left their previous job.

To carry out our analysis, we have created a matched employer-employee dataset based on individuals who responded to the CPS from 2003 through 2010 that combines information on

individuals' labor force status and personal characteristics from the CPS with information on their employment and earnings histories from unemployment insurance (UI) wage records. Similar to KCC 2014 and consistent with other evidence on job finding rates by duration of unemployment, we find that CPS respondents with longer unemployment durations are less likely than the short-term unemployed to be employed in subsequent quarters. To illustrate, in a model that controls for observable demographic characteristics such as age, gender and education as well as year, individuals unemployed five quarters (53-65 weeks) at the time of their CPS interview are 14.6 percentage points less likely to have UI jobs eight quarters later than those unemployed one quarter (1-13 weeks) as of the CPS interview date.

This pattern could be driven by unobserved differences in underlying employment propensities between the long-term and the short-term unemployed. In a balanced panel, it is possible to compare the employment rate in a period that follows a spell of unemployment to the pre-unemployment employment rate for the same set of people. In our empirical implementation, we compare the employment rate eight quarters after a group is observed in the CPS to the employment rate for the same group eight quarters earlier. The magnitude of any change in the employment rate for the long-term unemployed then can be contrasted with that for the short-term unemployed. The advantage of this approach is that it removes the effects of fixed unobservable factors on employment rates from the comparison of outcomes across the different unemployment duration groups. Adopting this approach tends to raise the estimated adverse impact of being long-term unemployed rather than short-term unemployed on subsequent employment—for example, the 14.6 percentage point estimate of the relative disadvantage associated with having been unemployed five quarters rather than one quarter cited in the previous paragraph increases to 19.8 percentage points. This is because, except for those reporting unemployment of six quarters or more, the longer-term unemployed were actually more likely to have been employed eight quarters prior to being observed in the CPS than those in the shortest-duration unemployment group.

Accounting for the effects of different underlying employment trajectories moves our estimates of the relative disadvantage experienced by the long-term unemployed in the opposite direction. This is mainly because a disproportionate share of the short term unemployed are young adults and employment rates for that age group are generally rising over time. Constructing a post-unemployment benchmark for employment that takes into account the differences in employment trajectories across the unemployment duration groups thus makes the post-unemployment outcomes of the short-term unemployed look relatively worse. Our preferred estimates, after allowing both for fixed individual differences in employment propensities and for differences in employment trajectories related to both observable and unobservable individual characteristics, imply that those unemployed five quarters (52-65 weeks) at the time of their CPS interview experience employment losses that are 13.5 percentage points greater than the losses experienced by those unemployed just one quarter (1-13 weeks), not very different than the 14.6 percentage point estimate obtained by simply comparing post-unemployment outcomes across the same unemployment duration groups.

An additional question on which we are able to provide new evidence is how experiencing a longer versus a shorter spell of unemployment affects an individual's subsequent earnings. Average earnings for a group of unemployed individuals in any later period depend

both on the probability that members of the group become employed and on the amount earned by those who are employed. In an analysis of average earnings that mimics the analysis of the relationship between unemployment duration and employment just summarized, we obtain very similar results. Our preferred estimates imply that the long term unemployed experience earnings losses that are much larger than those for the short term unemployed. This is attributable mainly to differences in the probability of having a job as opposed to differences in earnings conditional on having a job.

The plan of the paper is as follows. Section II describes the linked data used in our analysis. In Section III, a descriptive analysis of the employment and earnings impact of longer versus shorter spells of unemployment is presented in the form of graphs that show how these variables evolve from the period 20 quarters before through 11 quarters after the month in which individuals are observed in the CPS data in our linked sample. Section IV offers a more formal regression analysis focused on the same labor market outcomes, and Section V concludes.

## **II. Data**

The data used in our analysis consist of CPS records matched to Longitudinal Employer-Household Dynamics (LEHD) micro-data for the same individuals. The CPS collects information on labor force activity and demographics from the members of approximately 60,000 households per month. The LEHD is a longitudinally linked employer-employee dataset created at the U.S. Census Bureau. The LEHD source data are administrative records from states' unemployment insurance (UI) systems that contain information on the earnings of all covered workers in that quarter, often referred to as wage records, along with Quarterly Census of Employment and Wages (QCEW) data containing information about the establishments at which these workers are employed. Every quarter, employers who are subject to state UI laws—approximately 98% of all private sector employers, plus state and local governments—are required to submit information on their workers (the wage records) and their workplaces (the QCEW) to the state UI agencies. The wage records and the QCEW data submitted to the Census Bureau by the state agencies are combined and enhanced with various census and survey micro-data to create the LEHD. Abowd et.al. (2009) provide a thorough description of the source data and the methodology underlying the LEHD data.

For recent years, the LEHD comes close to being a universe of all workers with wage and salary employment in the U.S. private sector, state government or local government. Because states have joined the LEHD at different points in time, however, data for fewer states are available in earlier years. Data for 31 states accounting for approximately 70% of the U.S. population are available beginning in 1998 and our main analysis uses LEHD data for these 31 states.<sup>1</sup>

The CPS identifier needed to link the CPS and LEHD records—the Protected Identity Key or PIK—is available only for March of each year. Our sample consists of respondents to the 2003 through 2010 March CPS's who lived in one of the 31 states for which LEHD data are

---

<sup>1</sup> The 31 states are CA, CO, CT, FL, GA, HI, ID, IL, IN, KS, LA, ME, MD, MN, MO, MT, NV, NJ, NM, NC, ND, OR, PA, RI, SC, SD, TN, TX, WA, WV, and WI.

available from 1998 onwards and for whom we have a PIK. The LEHD records are the source of the prior and future employment histories used in our analysis. The linked data set incorporates LEHD data for the period from 1998 through 2012, meaning that we can follow unemployed individuals backwards in time for 20 quarters and forward in time for 11 quarters from the date of their CPS interview (provided they remain in the states for which we have LEHD data). In addition, the LEHD includes information on earnings for each reported job, so that we also are able to observe the prior and future earnings profiles of the CPS respondents.

Our interest lies primarily with those who are classified as unemployed in the CPS as of the date of the CPS-LEHD match. For comparison purposes, we also construct similar employment and earnings profiles for the CPS employed and for the CPS not in the labor force (NILF). One complication is that the PIK is missing for about 20% to 30% of CPS respondents. We use propensity score methods similar to those described in Abraham et.al. (2013) to re-weight the observations for those for whom we have a PIK so that they represent the entire population.<sup>2</sup> There are a small number of people who report being unemployed in two successive March CPS's. We have not attempted to exploit this feature of the data, instead treating person-year observations in the CPS as independent.

Weighted tabulations of the linked CPS observations in our 31 state sample match published national tabulations very closely. Appendix Figure A.1 presents the 2003 to 2010 time series of the CPS employment-to-population ratio and the unemployment rate from our weighted data (labeled as AHSS in the figure) and from the BLS website (labeled as BLS in the figure). The BLS data are non-seasonally adjusted statistics for the month of March. The AHSS line tracks the BLS line almost perfectly. Appendix Figure A.2 presents similar graphs for the percent of the unemployed in each of several unemployment duration categories (<5 weeks, 5-10 weeks, 11-14 weeks, 15-26 weeks, 27-51 weeks, and  $\geq 52$  weeks). Again, the weighted AHSS time series are essentially identical to the published estimates.

In our 31-state data set, we do not observe UI jobs reported by employers in the other 19 states or the District of Columbia. If, for example, the long-term unemployed were more likely than the short-term unemployed to move to a non-covered state in order to take a job, this could be a source of bias in our estimates. By 2005, the coverage of the LEHD had expanded to 48 states accounting for approximately 96% of the U.S. population. To assess whether our findings might have looked different had the coverage of the LEHD data used in our main analysis been broader, we also have carried out a sensitivity analysis using LEHD data for the 48 states available in later years. In this sensitivity analysis, we use records for the people living in our 31-state sample who responded to the March CPS in 2008, 2009 or 2010. In that shorter time period, the results are little affected by whether we use UI job information based on the 31-state sample or larger 48-state sample.<sup>3</sup> This suggests that we do not need to be concerned about

---

<sup>2</sup> More specifically, for each year, we estimate a regression model in which an indicator for having a PIK is regressed on indicators for age group, gender, race, education, marital status, foreign-born status and whether the person reported being employed in the relevant March CPS. The coefficients from this model are used to calculate each individual's probability of having a PIK and a weight adjustment equal to the inverse of this probability is applied to the CPS monthly estimation weight. Individuals with a PIK are retained in our sample regardless of whether we are able to locate any employment records for them in the LEHD data.

<sup>3</sup> For example, with the 2008-2010 CPS sample, the employment probability eight quarters later for individuals unemployed five quarters as of the CPS interview date is 41.6% when computed using UI job data for 31 states and

possible biases in our main analysis resulting from the long duration unemployed having different cross-state mobility propensities than the short duration unemployed.

Table 1 presents descriptive statistics of the outcome variables and many of the explanatory variables used in our subsequent analysis. We are interested in how subsequent labor market outcomes differ for people who have been unemployed for longer versus shorter lengths of time as of the date of the CPS-LEHD link. The descriptive statistics are reported separately for people who had been unemployed up to 13 weeks, 14-26 weeks, 27-39 weeks, 40-52 weeks, 52-65 weeks, or more than 65 weeks as of the March CPS reference week, referred to as the one, two, three, four, five and six-plus quarter unemployment duration groups. For comparison, descriptive statistics for people who reported in the CPS that they were employed at the time of the CPS-LEHD link also are reported. The numbers of underlying observations in each of these groups is shown at the bottom of the table.

The first two panels of Table 1 report employment and earnings variables based on the UI wage records contained in the LEHD for the periods eight quarters prior to and eight quarters following the CPS-LEHD link quarter. People who are employed in the CPS at  $q=0$  (the link quarter) have a 72-73% UI employment rate both two years before and two years after the link quarter. Conditional on having UI employment, the CPS employed have average real quarterly earnings of \$12,661 two years before and \$13,403 two years after the link quarter. The pre-unemployment employment rates and average earnings for persons who are unemployed at  $q=0$  are much lower than those for the employed, and this is even more true of their post-unemployment employment rates and earnings. In the  $q=-8$  quarter, those who later become longer-term unemployed generally have somewhat higher employment rates and earnings than those who later become short-term unemployed. In the  $q=+8$  quarter, the pattern of differences across the unemployment duration groups is reversed.

The table also reports values of the Inverse Hyperbolic Sine (IHS) of earnings as of  $q=-8$  and  $q=+8$ . The IHS of a variable  $x$  is defined as:

$$\text{IHS}(x) = \ln \{x + \sqrt{1 + (x*x)}\} \tag{1}$$

For  $x>0$  and especially for  $x$  far from zero, this expression is approximately equal to  $\ln(x)$  shifted by a constant (the  $\ln$  of 2). Unlike the  $\ln(x)$  transformation, however, the IHS transformation can be applied to zero values, with  $\text{IHS}(0)=0$ . Quarters with zero earnings are common in our data and, for that reason, we have adopted the IHS transformation of earnings for our analysis.<sup>4</sup>

The next few rows of Table 1 display information regarding the demographic characteristics of each of the different groups, as measured in the CPS as of  $q=0$ . These include age, education, race, gender, marital status and nativity. The differences in age across the different unemployment duration groups are especially notable. The proportion of people age

---

42.5% when computed using UI job data for 48 states. The effects of shifting from estimates based on UI data for 31 states to estimates based on UI data for 48 states are similarly small and very similar in magnitude for each of the other unemployment duration groups.

<sup>4</sup> Burbidge et.al.(1988) and Pence (2006) describe the advantages of the IHS transformation.

15-24 is much higher, and the proportion of people age 45 plus much lower, in the shorter unemployment duration groups than in the longer unemployment duration groups.

Finally, Table 1 reports average values for a set of variables intended to characterize individuals' job histories over the 12-quarter period  $q = \{-20, -9\}$ . These variables include the percent of the 12 quarters for which UI employment is recorded, average IHS(UI earnings) over the 12 quarters and, for the set of people for whom it can be calculated, average IHS(UI earnings) over the quarters with positive UI earnings. Consistent with the differences observed eight quarters prior to the CPS-LEHD link quarter, employment rates and earnings are consistently higher over the  $q = \{-20, -9\}$  period for those who are employed in the CPS at  $q=0$  than for those who are unemployed. The differences across unemployment duration groups are more modest, though there is some suggestion of stronger prior labor market attachment among those with longer unemployment durations as compared to those in the one-quarter unemployment duration group.

In addition, Table 1 displays estimates of the mean *trajectories* in employment rates and IHS(earnings) over the  $q = \{-20, -9\}$  period. These trajectories are estimated as the coefficient on a time trend variable in an individual-specific regression fit using the relevant set of observations—all twelve quarters for the employment rate and unconditional IHS(earnings) trajectories and quarters with positive earnings for the conditional-on-working IHS(earnings) trajectory. Estimation of the unconditional earnings trajectory requires at least one quarter of positive earnings and estimation of the conditional earnings trajectory requires at least two quarters of positive earnings; these trajectories are set to zero in cases where they cannot be estimated. The indicator variable “No Earnings Qtrs” takes the value of one for those with no UI earnings in any quarter over the  $q = \{-20, -9\}$  period; the indicator variable “<2 Earnings Qtrs” takes the value of one for those with no or just one quarter of UI earnings over the same period. There is little to indicate a systematic relationship between the estimated trajectories and how long a person has been unemployed, but those in the short-duration unemployment group are notably more likely not to have worked at all over the  $q = \{-20, -9\}$  period than those in any of the longer unemployment duration groups.<sup>5</sup>

### III. Employment and Earnings Profiles

A simple way to glean some initial insights into the labor market consequences of longer versus shorter spells of unemployment is to use our data to plot employment probabilities and measures of earnings for the time periods before and after the CPS-LEHD link quarter. Many of the conclusions to be drawn from the more formal analysis that follows can be previewed in these employment and earnings profiles.

---

<sup>5</sup> The negative trajectories for those in the group with six or more quarters of unemployment at  $q=0$  may be an artifact of some of these individuals having begun their spell of unemployment towards the end of the  $q = \{-20, -9\}$  period.

### *IIIa. Employment Profiles for the Full Matched Sample*

Figure 1 displays employment profiles based on our linked data for employed, NILF, and unemployed individuals. The horizontal axis in the figure is time period, normalized so that  $q=0$  is the quarter of the CPS-LEHD match (i.e., the quarter including the March in which the individual's labor force status is recorded in the CPS). Data for 2003 through 2010 have been pooled, so that  $q=0$  represents different calendar years for different people; in our more formal regression analyses, we estimate models that include year dummy controls. The points to the left of  $q=0$  along the horizontal axis are the 20 quarters preceding the quarter of the CPS-LEHD link  $q=\{-20,\dots,-1\}$  and those to the right are the 11 quarters following the quarter of the CPS-LEHD link  $q=\{1,\dots,11\}$ . The vertical axis in the figure is the weighted percentage of individuals with UI employment in the indicated quarter.

