Uncertain Futures? Youth Attachment to the Labor Market in the U.S. and New England

Appendix D. Data and Methodology

Data Sources

The data sources used to perform the analyses in this report include:

- 1980, 1990 and 2000 Decennial Census 5% Sample
- American Community Survey (ACS) 2005–2007 and 2009–2011 3-Year PUMS; ACS 2007 1-Year PUMS
- March Current Population Survey (CPS), 1976–2012
- Bureau of Labor Statistics (BLS) Current Employment Statistics (CES), December 2007–April 2013

The Decennial Census, ACS, and CPS data are from the Integrated Public Use Microdata Series (IPUMS-USA and IPUMS-CPS respectively) provided by the Minnesota Population Center at the University of Minnesota. IPUMS harmonizes these data to be compatible over time (e.g., industry and occupation classification codes) and with each other. [1]

Differences in Data Collection and Reference Periods

Despite harmonization of variable definitions, there are differences in data collection and survey methodology that impact the comparability of the estimates presented in the preceding report.

Decennial Census and ACS:

The American Community Survey has been replaced the Decennial Census Long Form for the collection of detailed socioeconomic information about the population as of Census 2010. However, the Decennial Census Long Form and ACS data collection periods are different, which may affect comparability of the estimates presented in Tables 1, 3–7, Appendix B: Tables B1–B3, Appendix C: Tables C1-C4, and Figure 8 for 1980–2000 vs. 2006 –2010. [2] The Decennial Census estimates can be thought of as reflecting reported characteristics at a "point in time," April 1st of the first year in each decade. The long form data collection occurred over a narrow time frame and many survey questions instructed respondents to fill out the form based on personal/household characteristics using the consistent April 1st reference point. The ACS, however, reflects characteristics of the population over a "period of time," with data collection occurring throughout the entire year.

Additionally, the sample size of the ACS 1-year PUMS is significantly smaller than the decennial census, resulting in the need to combine multiple years of data. To obtain a sufficient sample

---

[1] For more information on IPUMS-USA and IPUMS-CPS, visit http://www.ipums.org/.

[2] Tables in the report that report labor market attachment characteristics for 2006 and 2010 use the ACS 2005–2007 3-year PUMS and ACS 2009–2011 3-year PUMS respectively.
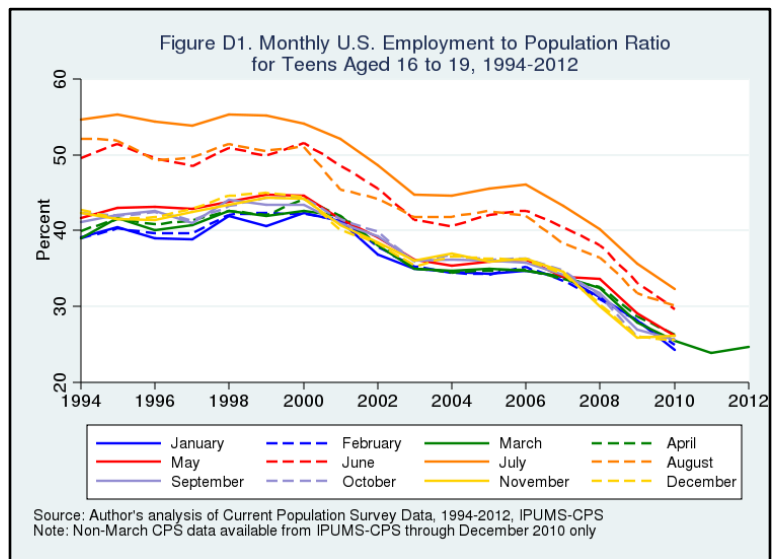
size for the estimates presented in this report, the ACS 3-year PUMS is used–lengthening the collection period to a three year period. As such, estimates presented in Tables 1, 3–7, B1–B3, C1-C4, and Figure 8 for 1980–2000 are not strictly comparable to those presented for 2006 and 2010.[3]

Decennial Census/ACS and CPS:

The labor market characteristics presented in the tables are not strictly comparable with those presented in the figures due to differences in the reference period between Census/ACS and CPS data. Most tables presented in the report are constructed using Decennial Census and ACS estimates, whereas the figures that analyze yearly labor market attachment over a continuous annual time series are constructed using CPS estimates.[4] The reference period for all labor market characteristics constructed using the Decennial Census and ACS is the week prior to the interview data–as such, these data are heavily influenced by the period of data collection. The CPS estimates provide measures of labor market attachment using a consistent reference period each year–the week containing the 12th of each month.

