Mismatch in the Labor Market?
Measuring the Supply of and Demand for Skilled Labor in New England

Methodology Appendix

The demand for skilled labor

To make comparisons over time, real average hourly wages are estimated in each year across education levels based on predictions from ordinary least squares (OLS) regressions. The OLS regressions account for changing characteristics of workers over time that are associated with earnings, and are estimated separately for men and women.

The dependent variable is the natural log of the hourly wage, calculated as annual earnings divided by the product of weeks worked per year and hours worked per week. Independent variables include education as a piecewise linear combination of dummies (a dummy for less than a high school diploma and linear years of schooling for less than high school, a dummy for some college and linear years of schooling for some college, a dummy for a bachelor's degree, a dummy for an advanced degree and linear years of schooling beyond a bachelor's degree). Other independent variables include a quartic in potential experience, as well as dummy variables for race, ethnicity, and foreign born status. Predictions are calculated for an individual with 15 years experience at the 2006-based average of all other characteristics.

The supply of skilled labor

We calculate the supply of skilled labor by education level for the U.S., New England, and each New England state in 1980, 1990, 2000, and 2006 using the 1980, 1990, and 2000 decennial Censuses, and the 2005-07 combined American Community Survey (ACS), respectively.[1] The total number of individuals in each year and education category represents all 25-64 year olds.

Future demand for skilled workers

Employment by education category

To obtain national and state estimates for the total number of jobs in 2006, 2009, and 2018 by education level, we first obtain non-military employment estimates by occupation for 2006, 2009, and 2008 from the U.S. Bureau of Labor Statistics (BLS) Occupational Employment Statistics (national and

[1] Some calculations use only the 2005 and 2006 ACS (combined). All decennial Census/ACS data obtained using the Integrated Public Use Microdata Series (IPUMS**:** Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2010.)

state cross-industry estimates).[2] To calculate employment by occupation in 2018, we apply projected national growth rates of employment by occupation for 2008-2018 to the 2008 geography-specific employment estimates. The projected growth rates are obtained from the BLS Employment Projections Program, and are calculated to account for new job growth and retiree replacement.

Employment by occupation for each year is then assigned education categories according to the current education distribution for each occupation, as calculated using employed 16-69 year-olds at the national level from the 2006-08 combined ACS.[3] For the 2018 estimates, this calculation represents how future demand will change given changes in employment *across* occupations, holding the future education distribution of workers in those occupations constant. These are the "lower" estimates of future labor demand. To account for changes in demand *within* occupations over time, we shift the education distribution to reflect past trends. We calculate the change in the education distribution by occupation for employed 16-69 year olds at the national level between the 2000 decennial Census and the 2005-07 combined ACS. The change is doubled to imitate the effects of a decade, and applied to the current education distribution. The 2018 employment estimates by occupation are then distributed into education categories from the projected education distribution to obtain "upper" estimates for 2018.

Total employment numbers by education level in 2006, 2009, and 2018 (both "lower" and "upper" estimates) are calculated by aggregating employment over all occupations by each education category.

Total vacancies

To obtain total number of jobs for 2006 and 2009 by education level, employment by education level needs to be summed with the number of vacancies by education level.[4] Vacancy data obtained from The Conference Board Help Wanted OnLine™ (HWOL) data series is calculated by occupation as state-level annual averages. The average number of vacancies by occupation are aggregated to the national and regional levels, and distributed into education categories using the current education distribution for each occupation. Vacancies are then summed over all occupations by each education level to generate total vacancies in each education category. Total employment and total vacancies are summed to obtain total labor demand in 2006 and 2009 for each education category.

Future supply of skilled labor

Baseline populations

[2] All "by occupation" estimates in the methodology are at the 6-digit Standard Occupational Classification (SOC) code level.

[3] The education categories are "less than high school", "high school", "some college", "associate's degree", "bachelor's degree", and "advanced degree".

[4] The total number of jobs for 2018 by education level is equivalent to the employment estimates for 2018, since the projected 2008-2018 growth rates used to generate 2018 employment by occupation already include vacancies through new job growth and retiree replacement considerations.

To estimate population changes, we rely on an approximated cohort component model to age the current population over time and model the U.S. and New England populations in 2019 and 2029. First, the current 5-64 year old population for the chosen geographies is obtained from the July 2009 U.S. Census Bureau population estimates, and allotted into five-year age groups (5-9, 10-14, etc.), gender, and race/ethnicity.[5] Each specific cohort is then split by nativity based on the percentage of that cohort that is native-born or foreign-born in the 2000 decennial Census.

To age the population, a 10-year "survival rate" is calculated as the percentage of individuals for each cohort appearing in both the 1990 and 2000 Census.[6] As an example, the number of 35-39 year old native white males in New England in 2000 divided by the equivalent cohort in 1990 (25-29 year old native white males in New England) is equal to 95 percent. The survival rates embody both mortality and net migration rates.

These 10-year survival rates are then applied to the corresponding cohort in the current baseline population to project each group into 2019, and are then applied again to project into 2029. Beginning with 5-64 year olds in 2009, we now have available estimates for 15-74 year olds in 2019 and 25-84 year olds in 2029. Since the focus is on the labor force, only 25-64 year olds in each year are used in aggregate estimates.