The solid line with filled in diamonds at the top of the figure shows the quarter-by-quarter UI employment probabilities for people who reported in the CPS that they were employed during the  $q=0$  March survey reference week. Of those reporting CPS employment in March, 79.9% had at least one UI job during the first quarter of that same year. One reason why this statistic is not 100% is that the UI data do not capture federal government employment or self-employment, meaning that someone with CPS employment of those types might not be recorded as employed in the UI data. In addition, from our previous research (Abraham et al 2013), we know that there are a significant share of CPS respondents who report CPS wage and salary employment in UI-covered sectors during the first quarter of the year but have no corresponding UI wage records in that quarter. The 79.9% UI employment rate for those classified as employed in the March CPS thus is generally in line with our priors.<sup>6</sup> Reflecting the fact that people may move into and out of employment, the UI employment probabilities for the  $q=0$  CPS employed resemble a shallow inverted “V” across time. Even in quarters other than  $q=0$ , however, the UI employment rates for the group classified as CPS employed at  $q=0$  consistently exceed 65%.

The dotted line at the bottom of Figure 1 shows the UI employment probabilities for people who report being NILF in the March CPS. Among those in this group, 9.5% have a UI-covered job at some point during the first quarter of the same year. We would not expect this percentage to be zero, since individuals who are working in January or February but not in March should be included here. Looking either backward or forward in time, the percentages of those NILF in  $q=0$  who have UI employment rises. Those who are NILF at  $q=0$  but employed in prior quarters will include some older people who have recently retired; those who are NILF at  $q=0$  but subsequently employed will include some younger people who are just entering the labor force. Because their experience is not obviously relevant for the questions in which we are interested, we have not incorporated the NILF in our subsequent analysis.

The most important information presented in Figure 1 is the pattern of UI employment probabilities for those who are classified as unemployed in the CPS. Separate employment

---

<sup>6</sup> In addition, someone who lives in one of our 31 CPS states could be employed in a jurisdiction not covered by our data. For example, a Maryland resident included in our 31 state sample might be employed in the District of Columbia or Virginia, jurisdictions for which we do not have UI data. As already discussed, however, this appears to have only a relatively minor effect on our estimates.

probability profiles are shown for those who had been unemployed up to 13 weeks (which we term the one quarter unemployed), 14-26 weeks (two quarter unemployed), 27-39 weeks (three quarter unemployed), 40-52 weeks (four quarter unemployed), 52-65 weeks (five quarter unemployed), or more than 65 weeks (six plus quarter unemployed) as of the March CPS reference week.

One initially surprising feature of these plots is the 10% to 20% employment rate in the  $q=0$  quarter among those who are classified in the March CPS as having been unemployed 14 or more weeks. Similar employment rates also are observed for prior quarters that, based on the duration of unemployment reported in the CPS, fall within the timespan of the individual's current unemployment spell (e.g., the 18.6% employment rate in  $q=-1$  for those in the three quarter unemployment group whose unemployment spells are reported to have started 27 or more weeks previously). A literal interpretation of what it means to be unemployed implies that no one in our sample should have done any work for pay during quarters falling wholly within the span of their unemployment spell.<sup>7</sup> Previous research has found, however, that reported unemployment durations do not always correspond strictly to the unemployment concept (see, for example, Clark and Summers 1979). Individuals who move from employment to unemployment or from out of the labor force to unemployment in the CPS commonly report initial unemployment durations of more than the 4 or 5 weeks elapsing between survey reference periods (Elsby, Hobijn and Sahin 2011, Kroft et al 2016). The positive UI employment rates during the timespan of reported unemployment may simply indicate that the employment in question is not considered by the respondent to be a "real" job. Consistent with this interpretation, the earnings on jobs held during the  $q=0$  quarter by individuals classified as unemployed in the CPS are substantially lower than the earnings on jobs held prior to the start of the same individuals' unemployment spells. Accepting that there will be some UI employment reported even during quarters in which no work is expected, the employment probabilities to the left of  $q=0$  seem very consistent with the unemployment durations recorded in the CPS, in the sense that, for each duration group, the number of quarters in the period prior to the CPS interview with a UI employment rate under 20% aligns precisely with the quarters of unemployment the group reports.

The economic questions motivating our paper relate to the subsequent labor market trajectories of those observed as unemployed in the CPS. Figure 1 makes clear that, consistent with the findings of KCC 2014, those with shorter unemployment spells are substantially more likely to be employed two to three years after their unemployment spell is recorded in the CPS than those with longer reported unemployment spells. Focusing on the period 8 to 11 quarters following the CPS link quarter, individuals with up to one quarter of unemployment (1 to 13 weeks) at  $q=0$  have an average employment rate of 56.6%, those with up to two quarters of unemployment (14-26 weeks) a 53.4% employment rate, those with up to three quarters of unemployment (27-39 weeks) a 51.4% employment rate, and those with up to four or five quarters of unemployment (40-52 or 53-65 weeks) a 42.3% employment rate. The transition

---

<sup>7</sup> Employment during the  $q=0$  quarter for those unemployed 13 weeks or less is not surprising, as these unemployment spells would have begun during that quarter. Similarly, employment during the  $q=-1$  quarter is not surprising for those unemployed 14-26 weeks, employment during the  $q=-2$  quarter is not surprising for those unemployed 27-39 weeks, and so on.

from unemployment to eventual employment also appears to occur somewhat more quickly for the short-term unemployed than for the long-term unemployed.

We are especially interested in the extent to which the substantial differences in post-spell outcomes across unemployment duration groups are driven by the differences in unemployment spell length (representing a causal effect) as opposed to being a result of heterogeneity across groups that pre-dated the start of the unemployment spell. Our approach to controlling for unobserved heterogeneity is to compare post-spell employment probabilities to pre-spell employment probabilities within duration groups. Figure 1 tells us that 55% to 60% of individuals with one to five quarters of unemployment at  $q=0$  were employed eight quarters earlier, before their spell of unemployment began. Further, for each of these unemployment duration groups, the probability of being employed was relatively steady with a slight upward drift over the several years before the start of the unemployment spell. Some of those with six or more quarters of unemployment could already have been unemployed even as far back as two years before the link quarter, but their employment rates three years before  $q=0$  are very similar to those of the other unemployment duration groups. In what follows, we focus mainly on the first five unemployment duration groups. The similarity in pre-spell employment rates across these duration groups stands in sharp contrast to their distinctly different post-spell rates. This is *prima facie* evidence that the differences in outcomes for those in the different unemployment duration groups are not simply a reflection of pre-existing heterogeneity in their employment propensities.

### *IIIb. Earnings Profiles*

In addition to learning about the pre-unemployment and post-unemployment employment profiles associated with short-term and long-term unemployment, we also are interested in the corresponding earnings profiles. We look separately at the path of UI earnings both overall (including people with zero earnings) and among those who succeed in finding a job (and thus have positive earnings). For groups defined on the basis of having CPS employment or belonging to one of the six unemployment duration groups as of the CPS link quarter, Figure 2 plots the mean of the inverse hyperbolic sine (IHS) of real UI earnings. Figure 2a includes everyone in our matched sample even if they had no earnings in a given quarter; Figure 2b includes only those people who have positive earnings in a given quarter, and thus speaks more directly to the magnitude of the earnings losses of the long-term unemployed relative to the short-term unemployed with respect to the earnings they eventually accept.

The profiles plotted in Figure 2a are very similar in appearance to the employment probability profiles we have already discussed. The profiles plotted in Figure 2b look quite different. Even among those with positive earnings, those who were employed at  $q=0$  have higher average IHS(real earnings) both before and after  $q=0$  than those in any of the unemployment by duration groups. Conditional on having a UI job, however, the difference in the earnings of the short-duration and long-duration unemployed are considerably smaller.<sup>8</sup>

---

<sup>8</sup> The top and bottom panels of Figure 2 have different scales along the vertical axis, chosen to span the values plotted in each. Appendix Figure A3 presents the top and bottom panels on consistent scales, meaning that a visual comparison of the panels shows directly how much of the earnings variation across duration groups is due to the extensive margin (employment versus non-employment) versus the intensive margin (earnings variation conditional

Note that there is an identity linking Figure 1 and the two panels of Figure 2. Since  $IHS(0)=0$ , the measure of earnings reported in Figure 2a for a given duration group and quarter is equal to the probability of being employed in that group and quarter in Figure 1 times the measure of earnings reported in Figure 2b for that same group and quarter. This identity along with the relatively small differences in Figure 2b compared to Figure 2a implies that most of the difference in earnings outcomes across unemployment duration groups apparent in Figure 2a occurs along the extensive margin (whether or not people are employed in a quarter) rather than along the intensive margin (how much they earn if they work). We explore this issue more systematically in the next section of the paper.

#### IV. Regression Results

Although there is a great deal to be learned from Figure 1 and Figure 2, visual inspection of the data can only take us so far. We turn now to a regression analysis that allows us to test for the statistical significance of observed pre-spell and post-spell differences across the duration groups, controlling for other factors that may differ across them. As noted, our methodology borrows from the related literature on job displacement, which grew out of a similar interest in learning about the effects of an event (in that case, job displacement) on subsequent labor market outcomes.

##### IVa. Estimation Framework

Rather than including all available quarters of data, the regression models we have estimated focus on employment and earnings eight quarters prior to and eight quarters following the CPS link quarter ( $q=0$ ). To help fix ideas, begin by considering the following simple specification:

$$Y_{idq} = \sum_{d=0}^{6+} \sum_{q=-8 \text{ or } +8} I_{idq} \delta_{dq} + \varepsilon_{idq} \quad (2)$$

where  $d = \{0, 1, 2, \dots, 5, 6+\}$  indexes the employment or unemployment duration group to which a person belongs, with  $d=0$  representing people who are CPS employed and  $d>0$  representing people with different unemployment durations in the CPS as of the CPS-LEHD link quarter;  $q=-8$  or  $+8$  indexes the quarter eight quarters before or eight quarters after the link quarter, and  $i$  indexes the person-year observation. We suppress calendar year time indices for convenience. Although this is not written out explicitly, we think of the  $\varepsilon_{idq}$  as potentially including both a fixed and a transitory individual component. Both of these components are assumed to have a mean value of zero for the target population as a whole, but there may be systematic differences in the mean value of the fixed component across duration groups. In the employment equations,  $Y_{idq}$  is a  $\{0, 1\}$  indicator of whether the person has one or more UI jobs in the quarter; in the

---

on employment). Note that Figure 2a is based on a balanced panel for each reference spell, whereas Figure 2b includes only those for a given reference spell with positive earnings in a particular quarter.

earnings equations, it is the person's IHS(real earnings) from any UI jobs held. The fourteen  $I_{idq}$  variables for any reference spell in period  $t$  are  $\{0,1\}$  indicators that take the value of 1 for the  $\{d,q\}$  combination pertaining to the given observation and 0 otherwise.<sup>9</sup>

Estimating equation (2) via OLS yields estimated  $\delta_{dq}$  value that are simply the mean values for  $q=-8$  and  $q=+8$  for each of the different duration groups as displayed in Figures 1 and 2. The differences across duration groups in the estimated  $\delta_{dq}$ 's for a given  $q$  may reflect both observable and unobservable heterogeneity. Controlling for observable heterogeneity is straightforward, with the modified specification given by:

$$Y_{idq} = \sum_{d=0}^{6+} \sum_{q=-8 \text{ or } +8} I_{idq} \delta_{dq} + X_i' \beta + \varepsilon_{idq} \quad (3)$$

where the  $X_i$  are individual characteristics defined as of  $q=0$  (and that thus do not vary over time for a given spell observed in progress in the CPS in year  $t$ ). These may include both demographic characteristics and job history variables such as the fraction of quarters over some prior period during which the person was employed. This is the specification most comparable to that of KCC 2014, though the models they estimate include only demographic controls and use data only for the post-unemployment period. Mimicking the KCC 2014 analysis, the differential effect on employment eight quarters later of being in any of the longer unemployment duration groups ( $N=2, 3, 4, 5$  or  $6+$ ) rather than the shortest duration group can be estimated as:

$$Diff_{N1} = \delta_{N,+8} - \delta_{1,+8} \quad (4)$$

We also want to allow, however, for the possibility of fixed but unobservable differences in individuals' underlying employment propensities. To do this, we focus on estimates of:

$$Loss_N = \delta_{N,+8} - \delta_{N,-8} \quad (5)$$

This loss is the change in the employment rate or in the mean of IHS(real earnings) from eight quarters before to eight quarters after  $q=0$  for a given unemployment duration group. Note that any fixed effects associated with either observable or unobservable individual characteristics that might have contributed to differences in the estimated  $\delta_{dq}$ 's across duration groups for a given  $q$  are differenced out in (5). In turn, the *difference* in these losses between any of the longer unemployment duration groups ( $N=2, 3, 4, 5$  or  $6+$ ) and the shortest duration group can be estimated as:

$$LossDiff_{N1} = (\delta_{N,+8} - \delta_{N,-8}) - (\delta_{1,+8} - \delta_{1,-8}) \quad (6)$$

---

<sup>9</sup> For example, for the  $q=-8$  observation for a person in unemployment duration group 3,  $I_{i3,-8}$  would equal 1 and all of the other  $I_{idq}$  variables would equal 0.

Equation (6) provides an estimate of the differential effect of being long-term unemployed rather than short-term unemployed on the outcome of interest that is uncontaminated by the influence of any time-invariant individual effects on that outcome.

As just described, the estimates of interest may be obtained by fitting (3) directly, using the (potentially biased)  $\delta_{dq}$  estimates to generate unbiased estimates of  $Loss_N$  as shown in (5) and then using those estimates to construct the  $LossDiff_N$ 's as shown in (6).<sup>10</sup> An alternative would be to difference equation (3) across  $q=+8$  and  $q=-8$  and then use the resulting equation to estimate the  $Loss_N$  values directly:

$$Y_{id,+8} - Y_{id,-8} = \sum_{d=0}^{6+} I_{id} (\delta_{d,+8} - \delta_{d,-8}) + (\varepsilon_{id,+8} - \varepsilon_{id,-8}) \quad (7)$$

where  $I_{id}$  is a dummy variable that equals 1 if the observations being differenced pertain to duration group  $d$  and otherwise equals 0. Observe that any time invariant individual controls are eliminated in this specification, along with the effects of any fixed but unobserved individual heterogeneity. It can be demonstrated that the two approaches are equivalent and will yield identical point estimates of the  $Loss_N$  and  $LossDiff_N$  effects of interest. The results we report are based on estimates of equation (3), but the interpretation of the specification shown in (7) may in some ways be more intuitive.