The CPS labor force estimates presented in the figures are from the March monthly labor force survey and therefore reflect characteristics of the respondent during the week in March that contains the 12th day of the month. As such, the point estimates presented in these figures may not accurately represent labor force attachment and school enrollment over the entire year. Figure D1 displays the teen employment to population ratio for each monthly survey over the

period 1995–2012. While the *level* of labor force attachment among youth varies throughout the year, the trends detailed in the report using March CPS data appear to hold regardless of the month chosen.

To more accurately reflect labor force characteristics over the entire year, point estimates presented in the paper are taken from the Census/ACS tables, rather than the CPS figures. The CPS figures are used to describe trends in the attachment and idleness over a continuous, annual time series.

Figure D1. Monthly U.S. Employment to Population Ratio for Teens Aged 16 to 19, 1994-2012

Source: Author's analysis of Current Population Survey Data, 1994-2012, IPUMS-CPS
Note: Non-March CPS data available from IPUMS-CPS through December 2010 only

---

[3] Tables in the report that report labor market attachment characteristics for 2006 and 2010 use the ACS 2005–2007 3-year PUMS and ACS 2009–2011 3-year PUMS respectively.
[4] Table 2 is constructed using the Current Population Survey as this question is not asked in the decennial census or ACS. Figure 8 is constructed using the decennial census and ACS.

Notes to Accompany Figures and Tables

Comparability Issues over Time

Many of the charts and figures using CPS data note that data from survey years prior to 1994 are not strictly comparable with those from 1994 onward due to a survey redesign. Much of the redesign was in response to dramatic increases in female labor force participation since the last major revision in 1967. Prior to 1994, respondents were asked about their "major activity" in the reference period week—a question that varied based on the gender and age of the respondent. Reports by the Bureau of Labor Statistics have concluded that estimates prior to the survey redesign led to underreporting of part time work—especially for females. Additionally, the BLS reports that prior to 1994, the proportion of those employed part-time for economic reasons was overestimated and reasons for unemployment were mismeasured.[5]

Attachment Measures by Race and Ethnicity

For purposes of reporting attachment measures by race, all individuals who self-report being of Hispanic origin are defined as Hispanic regardless of reported race.

Attachment Measures by Nativity

For purposes of reporting attachment measures by nativity, "native-born" is defined as born in the United States. Individuals born in Puerto Rico and other U.S. territories are defined as "foreign-born". "Second-generation immigrants" refers to individuals born in the United States with two foreign-born parents. Note that CPS data on nativity is not available prior to 1994.

Attachment Measures by Family Income

Family income categories are based on percentiles of people, not families or households. In other words, family income percentile is assigned relative to family income of other individuals. For youth, especially those aged 20 to 24, there exists considerable variation in family income conditional on whether the respondent has established his/her own residency or stills lives with a parent/guardian or other relative. Additionally, family income for most respondents who report living in group quarters (e.g. college dormitories) is equal to individual income.[6] For teens age 16 to 19, over 90 percent of respondents report living with a parent/guardian or other relative, compared to 48 percent of young adults aged 20 to 24. Average family income for all youth aged 16 to 24 who report having established his/her own residency is 62% lower than those who report living with a parent/guardian or other relative across the years studied (1976–2012).[7] As such, the figures displaying attachment measures by family income should be interpreted with caution.

---

[5] For additional information see http://www.bls.gov/osmr/pdf/ec950090.pdf.

[6] In instances where multiple family members live in the same group quarters, family income is the sum of individual income for all respondents.

[7] This 62% figure is an average across years, ranging from 53% in 1976 to 69% in 2012.

Cohort Analysis of Labor Market Attachment

Figures 9, 10, and B2 present a cohort analysis of youth labor market attachment over time. The data are plotted such that successive cohorts are followed over time. Eight synthetic cohorts are created using the CPS survey from every fifth year over the period 1976–2011. The years displayed in the legend for each figure represent the years that each cohort was aged 16 to 19. For example, the purple line in Figure 9a plots the employment to population ratio for the synthetic cohort that was aged 16 to 19 years in 1976. The employment to population ratio for this cohort is shown at five different points over their careers through 1996, the year in which this cohort was aged 35 to 39. The most recent cohort was aged 16 to 19 years in 2011, capturing the experiences of youth affected by the Great Recession.

Figure 10 presents this cohort analysis of youth labor market attachment over time for two native demographic groups. As data on nativity is not available prior to 1994, there are no cohorts displayed with attachment estimates for all five points in time. For example, the purple line in Figure 10a—representing native white males aged 16 to 19 in 1976—contains only one data point for data from 1996, representing the period when this cohort was aged 35 to 39.