Future educational attainment estimates

The current distribution of educational attainment for each geography and cohort is calculated from the 2006-08 combined ACS.[7] To determine the current educational composition of the population, the education distribution of each specific cohort is applied to its corresponding 2009 population estimate, and then the number of individuals is summed over the entire labor force (25-64 year-old cohorts) for each education category. Estimates of future educational attainment need to account for changes in the composition of the population and gains in educational attainment. To project the educational composition of the population under these circumstances, the education distribution will be assigned to the baseline populations' entering and current cohorts based on recent trends in educational attainment. Varying assumptions will create "lower bound" and "upper bound" estimates. Entering cohorts are defined as the age groups that would have entered the labor force over the course of a decade, and current cohorts are those who would have remained in the labor force over the course of a decade. As an example for the 2019 estimates, 15-24 year old individuals in 2009 will be entering cohorts in 2019 as 25-34 year olds. Likewise, 25-54 year olds in 2009 will be current cohorts as 35-64

[5] Included race/ethnicity categories are white, African-American, Hispanic, Asian (including Pacific Islander), and "other" (which includes American Indian and two or more races). Hispanic origin is coded as an ethnicity in the Census (a separate variable from race), so the Hispanic category includes all individuals of any race who identify as Hispanic. The other race categories only include individuals of non-Hispanic origin.

[6] Due to data limitations, all non-white race/ethnicity categories are combined to calculate survival rates for non-white cohorts. Also, younger age cohorts (5-9 and 10-14 year olds) assume the survival rates of 15-19 year olds.

[7] Due to data limitations, education distributions are not available for some cohorts in northern New England; specifically, 50-54, 60-64, and 65-69 year old African-American females and 60-64 year old males in the "other" race/ethnicity category.

year olds in 2019. Individuals 55 years and older are assumed to exit the labor force over the course of a decade.

"Lower bound" estimates only reflect changes in the composition of the population, with no effect from advancements in the rate of educational attainment. These are calculated by assigning entering cohorts the same education distribution as their equivalent predecessors. For example, since 28.4 percent of native white 25-29 year old males in New England have a bachelor's degree in the current education distribution, we assume that 28.4 percent of native white 15-19 year old males in New England in 2009 will have a bachelor's degree in 2019 (and likewise 28.4 percent of native white 5-9 year old males in New England in 2009 will have a bachelor's degree in 2029).

In addition to reflecting changes in population composition by assigning entering cohorts the education distribution of their predecessors, "upper bound" estimates also take into account changes in the educational attainment distribution over time as individuals obtain additional education. These changes are calculated as the change in educational attainment for each cohort between the 1990 and 2000 decennial Census. For example, in 2000 an additional 3 percent of 40-44 year old native white males in New England had an advanced degree compared to 30-34 year olds in 1990. Therefore, it is assumed that an additional 3 percent of native white 30-34 year old males in New England in 2009 will have an advanced degree in 2019 as 40-44 year olds. This change is applied to the education distribution for current cohorts less than 39 years old, after which it is assumed the probability of increased educational attainment decreases. For the 2029 "upper bound" estimates, the change is applied twice (first to the under-39 current cohorts in 2019 and then the under-39 current cohorts in 2029) to reflect two decades of shifts in the education distribution.

Once "lower bound" and "upper bound" estimates of population by education category are created for 2019 and 2029, the numbers are summed for all 25-64 year-old labor force cohorts by education category to obtain future educational attainment under varying assumptions.

Labor force participation rates

As a final step, labor force participation rates are generated from the 2005-07 combined ACS survey for each cohort and education level.[8] These labor force participation rates are applied to the corresponding population estimates to obtain the projected labor force for each set of estimates.

Policy exercises

To estimate potential labor market adjustments in New England as a result of mismatch, the following trends were applied to the "upper bound" population estimates, either separately or in combination:

[8] Similar to survival rate calculations, all non-white race/ethnicity categories are combined to calculate labor force participation rates for non-white cohorts due to data limitations.

- Increasing educational attainment of new labor force participants was estimated under the following assumptions[9]

	White	Asian	African-American	Hispanic	Other
Increase in high school graduation rate	2.0%	2.0%	3.0%	5.0%	3.0%
Increase in college continuation rate	5.0%	5.0%	1.0%	3.0%	1.0%
Increase in college completion rate:					
Associate's degree	2.0%	2.0%	0.0%	0.0%	0.0%
Bachelor's degree	5.0%	5.0%	0.0%	0.0%	0.0%

- Greater regional in-migration, which is historically highest amongst higher-educated and younger workers, is estimated as an increase in 25-39 year old labor force participants, with magnitudes varying by education category: 5 percent for cohorts with a bachelor's degree or higher, 2.5 percent for cohorts with some college or an associate's degree, and zero for high school or less.
- Increased labor force participation amongst older workers in response to the Great Recession is estimated as a 5 percent increase in labor force participation rates for 55-64 year old workers across all education levels.

[9] Source: author's calculations based on historical data from the National Center for Education Statistics.