By focusing on the losses in (5) and differences in losses in (6) we are controlling for fixed factors affecting individuals' employment and earnings propensities. These estimates do not, however, control for possible differences in employment and earnings trajectories across individuals. Age is one possible source of differences in employment trajectories. Life cycle dynamics imply that we should expect upward sloping employment and average earnings profiles for those at younger ages, as more of them complete their education and enter the labor force, and downward sloping employment and average earnings profiles for older age groups, reflecting the gradual withdrawal of group members from the labor force. Figure A.4 in the appendix contains plots similar to Figure 1 but broken out by age group (less than 30 years, 30-49 years or 50 plus years). There is a clear upward slope to the measured employment probabilities for the under-30 age group that is not present in the figures for the two older age groups. To control for such factors we consider a modified version of equations (3) given by:

$$Y_{idq} = \sum_{d=0}^{6+} \sum_{q=-8 \text{ or } +8} I_{idq} \delta_{dq} + X'_i \beta_{-8} * Z_{-8} + X'_i \beta_{+8} * Z_{+8} + \varepsilon_{idq} \quad (3')$$

where  $Z_K$  is an indicator variable equal to one if  $q=K$  for the reference spell. The innovation in this specification is that it allows the observable characteristics to have different effects on employment in  $q=-8$  and  $q=+8$ . This makes it equivalent to the first-differenced specification:

---

<sup>10</sup> Although the coefficients estimated in (3) are subject to omitted variable bias stemming from the omission of time-invariant individual effects, given that we are working with a balanced panel for each reference year, this bias is of the same size and magnitude for  $\delta_{d,-8}$  and  $\delta_{d,+8}$ , meaning that it drops out when the difference  $\delta_{d,+8} - \delta_{d,-8}$  is taken.

$$Y_{id,+8} - Y_{id,-8} = \sum_{d=0}^{6+} I_{id} (\delta_{d+8} - \delta_{d-8}) + X_i' \beta_q + \eta_{id} \quad (7')$$

where  $\eta_{id} = \varepsilon_{id,+8} - \varepsilon_{id,-8}$  and  $\beta_q = \beta_{+8} - \beta_{-8}$ . In other words, allowing the change in employment probabilities across the period  $q=-8$  to  $q=+8$  to depend on the X variables as in (7') is equivalent to allowing for time-varying coefficients in (3'), yielding precisely the same  $Loss_N$  and  $LossDiff_{NI}$  estimates.<sup>11</sup> Note that, as in equation (7), any time invariant unobserved individual fixed effects again are eliminated from this specification.

As already noted, all of the included X's in these specifications are fixed across the two data points related to each person-year CPS observation. The demographic controls available to us include gender, age (15-24, 25-34, 35-44, 45-54, 55-64, 65+), education (<HS, HS, some college, college, >college), race (white, black, other), marital status, and foreign born, all defined as of the date of the CPS link quarter ( $q=0$ ). Models estimated with demographic controls also include year dummies. In addition, we make use of the previously-described job history and employment trajectory variables derived from the employment histories contained in the LEHD data infrastructure.

#### *IVb. Employment Equation Estimates*

Estimates of the  $\delta$  parameters from a set of linear probability models with an indicator for whether the person is employed as the dependent variable are reported in the top two panels of Table 2. The  $q=-8$  parameters for the employment and six unemployment duration groups are reported in the top panel of the table; the  $q=+8$  parameters for the same groups are reported in the second panel. The next three panels report the  $Diff_{NI}$ ,  $Loss_N$  and  $LossDiff_{NI}$  estimates computed from those  $\delta$  values. To conserve space, we do not report standard errors, but note that all the estimated  $\delta$  parameters are estimated with considerable precision. The model in column (1), based on specification (2), includes only dummy variables for duration group ( $d=0$  for the employed and  $d=1, 2, 3, 4, 5$  or  $6+$  for the unemployed) by quarter ( $q=-8$  or  $q=+8$ ). The models in columns (2) and (3), based on specification (3), add demographic controls and year effects (column (2)) and the share of quarters  $q=\{-20,-9\}$  during which the person had UI employment (column (3)). The models in columns (4), (5) and (6) are based on specification (3'). Column (4) includes demographic and year effects; column (5) adds the previously-described  $q=\{-20,-9\}$  job history variable; and column (6) adds the employment trajectory control also based on  $q=\{-20,-9\}$ . Recall that specification (3') differs from specification (3) in allowing the coefficients on the various control variables to differ across the pre-link and post-link quarters.<sup>12</sup>

<sup>11</sup> The point estimates of the loss and loss differentials are identical using either (3') or (7') but the standard errors are slightly different. In unreported results, we find that there are no differences in statistical inferences between the two approaches.

<sup>12</sup> In sensitivity analyses available upon request, we considered models that added the full complement of previously-described earnings history and earnings trajectory variables for the period  $q=\{-20,-9\}$  to the employment history and employment trajectory variables included in models (3), (5) and (6) of Table 2. This had very little effect on the estimated coefficients and did not change any of our qualitative conclusions.

The parameter estimates in column (1) of the top two panels of Table 2 simply reproduce the employment probability values for  $q=-8$  and  $q=+8$  that can be read off the plot displayed in Figure 1. In both  $q=-8$  and  $q=+8$ , the probability of UI employment for people who are CPS employed in the link quarter ( $q=0$ ) are substantially higher than the corresponding probabilities for people who are unemployed.<sup>13</sup> Except for the group unemployed six quarters or more, the  $q=-8$  employment probability differs little across the unemployment duration groups; at  $q=+8$ , however, the probability of having UI employment is notably lower for those with longer  $q=0$  unemployment durations. To illustrate, those unemployed five quarters at  $q=0$  have a UI employment rate in  $q=+8$  that is 13.2 percentage points lower than the UI employment rate for those unemployed just one quarter at  $q=0$  (an employment rate of 0.4167 versus 0.5489). In the third panel of the table, we report the differences in the  $\delta$  parameters at  $q=+8$  between each of the longer unemployment duration groups and the one-quarter unemployment duration group. As indicated by the gray shading, all of these differences are statistically significant at the 0.01 level.

The models reported in columns (2) and (3) of Table 2 contain additional controls but constrain their coefficients to be the same in  $q=-8$  and  $q=+8$ . The addition of these controls—demographic and year effects in column (2) and, especially, job history controls in column (3)—raises the estimated  $q=-8$  and  $q=+8$  employment probabilities for all of the unemployment duration groups. In other words, had the observable characteristics of the unemployed been the same as those of the average person in the linked sample, both their  $q=-8$  and their  $q=+8$  employment rates would have been somewhat higher.<sup>14</sup> We are more interested, however, in what the models that include these added controls imply about the difference in the impact of being long-term unemployed rather than short-term unemployed on subsequent employment.

As already discussed, one way to approach this question, similar to the analysis in KCC 2014, is simply to compare the estimated  $\delta_{d,+8}$  coefficients in the second panel of Table 2 across the different unemployment duration groups. Recall that, in the model with no controls, the estimated  $q=+8$  employment rate for the five-quarter-unemployed group at  $q=0$  was 0.4167 and that for the one-quarter-unemployed group was 0.5489, a difference of negative 13.2 percentage points. The corresponding numbers in the model that adds demographic and year controls are 0.4284 (five quarter duration) and 0.5748 (one quarter duration), a difference of negative 14.6 percentage points; in the model incorporating job history controls, the figures are 0.4839 and 0.6414, a difference of negative 15.8 percentage points. Each of these  $Diff_{NI}$  estimates (negative 13.2, 14.6 and 15.8 percentage points) is statistically different from zero. Controlling for observables actually leads to slight increases (in absolute value) in the point estimates of the estimated differences between the  $\delta_{d,+8}$ 's for the long-duration and the short-duration unemployment groups.

As already discussed, one potential problem with the  $Diff_{NI}$  estimates based simply on the  $\delta_{d,+8}$  coefficients is that these coefficients may be biased due to unobserved heterogeneity that has a differential effect on the different unemployment duration groups. To address this

<sup>13</sup> Tests not presented in Table 2 show that these differences are statistically significant.

<sup>14</sup> Because there are so many more employed people than unemployed people in our sample, the characteristics of the full sample are very similar to the characteristics of the employed population.

problem, we use estimates of both  $\delta_{d,+8}$  and  $\delta_{d,-8}$  to determine the employment *losses* experienced by each unemployment duration group, and then compare the losses in each of the longer duration groups to those for the shortest unemployment duration group. Except for the six-quarter-or-longer unemployment duration group (a group for whom results based on a comparison of  $q=-8$  to  $q=+8$  are somewhat difficult to interpret, as some may already have been unemployed at  $q=-8$ ), the long duration unemployed fare relatively worse in the  $LossDiff_{NI}$  calculations than in the  $Diff_{NI}$  calculations that simply compare the values of  $\delta_{d,+8}$  across the duration groups.

To illustrate, consider the estimated losses and differences in losses implied by the coefficient estimates reported in the second column of Table 2. Similar to KCC 2014, this model controls for demographics and year. Other than for the six-plus-quarter group, the loss differences are all somewhat larger than the simple differences in the  $\delta_{d,+8}$  values. For example, the decline in the employment rate for the five-quarter unemployment group between  $q=-8$  and  $q=+8$  is approximately 19.8 percentage points larger than that for the one-quarter unemployment group. This compares to the previously-noted difference of approximately 14.6 percentage points between the  $q=+8$  employment rates for the two groups and the two estimates of the relative disadvantage of being five-quarter rather than one-quarter unemployed are significantly different statistically.

One thing to note in passing is that, by construction, starting with a model that has no controls and then introducing time invariant control variables whose coefficients are constrained to be the same in both  $q=-8$  and  $q=+8$  will have no effect on the estimated employment losses or loss differences across the two periods. This can be seen in the  $Loss_N$ 's and  $LossDiff_{NI}$ 's reported in the first three columns at the bottom of Table, which are identical across the columns.

The models in columns (4), (5) and (6) of Table 2 allow for the coefficients on the various control variable coefficients to differ across the pre-link ( $q=-8$ ) and post-link ( $q=+8$ ) quarters. This has a large effect on several of our estimates. For example, for the individuals who have one quarter (1-13 weeks) of unemployment in the link quarter, the estimates in the three right-hand columns show sizeable losses in employment (a 3.7 percentage point loss in column (4), a 7.0 percentage point loss in column (5) and a 7.6 percentage point loss in column (6)), rather than the small gains shown in columns (1) through (3). This occurs because many of the short-duration unemployed are young adults—people who are more likely than the average to have been in school rather than working two years earlier and to be working two years later. Taking into account this underlying dynamic and its implication that, based on their demographics, the employment rates for the short-term unemployed should be higher in  $q=+8$  than in  $q=-8$  makes the post-unemployment outcomes of the short-term unemployed look relatively worse. As a result, the differences in losses between the long term unemployed and the short term unemployed are smaller (in absolute value) in the models that allow for time-varying coefficients on the control variables. In column (2) of Table 2, for example, the reported difference in losses between those who are unemployed five quarters (52-65 weeks) at the time of their CPS interview and those unemployed just one quarter (1-13 weeks) is 19.8 percentage points; in the corresponding model in column (4) that allows the coefficients on the same

included control variables to vary with time, this difference falls to 14.9 percentage points (a statistically significant decline).

In summary, the employment probability regression results in Table 2 quantify the patterns that were evident in Figure 1. First, individuals who report CPS-LEHD link quarter CPS employment are substantially more likely to have UI employment eight quarters prior to the link quarter than any of the unemployed groups. Further, the corresponding pre-unemployment UI employment rates are similar for persons with unemployment durations of one to five quarters. Eight quarters after the CPS-LEHD match, those who report being unemployed in the CPS have lower employment probabilities than eight quarters before, and in our preferred specification the losses for all of the unemployment duration groups are statistically significant.

Most important, our results provide strong evidence that having been unemployed longer at the time of the CPS interview is associated with poorer employment outcomes that cannot be explained either by time invariant unobservable heterogeneity or by controls for heterogeneity in employment trajectories. Our findings thus are consistent with negative duration dependence in the rate of exit to employment across a spell of unemployment—that is, with extended unemployment being the cause of the poorer subsequent employment outcomes experienced by the long-term unemployed.

Although our results permit us to dismiss observed and unobserved time-invariant heterogeneity across individuals as an explanation for the association between long-term unemployment and lower employment probabilities, it could be that the poorer job finding prospects of the long-term unemployed are a consequence of their having experienced especially negative labor market shocks rather than a result of their extended unemployment. A full exploration of this hypothesis is beyond the scope of the present paper, but examining data broken out by reason for employment leads us to think it is not likely to be a major part of the explanation for our findings. In Figure A.5 we repeat the exercise of Figure 1 separately for groups defined by reason for unemployment. Holding reason for unemployment constant is a crude way to control for differences in the shocks that individuals have experienced. Figure A.5 shows that, regardless of the duration of their unemployment, job losers fare the worst, suggesting that as a group they are hit with worse shocks than those in the other reason groups. Among job losers, however, it remains the case that employment losses are notably larger for the long term unemployed than for the short term unemployed. This pattern holds within all of the reason-for-unemployment groups. Further research is needed, but this evidence that there is a strong association between unemployment duration and subsequent employment outcomes even within reason-for-unemployment groups is at least suggestive that the differential losses for the long-term versus short-term unemployed are not being driven primarily by differences in the events leading to their unemployment.

#### IVc. Earnings Equation Estimates

Estimates of the  $\delta$  parameters from two sets of earnings models are reported in the top two panels of Table 3 and Table 4. The model in Table 3 (which corresponds to Figure 2a) is fit using data for everyone in the sample whether they have earnings in a given quarter or not; in quarters with no UI earnings, a zero is assumed. The model in Table 4 (which corresponds to Figure 2b) uses only those observations for which positive UI earnings are reported in the quarter. In both cases, as already described, the dependent variable is IHS(real earnings). The structure of the tables reporting on the earnings models is the same as the structure of the employment probability table just discussed. The only difference in model specification is that in columns (3), (5) and (6), employment history and employment trajectory variables based on  $q=\{-20,-9\}$  are replaced with the earnings history and earnings trajectory variables that correspond to the model's dependent variable. In column (6), the indicator variable for whether the earnings trajectory measure could be constructed also is included.<sup>15</sup>

The estimates that appear in Table 3 are in many respects very similar to those in Table 2. Looking at the full linked sample, at  $q=-8$ , the IHS(real earnings) of those who will be employed in the CPS at  $q=0$  are significantly higher than those for any of the groups that will be unemployed in  $q=0$ . The point estimates of the coefficients for the  $q=-8$  measure of earnings across the first five unemployment duration groups are generally similar. At  $q=+8$ , it continues to be the case that those who were employed in the CPS at  $q=0$  have higher IHS(real earnings) than those in any of the unemployment duration groups, but now there is a systematic decline in the measure of earnings with the duration of unemployment at  $q=0$ . Looking at the  $Loss_N$  statistics in column (1), the quarterly earnings of employed individuals rise on average between  $q=-8$  and  $q=+8$  and quarterly earnings for the group unemployed 1-13 weeks as of the link quarter do not change significantly. In contrast, quarterly earnings for all of the longer-term unemployment groups exhibit a significant decline and the magnitude of these earnings losses increases monotonically with unemployment duration through the group with five quarters of unemployment at the link quarter; the losses of those unemployed six plus quarters are smaller because that group was doing poorly even at  $q=-8$ .

Controls for demographics and year are added in column (2); those variables plus a measure of the individual's earnings history over the period  $q=\{-20,-9\}$  appear in column (3). In these models, the coefficients on the added variables are constrained to be the same in  $q=-8$  and  $q=+8$ . The same variables are added in column (4) and column (5) but with their coefficients permitted to differ across the two periods. Column (6) controls for the  $q=\{-20,-9\}$  earnings trajectory as previously described, along with an indicator variable for those for whom this trajectory could not be constructed. As before, by construction, adding fixed controls with coefficients constrained to be equal across the two time periods may raise or lower the level of the estimated  $\delta$ 's but by the same amount in both  $q=-8$  and  $q=+8$ , so that the estimated  $Loss_N$ 's are not affected. This explains why the losses reported in column (2) and column (3) are

---

<sup>15</sup> In sensitivity analyses available upon request, we considered models that added all of the employment history and employment trajectory for the period  $q=\{-20,-9\}$  to the earnings history and earnings trajectory variables corresponding to the dependent variable that are included in the models reported in columns (3), (5) and (6) of Table 3 and Table 4. This had very little effect on the estimated coefficients and did not change any of our qualitative conclusions.

identical to those reported in column (1). To affect the estimated losses, the coefficients on these variables must be allowed to differ across time periods so that they can pick up differences in the earnings trajectories associated with observable characteristics that might be correlated with unemployment duration. When this is done in columns (4), (5), and (6) we observe larger estimated earnings losses for several of the unemployment duration groups (those with one to four quarters of unemployment as of the date of the CPS-LEHD match). The implication is that, given the characteristics of the members of these groups, their earnings would have been expected to grow and the declines actually experienced between  $q=-8$  and  $q=+8$  look worse relative to this expected-growth benchmark than relative to the no-change benchmark implicit in column (1).