Top Youth Industries and Occupations

All industry analysis in Tables 6, 7, and A1–A4 is conducted at the subsector level. To provide a consistent long-term classification, the analysis is conducted with the IPUMS industry variable that uses a modified 1990 industrial classification system (IND1990). All occupation analysis is conducted at the broad occupation (5 digit) level with the IPUMS occupation variable that uses a modified 1990 occupational scheme (OCC1990).

Table 6 displays an industry breakdown of youth employment for all industries and "top" youth industries over the period 2000–2006. It also separates employment by industry into two broad categories: industries for which youth share is falling between 2000 and 2006 and industries for which youth share is increasing between 2000 and 2006. These categories are further decomposed into industries that are a decreasing or an increasing share of the economy. Industries that are a decreasing share of the economy are those where industry-level employment is growing less than average aggregate employment; industries that are an increasing share of the economy are those where industry-level employment is growing more than average aggregate employment.

"Top" or "youth intensive" industries and occupations are defined as those ranked in the top 50 in terms of number of youth employed or share of youth employment within the industry/occupation in 2000 or 2006. Additionally, the industry/occupation must be present in both years. Some occupations and industries are not available in all years, due to the "majority mapping" methodology used to construct the harmonized industry/occupation variables.[8]

Table 7 lists only "top" industries/occupations where the youth share (ages 16–19) is falling and young adult share (20–24) is rising. Tables A1–A4 display all "top" youth industries and occupations for each youth age group, teens aged 16 to 19 and young adults aged 20 to 24.

---

[8] See https://usa.ipums.org/usa/volii/occ_ind.shtml for more information.

Changes in Employment for Top Youth Industries by Quadrant

All industry analysis in Table 8 is based on the methodology developed by Groshen and Potter (2003) and is conducted at the 3-digit NAICS code (subsector) level[9]. To estimate the shares of peak employment by age cohort and industry prior to the Great Recession, the 2007 ACS 1-year PUMS is used. Age shares are calculated by collapsing over adjusted 3-digit NAICS codes.

The 2007 1-year ACS uses the 2002 NAICS Classification System; therefore, several cross walking steps are taken to match the modified 2012 NAICS classification system used by the BLS Current Employment Statistics. Additionally, in order to match the CES universe for total nonfarm employment, active duty armed forces observations are dropped, public jobs are stripped out of NAICS codes into government categories, and agriculture and private household industries are excluded.

These calculated employment shares by age cohort and industry are merged with the BLS Current Employment Statistics Nonfarm Employment by Industry to estimate peak employment by industry among youth aged 16–19 and 20–24. To maintain a consistent universe with all other analysis presented in the report (i.e. total population aged 16 to 65), an adjusted universe is calculated by removing the estimated share of peak employment for those aged 66 and older by industry. As such, the shares presented in Table 8 are expressed as a fraction of estimated peak employment for those aged 16 to 65.

The following definitions are used in conducting the analysis:

- Percent growth in employment over the course of the recession is calculated with recession defined as month of peak to month of trough, inclusive.
- Percent growth during the recovery period is calculated as month of trough to end of recovery month, inclusive. The end of recovery definition used is "return to peak," defined as the time period when total nonfarm employment returns to the level of employment at the NBER recession peak.

For purposes of performance over the Great Recession, the NBER definitions business cycle peak (December 2007) and trough (June 2009) dates are used.   "Return to peak" employment has yet to occur for the Great Recession, so the most recent employment data available at the time the analysis was conducted (April 2013) is used as the end of recovery month.

To classify industry performance over the business cycle, we follow the Groshen and Potter (2003) methodology. Table 8 categorizes top youth industries by performance over the business cycle during the Great Recession. Industries are classified as "structural" or "cyclical" based on axes defined by the overall change in total nonfarm employment during the recession versus the recovery, which centers the industry distribution around the specific origin for each recession/recovery. In other words, the percent change in total nonfarm employment is subtracted from each industry's percent change in employment. Performance is further categorized by separating "structural" changes into two categories: "structural loss" and

[9] Groshen, Erica and Simon Potter. 2003. "Has Structural Change Contributed to a Jobless Recovery?" *Current Issues In Economics and Finance*, Federal Reserve Bank of New York.

"structural gain." The "cyclical" definition is separated into "procyclical" and "countercyclical" categories.

The specific industries listed in Table 8 are top youth industries that underwent structural changes during the recession/recovery and account for at least 0.5 percent of peak employment among youth aged 16–19 years or 20–24 years.