Table 4 repeats this exercise conditional on positive earnings in  $q=-8$  and  $q=+8$ . We undertake this exercise to take advantage of the identity noted above in our discussion of Figures 1 and 2. The following identity relates the top two panels of Tables 2, 3 and 4:

$$\delta_{dq}^3 = \delta_{dq}^2 * \delta_{dq}^4 \quad (8)$$

where the numbered superscripts reference the table number. That is, for duration group  $d$  and quarter  $q$ , the estimated overall mean of IHS(real earnings) in Table 3 denoted by  $\delta_{dq}^3$  is equal to the product of the estimated employment probability from the same cell in Table 2 (denoted by  $\delta_{dq}^2$ ) and the estimated mean of IHS(earnings) for those who are working from the same cell in Table 4 (denoted by  $\delta_{dq}^4$ ). This identity permits a decomposition (described in detail below) of the earnings losses in Table 3 into the portions of the loss occurring along the extensive and the intensive margins.<sup>16</sup>

One caution in interpreting the results in Table 4 in isolation is that the estimation of equation (3) and (3') is no longer based on a balanced panel of observations for individuals who have an unemployment spell in a given reference year. The reason is that the dependent variable is contingent on positive earnings in a given quarter. As such, the estimated losses from taking the differences in the top two panels of Table 2 are no longer equivalent to the losses estimated from first difference specifications (7) and (7)'. With this caveat in mind, we note that the qualitative patterns of Table 4 are similar to those in Table 3 but the quantitative impacts are substantially reduced. Even with smaller magnitudes, the estimated earnings loss effects are still substantial. For example, in column 6 of Table 4 in the bottom panel, among those who have earnings, the differential earnings loss between  $q=-8$  and  $q=+8$  for the long term (5 quarter) unemployed compared to the short term (one quarter) unemployed is approximately 16 ln points (given that the differences in the IHS values are approximately ln differences at the values reported in Table 4).

The identity in (8) permits an exact decomposition of the earnings losses in Table 3:

---

<sup>16</sup> The identity in equation (8) is exact for column (1) of Tables 2, 3 and 4. The addition of control variables to the models reported in those tables breaks the exact identity, but the resulting "error" is small. Defining error as the absolute value of the difference between the  $\delta$  estimates shown in Table 3 and those calculated using equation (8), this error never exceeds five percent of the actual  $\delta$  estimates in Table 3.

$$\delta_{a,+8}^3 - \delta_{a,-8}^3 = (\delta_{a,+8}^2 - \delta_{a,-8}^2) * \delta_{am}^4 + (\delta_{a,+8}^4 - \delta_{a,-8}^4) * \delta_{am}^2 \quad (9)$$

where  $\delta_{am}^4 = (\delta_{a,+8}^4 + \delta_{a,-8}^4)/2$  and  $\delta_{am}^2 = (\delta_{a,+8}^2 + \delta_{a,-8}^2)/2$ . The first term on the right hand side of (9) can be interpreted as the earnings loss accounted for by the loss in employment probabilities as estimated in Table 2 and the second term as the earnings loss accounted for by the lower earnings among those who are employed as estimated in Table 4. This decomposition can be done for each duration group which in turn implies that the differential loss in earnings can be decomposed into the extensive component (differences in the first term on the right hand side of (9) across duration groups) and intensive component (differences in the second term on the right hand side of (9) across duration groups). Exploiting this decomposition, we find that the differential between the long term and short term unemployed in their earnings losses is accounted for mainly by the extensive margin. For example, applying this decomposition using column (1) of Tables 2, 3, and 4 we find that about 92% of the differential loss between the long-term (5 quarter) and short-term (1 quarter) unemployed is accounted for by the extensive margin. For column (6), the equivalent calculation is about 99%. We conclude that the primary driving force behind the earnings losses shown in Figure 2a and Table 3—and the greater magnitude of those earnings losses for the longer-term unemployed—is the lower probability of employment in  $q=+8$  rather than the reduction in what unemployed persons earn when they eventually do get a job.

## V. Concluding Remarks

The sharp rise in long term unemployment in the Great Recession and the slow accompanying recovery have heightened concerns about the distressingly bad labor market outcomes experienced by the long-term unemployed. Efforts to form a clear understanding of the reasons for this group's poor success in the labor market have long been plagued, however, by the challenge of distinguishing the effects of state dependence from those of unobserved heterogeneity, in particular the possibility that the long-term unemployed are individuals whose employment probabilities and earnings outcomes would have been worse than those of the short-term unemployed even absent their experience of extended unemployment.

We use integrated survey and administrative data to overcome this challenge. From the CPS, we identify currently-unemployed individuals who have been unemployed for different lengths of time. We integrate these survey data with matched employer-employee data from the LEHD data infrastructure. This enables us to generate employment and earnings profiles for those in a current spell of unemployment over an interval from 20 quarters prior to 11 quarters after the date they are observed as unemployed in the CPS.

Using these integrated survey and administrative data, we evaluate the impact of the duration of unemployment through the lens of differences in the pre-spell and post-spell employment and earnings outcomes. More specifically, our linked data enable us to control for fixed effects attributable to individual characteristics measured as of the date an unemployed person is observed in the CPS as well as fixed effects attributable to unobserved but unchanging individual characteristics. For a given group of individuals, any such effects should impact both pre-spell and the post-spell employment and earnings outcomes in the same way, allowing us to

draw meaningful inferences from the differences in the pre-spell and post-spell outcomes for a given duration group. In turn, we consider the differences in these differences across duration groups. We find that, even after accounting for differences in their observed and unobserved characteristics, the longer-term unemployed experience substantially worse employment and earnings losses than the short term unemployed.

This fixed effects approach does not control for different employment trajectories across individuals that might be driving differences across unemployment duration groups. Given our integrated survey and administrative data set, we are able to control for differences in trajectories associated with observable characteristics and, at least to some extent, for differences in trajectories observable in each individual's own history. Controlling for such trajectories somewhat mitigates the differences in employment and earnings losses between the long term and short termed unemployed, but they remain substantial.

These results allow us to rule out the “bad apple” explanation for why the long-term unemployed fare worse in the labor market than the short-term unemployed and are consistent with duration dependence as the explanation for their poorer outcomes. It should be acknowledged that this is not the only possible explanation for what we see. In particular, it is possible that differences in the shocks experienced by the long term and the short term unemployed could play some role in these differences. The fact, however, that we see very similar patterns in our data even when the sample is restricted to those in a particular reason-for-unemployment group (e.g. those who lost a previous job) suggests to us that differences in the shocks affecting the long-term and the short-term unemployed are unlikely to be the whole explanation for what we see.

Using much the same approach as just described for studying the effect of long-term unemployment on subsequent employment, we also examine the impact of unemployment duration on subsequent earnings. We document a significant adverse impact of long term unemployment as compared to short-term unemployment on what individuals later earn. Most of this effect is realized through the extensive margin, that is, through its effect on the probability of being employed. The long-term unemployed do experience greater earning losses conditional on being employed, but the large majority of the adverse impact of long-term unemployment on earnings takes place along the extensive margin.

An interesting question is whether the adverse effects of long-term unemployment on subsequent employment and earnings vary with labor market conditions. To the extent that the adverse effects of long-term unemployment reflect employers' interpretation of such unemployment as a negative signal, one might expect these effects to be weaker when the labor market is weaker. In a weak labor market, many people cannot easily find work, so that having been out of work for an extended period might be a less negative signal than in a stronger labor market. In some preliminary analyses that are not reported in the paper, we explored whether our findings varied across years in which the national labor market could be characterized as stronger or weaker. In those analyses, although the point estimates of the  $LossDiff_{NI}$ 's were generally consistent with this speculation, the differences across years we characterized as “good years” and those we characterized as “bad years” were not statistically significant. Whether the impact of being long-term unemployed varies with labor market conditions is, however, a

question that merits further investigation. This could be done by extended the analysis to use additional years of data as they become available or, perhaps more promising, exploiting the variation in labor market conditions across states in the data we already have.

Another important avenue for future research will be to develop a better understanding of the relative importance of the specific factors that may be contributing to duration dependence in the effects of unemployment on later employment and earnings—including human capital depreciation during spells away from work, decreases in search intensity over the spell of unemployment and employer discrimination against the long-term unemployed in hiring, among other possibilities. There also would be value in a more complete exploration of the possible role of differential shocks as a contributor to the more negative outcomes experienced by the long-term unemployed.

Our findings highlight the advantages of integrated survey and administrative data for questions related to the consequences of passing through a spell of unemployment. There is no obvious way to identify and track spells of unemployment (as opposed to non-employment) in administrative data. The primary source for information on unemployment is the CPS, but the CPS contains limited information on the experiences of unemployed workers prior to or following their unemployment spell. Integration of administrative data such as that from the LEHD data infrastructure with the CPS records permits the construction of employment and earnings histories that greatly enhance what can be said about the consequences of being unemployed.

## References

- Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in *Producer Dynamics*, Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts, eds., Chicago: University of Chicago Press, 149-230.
- Acemoglu, Daron and Robert Shimer. 2000. "Productivity gains from unemployment insurance," *European Economic Review* 44, 1195-1224.
- Abraham, Katharine G., John Haltiwanger, Kristin Sandusky, and James R. Spletzer. 2013. "Exploring Differences in Household vs. Establishment Measures of Employment," *Journal of Labor Economics* 31(2, pt. 2), S129-S172.
- Acemoglu, Daron. 1995. "Public Policy in a Model of Long-term Unemployment," *Economica* 62, 161-78.
- Albrecht, James W., Per-Anders Edin, Marianne Sundstrom and Susan B. Vroman. 1999. "Career Interruptions and Subsequent Earnings: A Reexamination Using Swedish Data," *Journal of Human Resources* 34(2), 294-311.
- Alvarez, Fernando, Katarina Borovickova and Robert Shimer. 2015. "Decomposing Duration Dependence in a Stopping Time Model," unpublished working paper. March.
- Burbidge, John, Lonnie Magee and A. Leslie Robb. 1988. "Alternative Transformations to Handle Extreme Values of the Dependent Variable" *Journal of the American Statistical Association* 83(401), 123-127.
- Clark, Kim B. and Lawrence H. Summers. 1979. "Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity*, No. 1, 13-60.
- Couch, Kenneth A. and Dana W. Placzek. 2010. "Earnings Losses of Displaced Workers Revisited," *American Economic Review* 100(1), 572-589.
- Cripps, T.F. and R.J. Tarling. 1974. "An Analysis of the Duration of Male Unemployment in Great Britain 1932-73," *Economic Journal* 84(334), 289-316.
- Davis, Steven J. and Till von Wachter. 2011. "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity*, Fall, 1-55.
- Ehrenberg, Ronald G. and Ronald L. Oaxaca. 1976. "Unemployment Insurance, Duration of Unemployment, and Subsequent Wage Gain," *American Economic Review* 66(5), 754-766.
- Elsby, Michael, Bart Hobijn and Ayşegül Şahin. 2011. "The Labor Market in the Great Recession: An Update to September 2011," *Brookings Papers on Economic Activity*, Spring, 1-69.

- Eriksson, Stefan and Dan-Olof Rooth. 2011. "Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment," IZA Discussion Paper No. 6235, December.
- Faberman, R. Jason and Marianna Kudlyak. 2014. "The Intensity of Job Search and Search Duration," Working Paper 14-12, Federal Reserve Bank of Richmond.
- Farber, Henry S. 1993. "The Incidence and Costs of Job Loss: 1982-1991," *Brookings Papers on Economic Activity: Microeconomics*, 73-132.
- Farber, Henry S. 1997. "The Changing Face of Job Loss in the United States, 1981-1995," *Brookings Papers on Economic Activity: Microeconomics*, 55-128.
- Farber, Henry S. 2005. "What do we know about Job Loss in the United States? Evidence from the Displaced Workers Survey, 1981-2004," *Economic Perspectives*, Federal Reserve Bank of Chicago, Second Quarter, 13-28
- Farber, Henry S. 2013. "Job Loss: Historical Perspective from the Displaced Workers Survey." in *Unexpected Lifecycle Events and Economic Security: the Roles of Job Loss, Disability, and Changing Family Structure*, Kenneth A. Couch, Mary C. Daly, and Julie Zissimopoulos, eds., Stanford University Press, 11-33.
- Farber, Henry S. 2015. "Job Loss in the Great Recession and Its Aftermath: U.S. Evidence from the Displaced Workers Survey," Working Paper No. 589, Industrial Relations Section, Princeton University.
- Farber, Henry S., Dan Silverman and Till von Wachter. 2015. "Factors Determining Callbacks to Job Applications by the Unemployed: An Audit Study," NBER Working Paper No. 21689, October.
- Ghayad, Rand. 2013. "The Jobless Trap," unpublished working paper.
- Gorlich, Dennis and Andries de Grip. 2009. "Human Capital Depreciation During Hometown," *Oxford Economic Papers* 61(Supplement 1), 98-121.
- Heckman, James, and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 52(2), 271-320.
- Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers," *The American Economic Review* 83(4), 685-709.
- Kaitz, Hyman. 1970. "Analyzing the Length of Spells of Unemployment," *Monthly Labor Review* 93(11), 11-20.

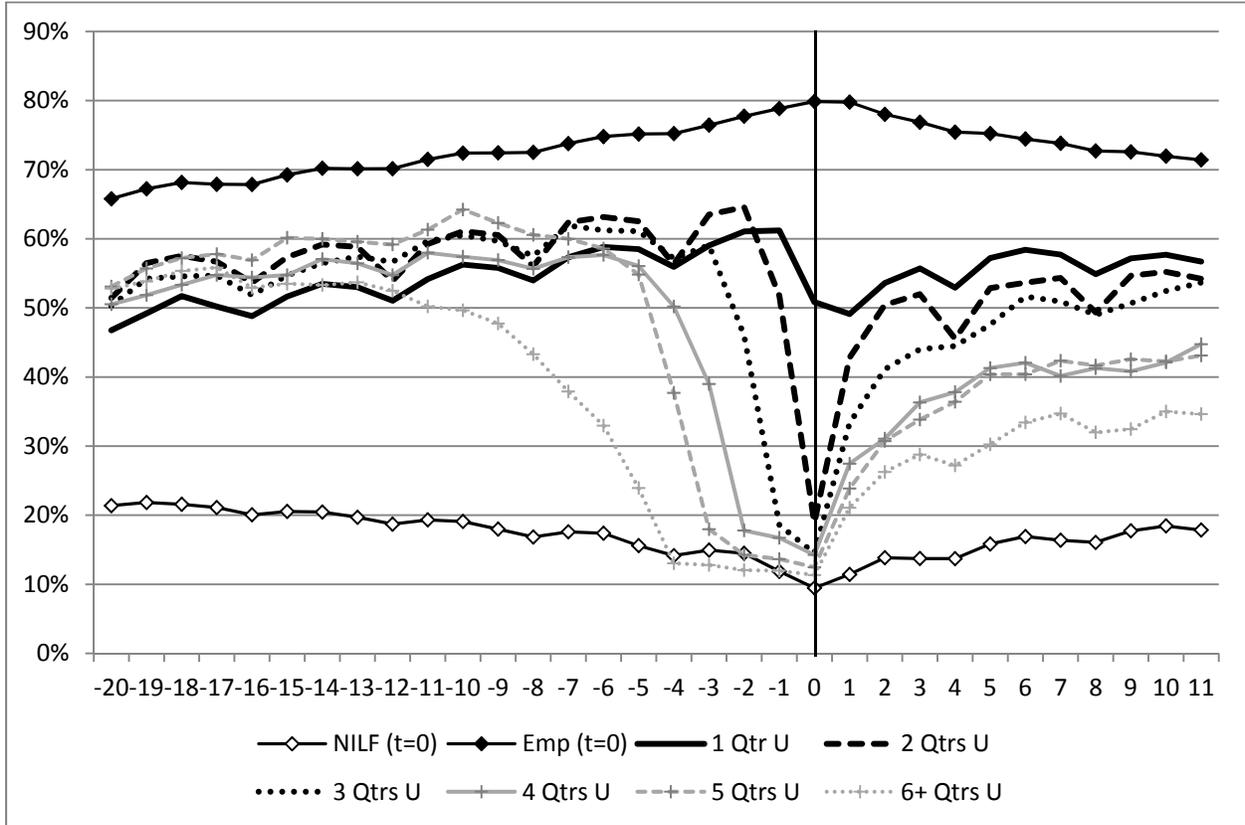
- Kroft, Kory, Fabian Lange and Matthew J. Notowidigdo. 2013. "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment." *Quarterly Journal of Economics* 128(3), 1123-1167.
- Kory Kroft, Fabian Lange, Matthew J. Notowidigdo and Lawrence F. Katz. 2016. "Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Nonparticipation," *Journal of Labor Economics* 34(Suppl. 1), S7-S54.
- Krueger, Alan B., Judd Cramer, and David Cho. 2014. "Are the Long-Term Unemployed on the Margins of the Labor Market?" *Brookings Papers on Economic Activity*, Spring, 229-280.
- Krueger, Alan B. and Andreas Mueller. 2011. "Job Search, Emotional Well-Being and Job Finding in a Period of Mass Unemployment: Evidence from High-frequency Longitudinal Data," *Brookings Papers on Economic Activity*, Spring, 1-57.
- Machin, Stephen and Alan Manning. 1999. "The Causes and Consequences of Longterm Unemployment in Europe," in *Handbook of Labor Economics: Volume 3*, O. Ashenfelter and D. Card, eds., 3085-3139.
- Mincer, Jacob and Haim Ofek. 1982. "Interrupted Work Careers: Depreciation and Restoration of Human Capital," *Journal of Human Resources* 17(1), 3-24.
- Pence, Karen. 2006. "The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Savings," *Contributions to Economic Analysis and Policy* 5(1), 1-20.
- Schmieder, Johannes F., Till von Wachter, and Stefan Bender. 2014. "The Causal Effect of Unemployment Duration on Wages: Evidence from Unemployment Insurance Extensions," unpublished working paper.
- Stratton, Leslie S. 1995. "The Effect Interruptions in Work Experience Have on Wages," *Southern Economic Journal* 61(4), 955-970.
- Von Wachter, Till, Jae Song, and Joyce Manchester. 2011. "Long-Term Earnings Losses Due to Mass Layoffs during the 1982 Recession: An Analysis Using Longitudinal Administrative Data from 1974 to 2008," unpublished working paper.

**Table 1: Weighted Sample Characteristics**

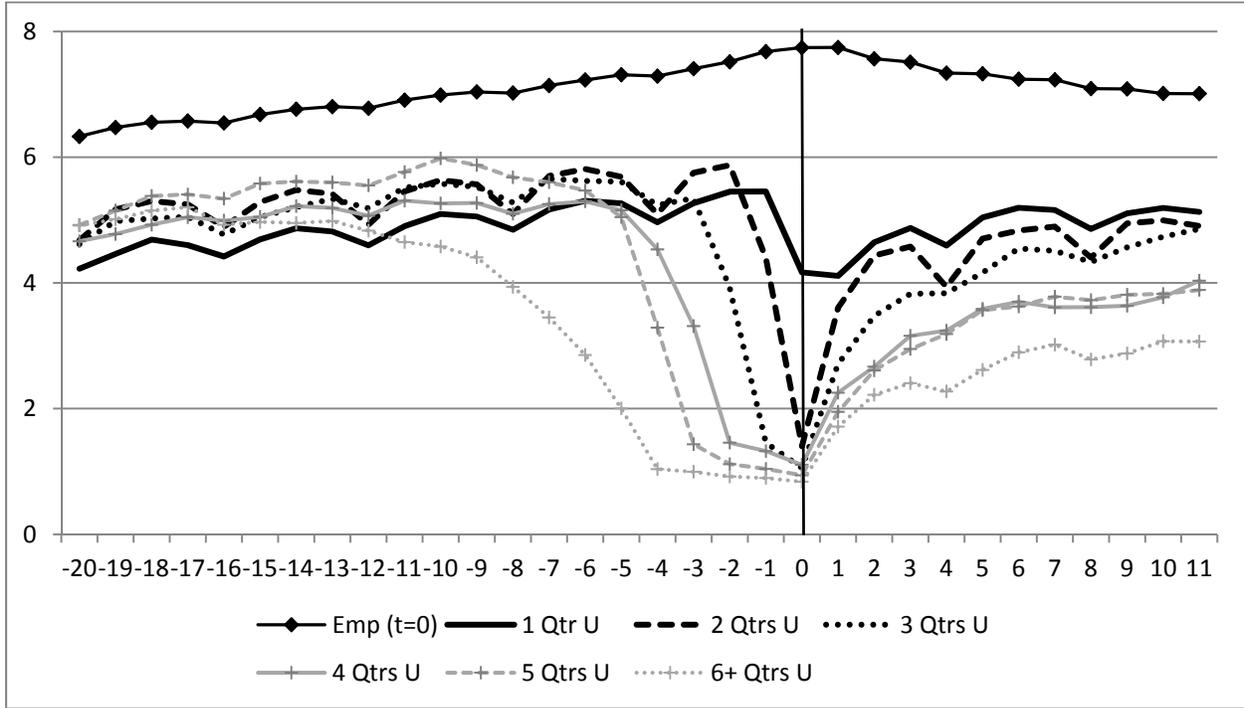
	Employed (in CPS) at q=0	Unemp. 1 Qtr at q=0	Unemp. 2 Qtrs at q=0	Unemp. 3 Qtrs at q=0	Unemp. 4 Qtrs at q=0	Unemp. 5 Qtrs at q=0	Unemp. 6+ Qtrs at q=0
<u>Work Outcome q=-8</u>							
UI employment	.7251	.5396	.5601	.5751	.5566	.6057	.4330
UI real \$ ( $\geq 0$ )	9,181	4,031	4,729	4,995	4,593	5,948	3,501
UI real \$ ( $> 0$ )	12,661	7,471	8,443	8,686	8,253	9,820	8,086
IHS UI real \$ ( $\geq 0$ )	7.021	4.847	5.097	5.279	5.100	5.679	3.936
IHS UI real \$ ( $> 0$ )	9.683	8.984	9.101	9.180	9.163	9.377	9.091
<u>Work Outcome q=8</u>							
UI employment	.7272	.5489	.4933	.4899	.4128	.4167	.3198
UI real \$ ( $\geq 0$ )	9,746	3,382	3,166	3,078	2,356	2,842	1,764
UI real \$ ( $> 0$ )	13,403	6,161	6,420	6,282	5,709	6,821	5,517
IHS UI real \$ ( $\geq 0$ )	7.092	4.859	4.403	4.333	3.613	3.725	2.783
IHS UI real \$ ( $> 0$ )	9.753	8.852	8.927	8.845	8.754	8.940	8.701
<u>CPS {0,1} Variables</u>							
Age 15-24 yrs	.1306	.3244	.2443	.2755	.2006	.1554	.1362
Age 25-34 yrs	.2170	.2377	.2342	.1943	.2371	.1971	.1739
Age 35-44 yrs	.2427	.1870	.2113	.1912	.1986	.2263	.2238
Age 45-54 yrs	.2367	.1493	.1904	.2133	.2055	.2668	.2377
Age 55+ yrs	.1730	.1016	.1198	.1256	.1581	.1544	.2285
Education <12 yrs	.1090	.2668	.2461	.2278	.2275	.1929	.2115
Education =12 yrs	.2871	.3415	.3806	.3593	.3794	.3764	.3483
Education 13-15 yrs	.2912	.2637	.2355	.2596	.2558	.2818	.2678
Education $\geq 16$ yrs	.3127	.1280	.1378	.1532	.1373	.1490	.1724
White	.8274	.7605	.7430	.7196	.6982	.7234	.6554
Male	.5307	.5668	.6069	.6029	.6012	.6113	.5944
Married	.5847	.3705	.3804	.3425	.3905	.3822	.5847
Foreign Born	.1678	.1615	.1814	.1639	.1747	.1926	.2238
<u>UI Job History</u>							
Employment rate	.6942	.5184	.5716	.5589	.5501	.5895	.5261
IHS UI real \$ ( $\geq 0$ )	6.704	4.703	5.255	5.153	5.065	5.514	4.882
IHS UI real \$ ( $> 0$ )	9.454	8.709	8.863	8.877	8.849	8.959	8.913
<u>UI Job Trajectory</u>							
Employment rate	.0057	.0070	.0055	.0078	.0056	.0077	-.0050
IHS UI real \$ ( $\geq 0$ )	.0698	.0821	.0657	.0945	.0649	.0940	-.0659
IHS UI real \$ ( $> 0$ )	.0267	.0269	.0412	.0156	.0264	.0342	-.0286
<u>UI {0,1} Variables</u>							
No Earnings Qtrs.	.1604	.2456	.1889	.2104	.2112	.1762	.2185
<2 Earnings Qtrs.	.1837	.2926	.2325	.2460	.2574	.2314	.2666
Sample Size	267,627	9,943	3,743	1,423	1,199	776	1,030

Note: Sample is 285,741 person-level observations for individuals who are employed or unemployed at q=0 when interviewed for the Current Population Survey (CPS) and for whom a PIK is available for linking to the LEHD. UI job history, job trajectory and {0,1} variables are defined over the time period q=[-20,-9]. Real earnings are in 2012:Q2 dollars using the CPI-U-RS as a deflator.

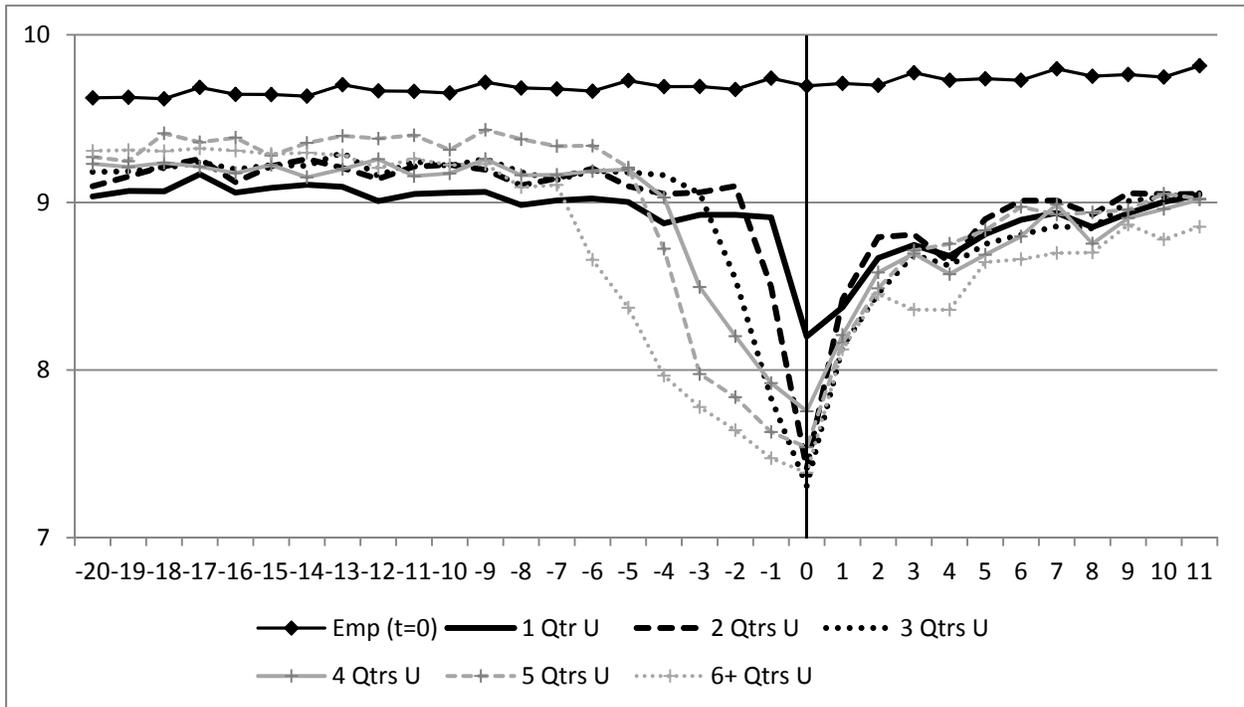
**Figure 1: Employment Probabilities**



**Figure 2a: Real Earnings (IHS Earnings<sub>≥0</sub>)**



**Figure 2b: Real Earnings (IHS Earnings<sub>>0</sub>)**



**Table 2: Employment Probability Regressions**

	(1)	(2)	(3)	(4)	(5)	(6)
$\delta$ parameters (q=-8)						
$\delta_{d=0,q=-8}$	.7251	.7235	.7192	.7224	.7169	.7163
$\delta_{d=1,q=-8}$	.5396	.5654	.6320	.5886	.6715	.6810
$\delta_{d=2,q=-8}$	.5601	.5796	.6263	.5923	.6504	.6599
$\delta_{d=3,q=-8}$	.5751	.5960	.6476	.6104	.6746	.6781
$\delta_{d=4,q=-8}$	.5566	.5751	.6457	.5801	.6680	.6746
$\delta_{d=5,q=-8}$	.6057	.6174	.6729	.6160	.6852	.6815
$\delta_{d=6,q=-8}$	.4330	.4559	.5519	.4519	.5714	.6105
$\delta$ parameters (q=+8)						
$\delta_{d=0,q=+8}$	.7272	.7256	.7212	.7267	.7234	.7232
$\delta_{d=1,q=+8}$	.5489	.5748	.6414	.5516	.6019	.6047
$\delta_{d=2,q=+8}$	.4933	.5128	.5595	.5001	.5354	.5382
$\delta_{d=3,q=+8}$	.4899	.5109	.5624	.4965	.5354	.5365
$\delta_{d=4,q=+8}$	.4128	.4313	.5019	.4263	.4796	.4815
$\delta_{d=5,q=+8}$	.4167	.4284	.4839	.4297	.4717	.4706
$\delta_{d=6,q=+8}$	.3198	.3428	.4388	.3468	.4193	.4307
q=+8 Employment Differences						
Diff21: $\delta_{d=2,q=+8} - \delta_{d=1,q=+8}$	<b>-0.557</b>	<b>-0.620</b>	<b>-0.819</b>	<b>-0.515</b>	<b>-0.665</b>	<b>-0.665</b>
Diff31: $\delta_{d=3,q=+8} - \delta_{d=1,q=+8}$	<b>-0.590</b>	<b>-0.640</b>	<b>-0.789</b>	<b>-0.551</b>	<b>-0.665</b>	<b>-0.682</b>
Diff41: $\delta_{d=4,q=+8} - \delta_{d=1,q=+8}$	<b>-1.362</b>	<b>-1.435</b>	<b>-1.395</b>	<b>-1.253</b>	<b>-1.223</b>	<b>-1.232</b>
Diff51: $\delta_{d=5,q=+8} - \delta_{d=1,q=+8}$	<b>-1.323</b>	<b>-1.464</b>	<b>-1.575</b>	<b>-1.219</b>	<b>-1.303</b>	<b>-1.341</b>
Diff61: $\delta_{d=6,q=+8} - \delta_{d=1,q=+8}$	<b>-2.291</b>	<b>-2.320</b>	<b>-2.206</b>	<b>-2.049</b>	<b>-1.826</b>	<b>-1.740</b>
Employment Losses						
Loss0: $\delta_{d=0,q=+8} - \delta_{d=0,q=-8}$	.0020	.0020	.0020	<b>.0043</b>	<b>.0065</b>	<b>.0069</b>
Loss1: $\delta_{d=1,q=+8} - \delta_{d=1,q=-8}$	.0094	.0094	.0094	<b>-.0370</b>	<b>-.0696</b>	<b>-.0763</b>
Loss2: $\delta_{d=2,q=+8} - \delta_{d=2,q=-8}$	<b>-.0668</b>	<b>-.0668</b>	<b>-.0668</b>	<b>-.0922</b>	<b>-.1150</b>	<b>-.1218</b>
Loss3: $\delta_{d=3,q=+8} - \delta_{d=3,q=-8}$	<b>-.0851</b>	<b>-.0851</b>	<b>-.0851</b>	<b>-.1139</b>	<b>-.1391</b>	<b>-.1416</b>
Loss4: $\delta_{d=4,q=+8} - \delta_{d=4,q=-8}$	<b>-.1438</b>	<b>-.1438</b>	<b>-.1438</b>	<b>-.1539</b>	<b>-.1884</b>	<b>-.1930</b>
Loss5: $\delta_{d=5,q=+8} - \delta_{d=5,q=-8}$	<b>-.1890</b>	<b>-.1890</b>	<b>-.1890</b>	<b>-.1863</b>	<b>-.2135</b>	<b>-.2110</b>
Loss6: $\delta_{d=6,q=+8} - \delta_{d=6,q=-8}$	<b>-.1132</b>	<b>-.1132</b>	<b>-.1132</b>	<b>-.1052</b>	<b>-.1522</b>	<b>-.1798</b>
Employment Loss Differences						
LossDiff21: loss2 – loss1	<b>-.0762</b>	<b>-.0762</b>	<b>-.0762</b>	<b>-.0552</b>	<b>-.0455</b>	<b>-.0455</b>
LossDiff31: loss3 – loss1	<b>-.0945</b>	<b>-.0945</b>	<b>-.0945</b>	<b>-.0769</b>	<b>-.0696</b>	<b>-.0654</b>
LossDiff41: loss4 – loss1	<b>-.1532</b>	<b>-.1532</b>	<b>-.1532</b>	<b>-.1169</b>	<b>-.1189</b>	<b>-.1167</b>
LossDiff51: loss5 – loss1	<b>-.1984</b>	<b>-.1984</b>	<b>-.1984</b>	<b>-.1493</b>	<b>-.1439</b>	<b>-.1347</b>
LossDiff61: loss6 – loss1	<b>-.1225</b>	<b>-.1226</b>	<b>-.1225</b>	<b>-.0682</b>	<b>-.0826</b>	<b>-.1035</b>
R-squared	.0131	.0377	.3600	.0475	.3891	.4389
CPS Demographics controls		Yes	Yes	Yes	Yes	Yes
Year dummy controls		Yes	Yes	Yes	Yes	Yes
Employment history controls			Yes		Yes	Yes
Employment trajectory controls						Yes
Coefficient vector $\beta$ or $\beta(q)$	$\beta$	$\beta$	$\beta$	$\beta(q)$	$\beta(q)$	$\beta(q)$

Sample is 285,741 persons who are employed or unemployed in CPS at q=0 and for whom a PIK is available for linking to the LEHD. Number of observations in regression is 571,482. All control variables are deviations from means. Dependent variable is {0,1} employment indicator; mean=.7130. All estimated  $\delta$  parameters are statistically different from zero. Bold with grey shading in the lower three panels indicates reported estimate is statistically significant at the .01 level.

**Table 3: Earnings Regressions (IHS, Earnings $\geq$ 0)**

	(1)	(2)	(3)	(4)	(5)	(6)
$\delta$ parameters (q=-8)						
$\delta_{d=0,q=-8}$	7.021	6.993	6.944	6.982	6.922	6.918
$\delta_{d=1,q=-8}$	4.847	5.320	6.054	5.549	6.443	6.517
$\delta_{d=2,q=-8}$	5.097	5.448	5.999	5.568	6.238	6.264
$\delta_{d=3,q=-8}$	5.279	5.647	6.230	5.785	6.495	6.481
$\delta_{d=4,q=-8}$	5.100	5.422	6.218	5.463	6.432	6.443
$\delta_{d=5,q=-8}$	5.679	5.881	6.514	5.858	6.629	6.508
$\delta_{d=6,q=-8}$	3.936	4.249	5.311	4.195	5.488	5.786
$\delta$ parameters (q=+8)						
$\delta_{d=0,q=+8}$	7.092	7.064	7.015	7.075	7.036	7.035
$\delta_{d=1,q=+8}$	4.859	5.332	6.066	5.103	5.678	5.704
$\delta_{d=2,q=+8}$	4.403	4.754	5.305	4.635	5.066	5.080
$\delta_{d=3,q=+8}$	4.333	4.702	5.285	4.564	5.020	5.020
$\delta_{d=4,q=+8}$	3.613	3.935	4.731	3.894	4.517	4.527
$\delta_{d=5,q=+8}$	3.725	3.927	4.560	3.950	4.445	4.414
$\delta_{d=6,q=+8}$	2.783	3.096	4.157	3.150	3.980	4.084
q=+8 Earnings Differences						
Diff21: $\delta_{d=2,q=+8} - \delta_{d=1,q=+8}$	<b>-0.456</b>	<b>-0.578</b>	<b>-0.761</b>	<b>-0.468</b>	<b>-0.612</b>	<b>-0.624</b>
Diff31: $\delta_{d=3,q=+8} - \delta_{d=1,q=+8}$	<b>-0.526</b>	<b>-0.631</b>	<b>-0.782</b>	<b>-0.539</b>	<b>-0.658</b>	<b>-0.684</b>
Diff41: $\delta_{d=4,q=+8} - \delta_{d=1,q=+8}$	<b>-1.246</b>	<b>-1.397</b>	<b>-1.335</b>	<b>-1.209</b>	<b>-1.161</b>	<b>-1.177</b>
Diff51: $\delta_{d=5,q=+8} - \delta_{d=1,q=+8}$	<b>-1.135</b>	<b>-1.405</b>	<b>-1.506</b>	<b>-1.153</b>	<b>-1.232</b>	<b>-1.289</b>
Diff61: $\delta_{d=6,q=+8} - \delta_{d=1,q=+8}$	<b>-2.076</b>	<b>-2.237</b>	<b>-1.909</b>	<b>-1.953</b>	<b>-1.697</b>	<b>-1.620</b>
Earnings Losses						
Loss0: $\delta_{d=0,q=+8} - \delta_{d=0,q=-8}$	<b>0.071</b>	<b>0.071</b>	<b>0.071</b>	<b>0.093</b>	<b>0.114</b>	<b>0.116</b>
Loss1: $\delta_{d=1,q=+8} - \delta_{d=1,q=-8}$	0.012	0.012	0.012	<b>-0.446</b>	<b>-0.766</b>	<b>-0.813</b>
Loss2: $\delta_{d=2,q=+8} - \delta_{d=2,q=-8}$	<b>-0.694</b>	<b>-0.694</b>	<b>-0.694</b>	<b>-0.934</b>	<b>-1.173</b>	<b>-1.184</b>
Loss3: $\delta_{d=3,q=+8} - \delta_{d=3,q=-8}$	<b>-0.946</b>	<b>-0.946</b>	<b>-0.946</b>	<b>-1.222</b>	<b>-1.475</b>	<b>-1.461</b>
Loss4: $\delta_{d=4,q=+8} - \delta_{d=4,q=-8}$	<b>-1.486</b>	<b>-1.486</b>	<b>-1.486</b>	<b>-1.569</b>	<b>-1.915</b>	<b>-1.916</b>
Loss5: $\delta_{d=5,q=+8} - \delta_{d=5,q=-8}$	<b>-1.954</b>	<b>-1.954</b>	<b>-1.954</b>	<b>-1.909</b>	<b>-2.184</b>	<b>-2.094</b>
Loss6: $\delta_{d=6,q=+8} - \delta_{d=6,q=-8}$	<b>-1.153</b>	<b>-1.153</b>	<b>-1.153</b>	<b>-1.046</b>	<b>-1.507</b>	<b>-1.702</b>
Earnings Loss Differences						
LossDiff21: loss2 – loss1	<b>-0.706</b>	<b>-0.706</b>	<b>-0.706</b>	<b>-0.487</b>	<b>-0.407</b>	<b>-0.371</b>
LossDiff31: loss3 – loss1	<b>-0.958</b>	<b>-0.958</b>	<b>-0.958</b>	<b>-0.775</b>	<b>-0.710</b>	<b>-0.649</b>
LossDiff41: loss4 – loss1	<b>-1.498</b>	<b>-1.498</b>	<b>-1.498</b>	<b>-1.122</b>	<b>-1.149</b>	<b>-1.104</b>
LossDiff51: loss5 – loss1	<b>-1.966</b>	<b>-1.966</b>	<b>-1.966</b>	<b>-1.462</b>	<b>-1.418</b>	<b>-1.281</b>
LossDiff61: loss6 – loss1	<b>-1.165</b>	<b>-1.165</b>	<b>-1.165</b>	<b>-0.599</b>	<b>-0.742</b>	<b>-0.889</b>
R-squared	.0178	.0577	.4167	.0686	.4446	.4957
CPS Demographics controls		Yes	Yes	Yes	Yes	Yes
Year dummy controls		Yes	Yes	Yes	Yes	Yes
Earnings history controls			Yes		Yes	Yes
Earnings trajectory controls						Yes
Coefficient vector $\beta$ or $\beta(q)$	$\beta$	$\beta$	$\beta$	$\beta(q)$	$\beta(q)$	$\beta(q)$

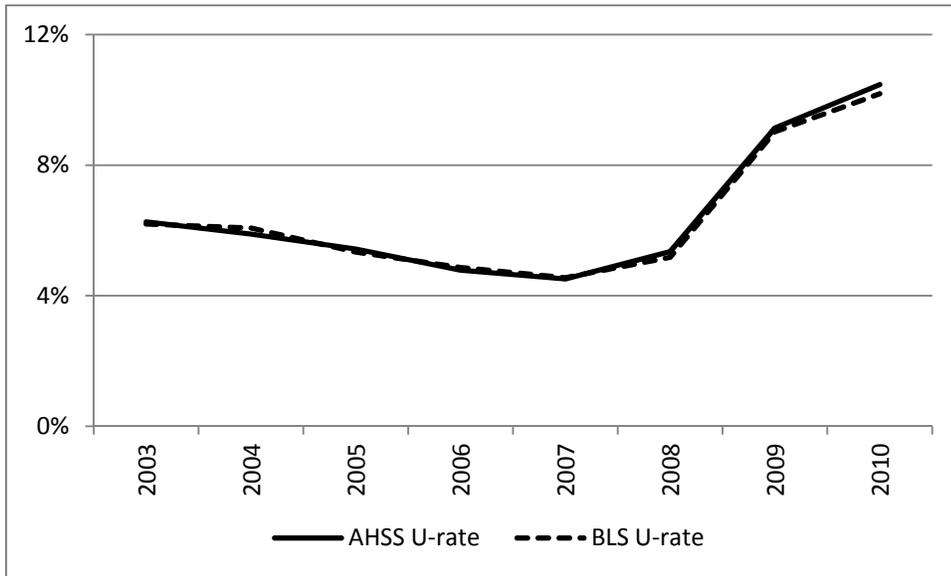
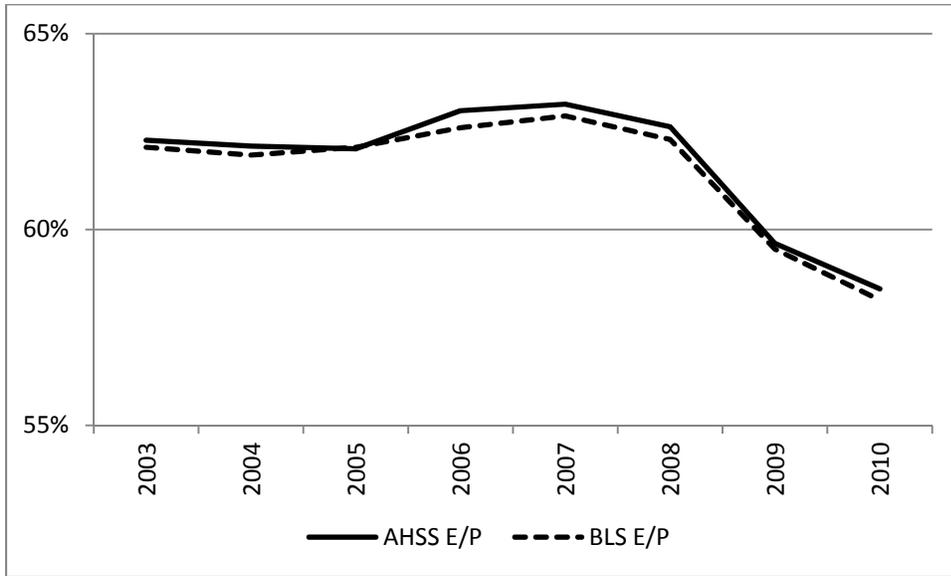
Sample is 285,741 persons who are employed or unemployed in CPS at q=0 and for whom a PIK is available for linking to the LEHD. Number of observations in regression is 571,482. All control variables are deviations from means. Dependent variable is IHS(Real Quarterly Earnings  $\geq$ 0); mean=6.903. Inverse Hyperbolic Sine is defined as  $IHS(x)=\ln\{x+\sqrt{1+(x*x)}\}$ . All estimated  $\delta$  parameters are statistically different from zero. Bold with grey shading in the lower three panels indicates reported estimate is statistically significant at the .01 level.

**Table 4: Earnings Regressions (IHS, Earnings>0)**

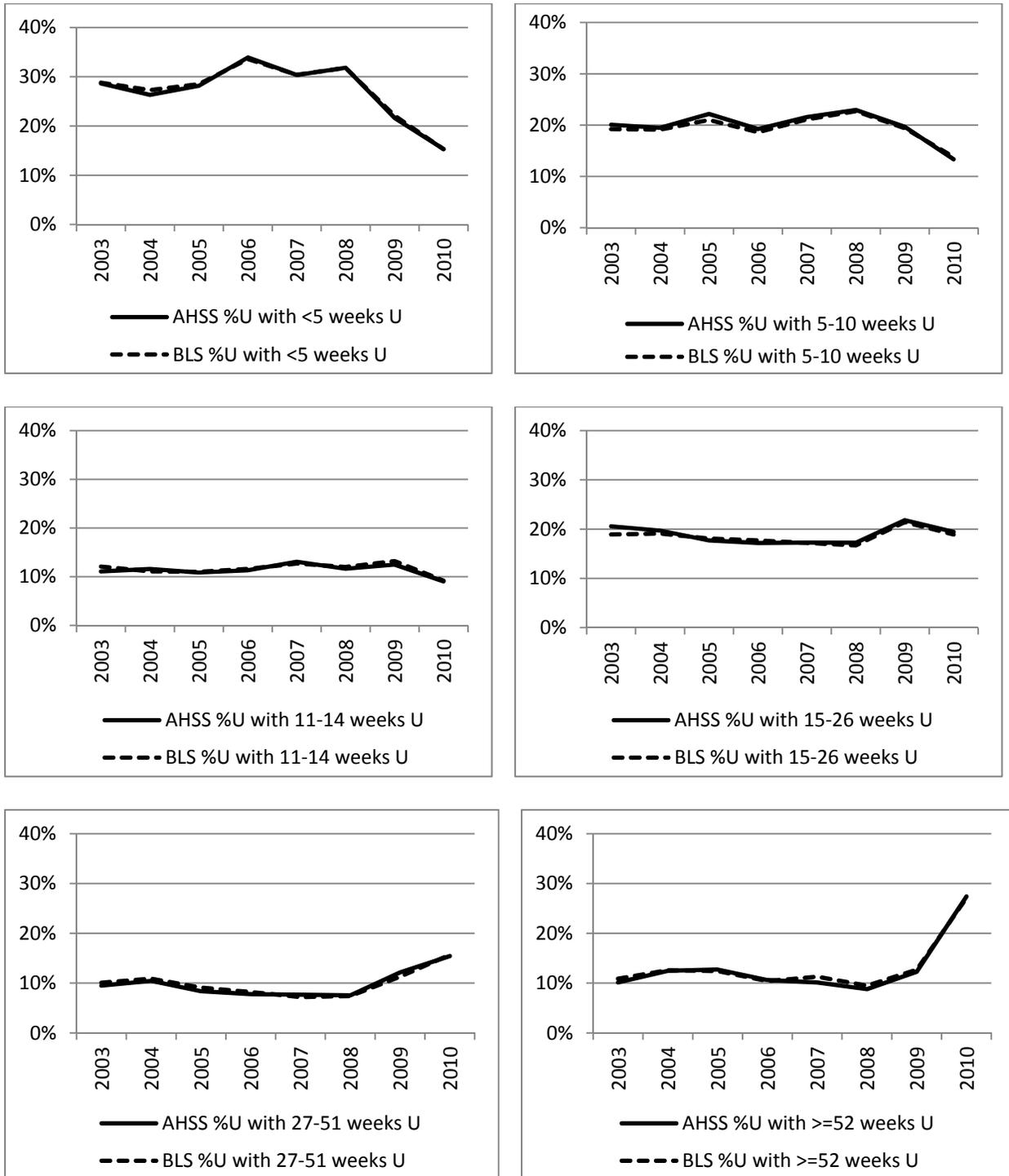
	(1)	(2)	(3)	(4)	(5)	(6)
$\delta$ parameters (q=-8)						
$\delta_{d=0,q=-8}$	9.683	9.652	9.619	9.671	9.659	9.659
$\delta_{d=1,q=-8}$	8.984	9.217	9.388	9.271	9.511	9.518
$\delta_{d=2,q=-8}$	9.101	9.270	9.415	9.302	9.509	9.504
$\delta_{d=3,q=-8}$	9.180	9.314	9.451	9.340	9.536	9.540
$\delta_{d=4,q=-8}$	9.162	9.300	9.462	9.312	9.539	9.535
$\delta_{d=5,q=-8}$	9.377	9.455	9.538	9.461	9.595	9.581
$\delta_{d=6,q=-8}$	9.091	9.183	9.307	9.182	9.365	9.376
$\delta$ parameters (q=+8)						
$\delta_{d=0,q=+8}$	9.753	9.758	9.773	9.740	9.733	9.733
$\delta_{d=1,q=+8}$	8.852	9.228	9.437	9.162	9.312	9.313
$\delta_{d=2,q=+8}$	8.927	9.220	9.420	9.175	9.316	9.313
$\delta_{d=3,q=+8}$	8.845	9.137	9.292	9.086	9.195	9.199
$\delta_{d=4,q=+8}$	8.754	9.010	9.207	8.972	9.114	9.113
$\delta_{d=5,q=+8}$	8.940	9.106	9.307	9.078	9.220	9.221
$\delta_{d=6,q=+8}$	8.701	8.847	9.113	8.831	9.023	9.020
q=+8 Earnings Differences						
Diff21: $\delta_{d=2,q=+8} - \delta_{d=1,q=+8}$	<b>.075</b>	- .008	- .016	.012	.004	.000
Diff31: $\delta_{d=3,q=+8} - \delta_{d=1,q=+8}$	- .007	- .091	<b>-.145</b>	- .076	<b>-.117</b>	<b>-.114</b>
Diff41: $\delta_{d=4,q=+8} - \delta_{d=1,q=+8}$	- .099	<b>-.219</b>	<b>-.230</b>	<b>-.190</b>	<b>-.198</b>	<b>-.200</b>
Diff51: $\delta_{d=5,q=+8} - \delta_{d=1,q=+8}$	.087	- .122	<b>-.130</b>	- .084	- .092	- .092
Diff61: $\delta_{d=6,q=+8} - \delta_{d=1,q=+8}$	- .151	<b>-.381</b>	<b>-.324</b>	<b>-.331</b>	<b>-.289</b>	<b>-.292</b>
Earnings Losses						
Loss0: $\delta_{d=0,q=+8} - \delta_{d=0,q=-8}$	<b>.070</b>	<b>.105</b>	<b>.153</b>	<b>.069</b>	<b>.073</b>	<b>.074</b>
Loss1: $\delta_{d=1,q=+8} - \delta_{d=1,q=-8}$	- .132	.011	<b>.049</b>	- .109	- .199	- .205
Loss2: $\delta_{d=2,q=+8} - \delta_{d=2,q=-8}$	- .174	- .050	.005	- .128	- .193	- .192
Loss3: $\delta_{d=3,q=+8} - \delta_{d=3,q=-8}$	- .335	- .177	- .159	- .255	- .341	- .341
Loss4: $\delta_{d=4,q=+8} - \delta_{d=4,q=-8}$	- .409	- .290	- .255	- .340	- .425	- .423
Loss5: $\delta_{d=5,q=+8} - \delta_{d=5,q=-8}$	- .437	- .348	- .231	- .383	- .375	- .361
Loss6: $\delta_{d=6,q=+8} - \delta_{d=6,q=-8}$	- .390	- .336	- .195	- .351	- .342	- .356
Earnings Loss Differences						
LossDiff21: loss2 – loss1	- .043	- .061	- .043	- .019	.006	.014
LossDiff31: loss3 – loss1	<b>-.203</b>	<b>-.188</b>	<b>-.207</b>	<b>-.146</b>	<b>-.142</b>	<b>-.136</b>
LossDiff41: loss4 – loss1	<b>-.277</b>	<b>-.301</b>	<b>-.304</b>	<b>-.231</b>	<b>-.226</b>	<b>-.217</b>
LossDiff51: loss5 – loss1	<b>-.306</b>	<b>-.359</b>	<b>-.280</b>	<b>-.274</b>	<b>-.176</b>	<b>-.156</b>
LossDiff61: loss6 – loss1	<b>-.258</b>	<b>-.347</b>	<b>-.243</b>	<b>-.243</b>	<b>-.143</b>	<b>-.151</b>
R-squared	.0238	.2687	.5125	.2801	.5419	.5496
CPS Demographics controls		Yes	Yes	Yes	Yes	Yes
Year dummy controls		Yes	Yes	Yes	Yes	Yes
Earnings history controls			Yes		Yes	Yes
Earnings trajectory controls						Yes
Coefficient vector $\beta$ or $\beta(q)$	$\beta$	$\beta$	$\beta$	$\beta(q)$	$\beta(q)$	$\beta(q)$

Sample is 235,691 persons who are employed or unemployed in CPS at q=0, with LEHD employment observations at q=-8 and/or q=+8. Number of observations in regression is 406,200. All control variables are deviations from means. Dependent variable is IHS(Real Quarterly Earnings >0); mean=9.682. Inverse Hyperbolic Sine is defined as  $IHS(x)=\ln\{x+\sqrt{1+(x^2)}\}$ . All estimated  $\delta$  parameters are statistically different from zero. Bold with grey shading in the lower three panels indicates reported estimate is statistically significant at the .01 level.

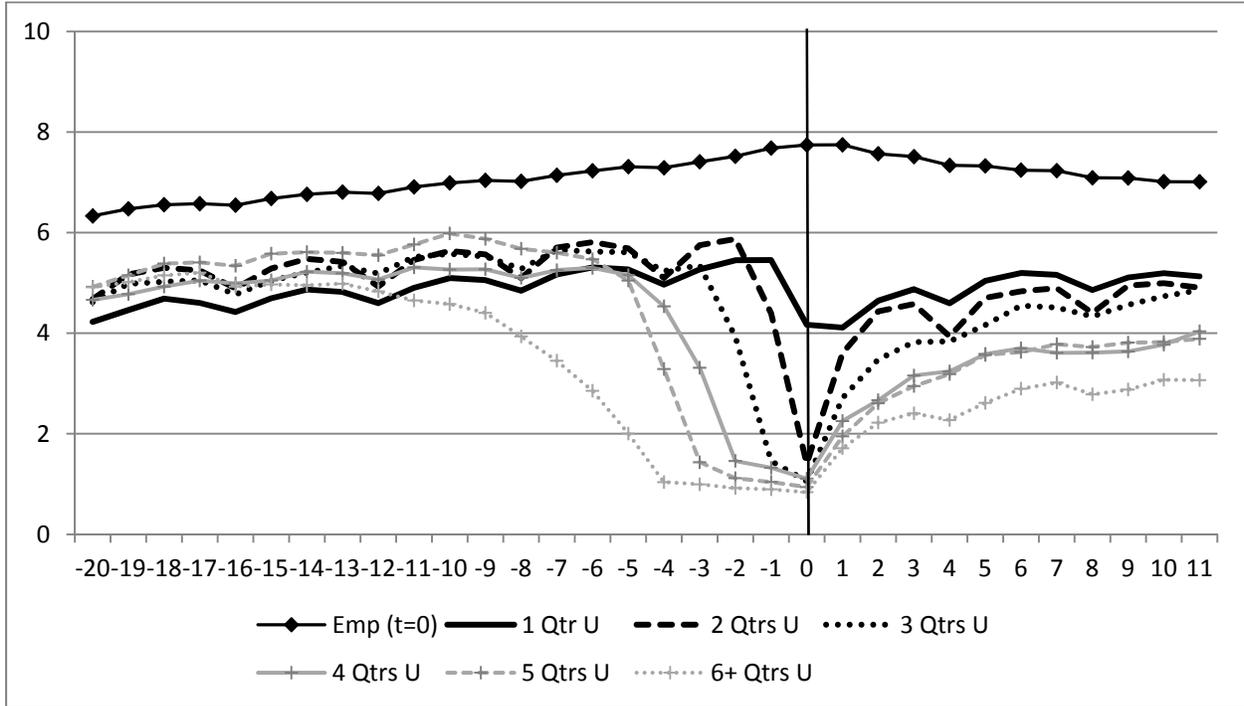
**Figure A.1: Employment-to-Population Rate and Unemployment Rate, 2003-2010**  
**CPS-LEHD weighted microdata (“AHSS”)**  
**Tabulations from BLS website**



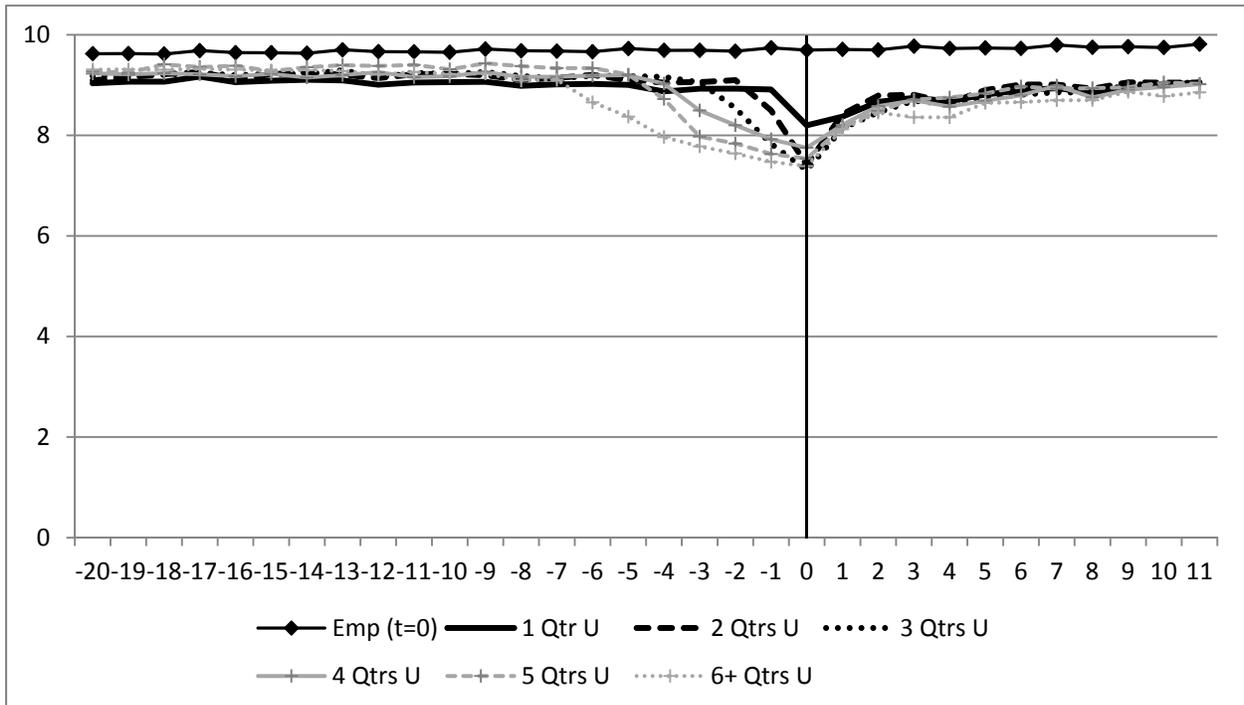
**Figure A.2: Duration of Unemployment, 2003-2010**  
**CPS-LEHD weighted microdata (“AHSS”)**  
**Tabulations from BLS website**



**Figure A.3a: Real Earnings (IHS Earnings<sub>≥0</sub>)**

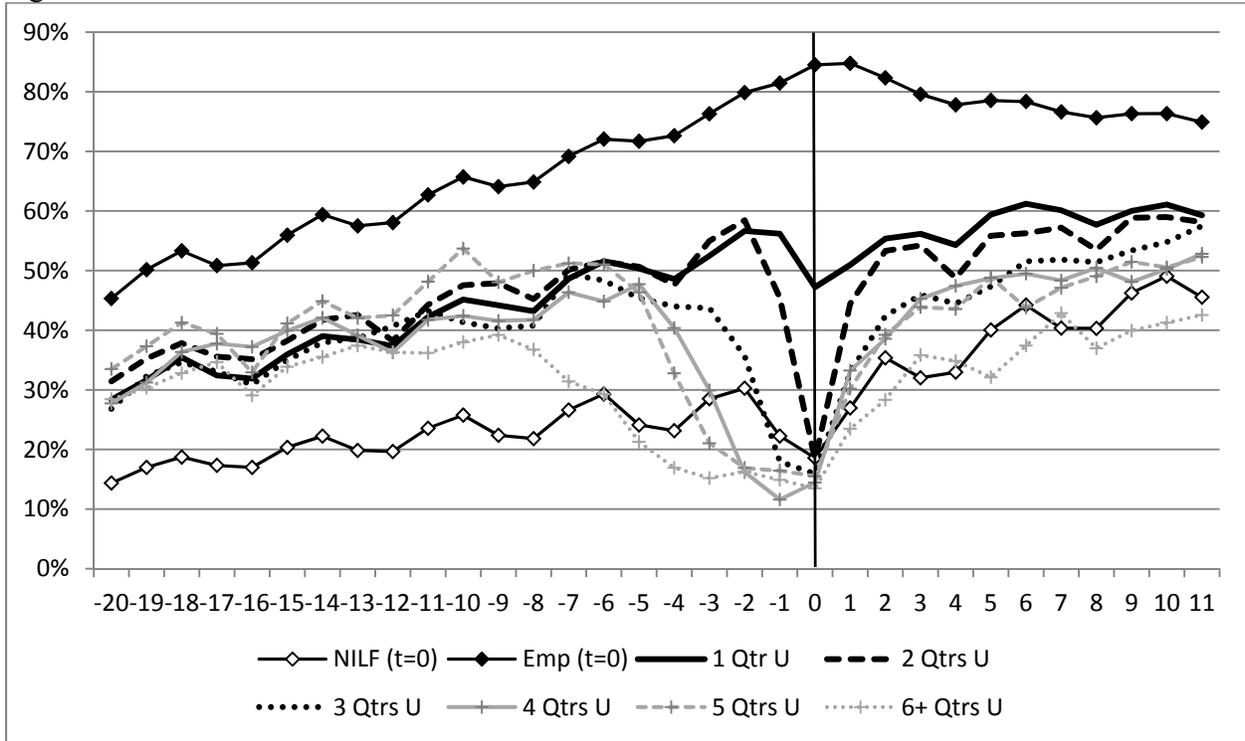


**Figure A.3b: Real Earnings (IHS Earnings<sub>>0</sub>)**

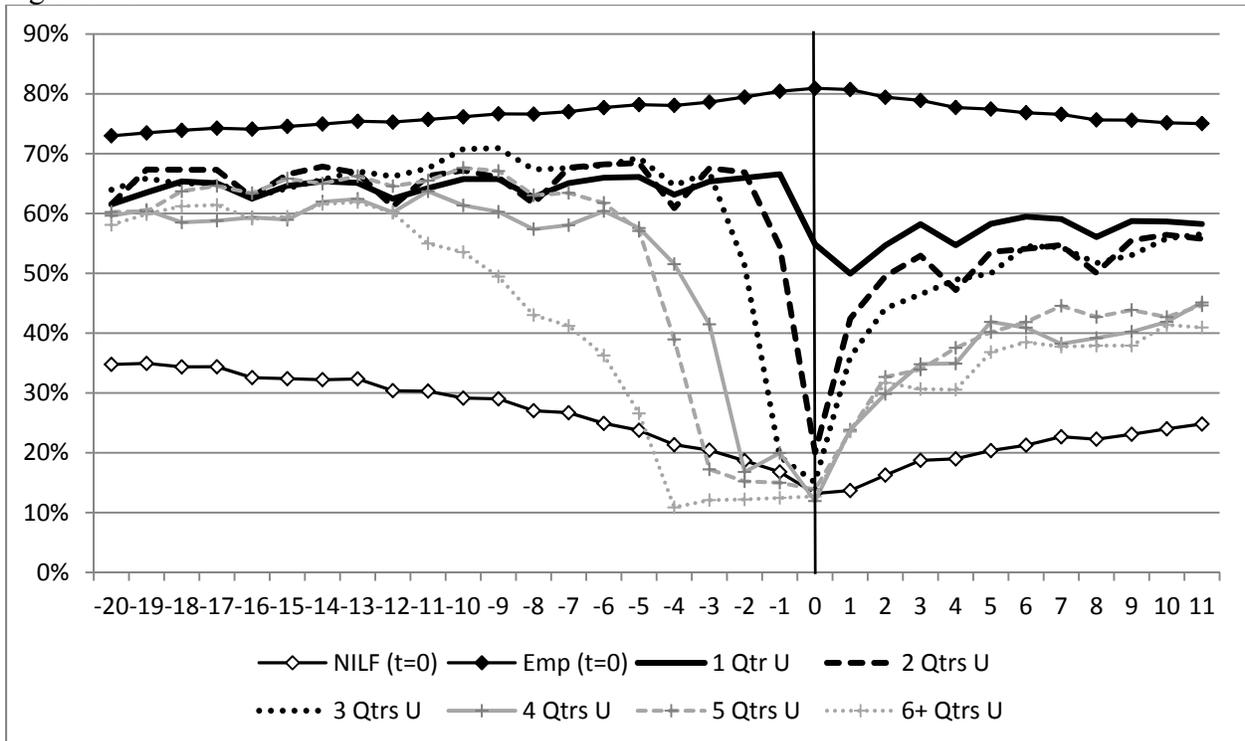


**Figure A.4: Employment Probabilities, by Age {<30, 30-49, ≥50}**

Age <30

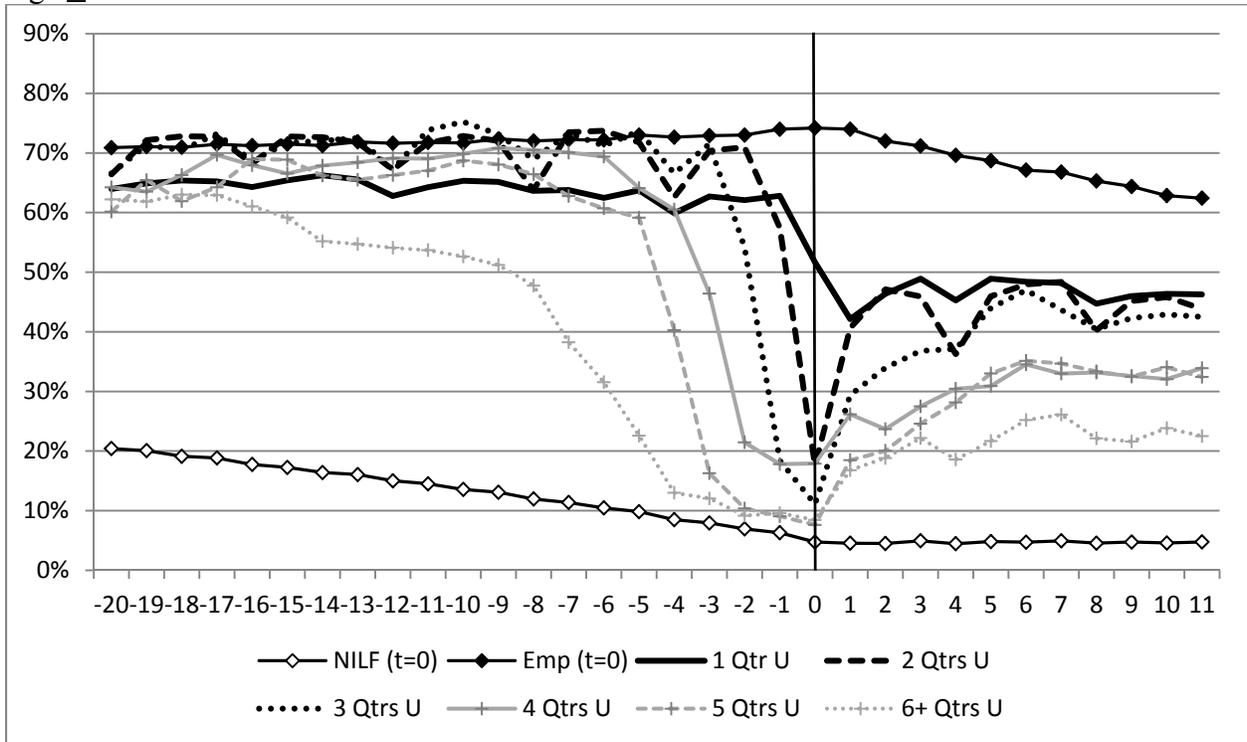


Age 30-49



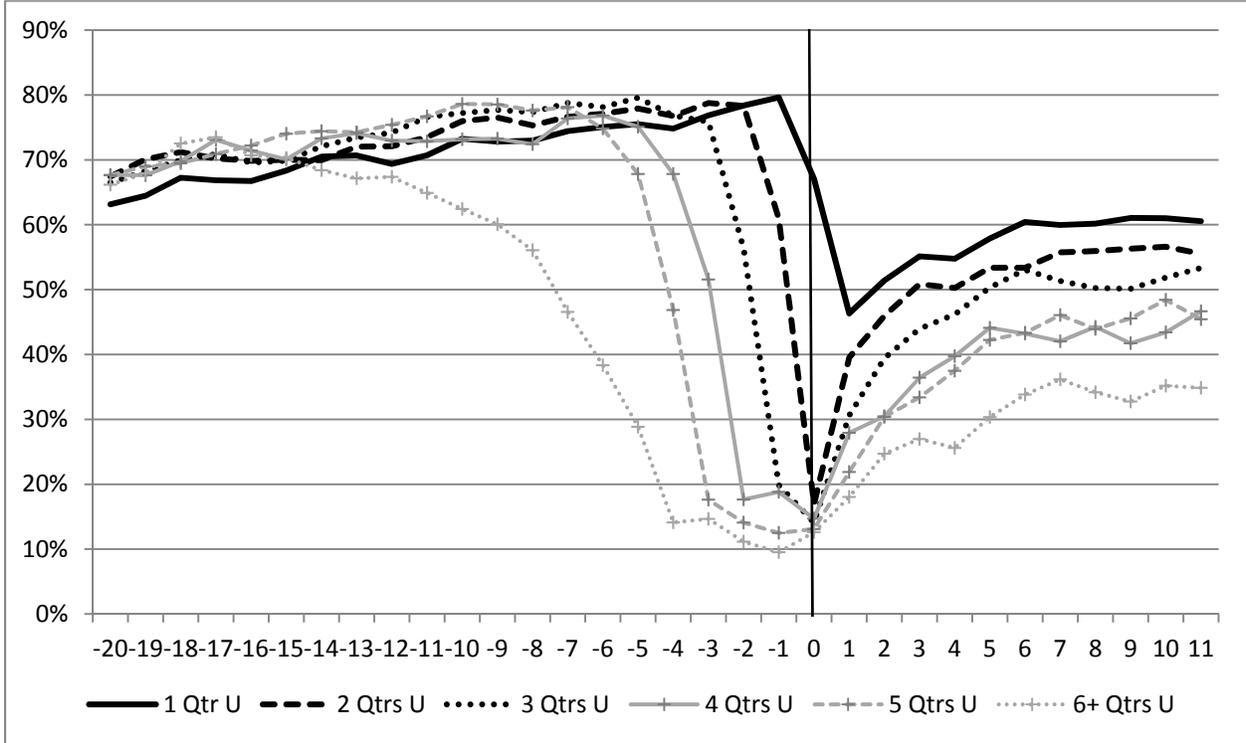
**Figure A.4 (continued)**  
**Employment Probabilities, by Age {<30, 30-49, ≥50}**

Age ≥50



**Figure A.5: Employment Probabilities, by Reason for U {Job Loser, Other EU, NU}**

**Job Loser**



**Other EU**

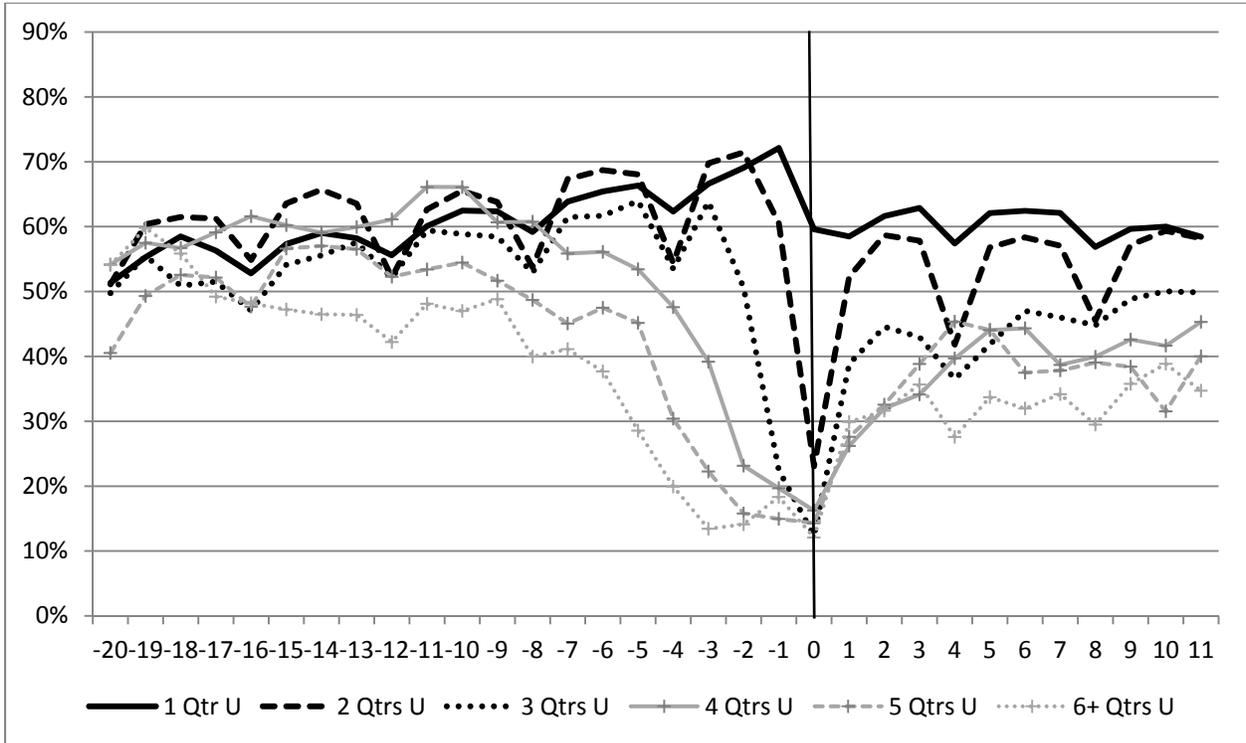


Figure A.5 (continued)

Employment Probabilities, by Reason for U {Job Loser, Other EU, NU}

Entrants NU

