Methodological Appendix
Demand for H-1B Visas in New England: An Analysis of Employer Requests for Highly-Skilled Guest Workers (New England Public Policy Center - Policy Report 14-1)

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

- Massachusetts Executive Office of Labor and Workforce Development, Job Vacancy Survey, Second Quarter 2012.
- U.S. Bureau of Labor Statistics, Job Openings and Labor Turnover Survey, October 2011 to September 2012.
- U.S. Bureau of Labor Statistics, Occupational Employment Survey, 2008-2012.
- U.S. Department of Labor, Employment and Training Administration, CareerOneStop, EmployerLocator.
- U.S. Department of Labor, Employment and Training Administration, Office of Foreign Labor Certification, OFLC Performance Data, LCA Programs Disclosure Data, FY 2008-FY 2012

Intensity of Demand

The intensity of demand measures used in this report are the number of Labor Condition Application (LCA) requests per 1,000 payroll employees. This concept came from the intensity of demand measures used by Ruiz, Wilson, and Choudhry (2012).These measures required the alignment of geographic and occupational concepts from the Labor Condition Application with the Occupational Employment Survey (OES).

Labor Condition Application

Since 2000, the Office of Foreign Labor Certification (OFLC), within the Department of Labor's Employment and Training Administration, has been making every LCA request available in a public database of administrative records. The data is available by fiscal year (October 1-September 31) when the request is filed, and includes the number of workers requested, the prospective start and end dates of employment, occupational and industry details, wages, company name and address, and the city and state where the worker will be employed. The granularity of this data allows for analysis of the H-1B visas at a local level, a major strength over other H-1B visa data sources that provide only national level information.

However, there are a number of limitations to this dataset. First, the LCA is only the first step in the H-1B visa application process and does not require the payment of any fee to file. As a result the number of LCAs filed in a fiscal year far exceeds the number of I-129 forms, the next step in the application process, and there is no way to track an LCA through the application process (Figure 1). Further, the LCA does not distinguish between a new application and an application for continuation of employment of a current visa holder. Due to these limitations the LCA is likely a better indicator of employer demand for H-1B workers, than a measure of supply. Both Kerr and Lincoln (2010) and Ruiz, Wilson, and Choudhury (2012) use this data as a dependency or demand measure.

The other major limitation to using the LCA is the data quality issues that arise in the administrative records. The OFLC posts the LCA records as reported by employers with little to no cleaning of the data. In particular, data quality for certain years between 2000 and 2007 made the construction of a time series of the LCA impossible for these years. As a result, the analysis of the LCA was limited to 2008 through 2012. Data quality issues still persist in these years as well, such as numerous records with missing observations, misspellings of employers, cities, and occupations, and

incorrect assignment of numbers of workers requested, occupational codes, and wages values. To be confident in the analysis of the data three methodological concerns needed to be addressed: (1) the status of the LCA, (2) assigning the request to the correct geography, and (3) aligning occupational codes to compare with other data sources. Additional assumptions needed to be made for the interpretation of requests and the analysis of employer requests.

LCA Status

The DOL reviews LCAs for completeness and only rejects those applications with obvious errors. In the LCA data there is a variable regarding the status of the application. The status is certified, denied, or withdrawn. Withdrawn applications are typically those submitted by employers but then removed from consideration. Withdrawn LCAs were removed from this analysis because they were stopped before the final stage of review and are missing many key pieces of information. Therefore the analysis is only for those applications that completed the review process and were either certified or denied. A majority of the applications are certified (See Table 1).

Geography

The geographic variable of interest in the LCA is the prospective place of work. In particular, we wanted to know which metropolitan area the potential recipient would work in to be able to compare with other data sources, such as the OES. Unfortunately, the LCA only lists the city name and state for the place of work but does not provide a zip code. In a number of instances city names are misspelled, city or state names are missing, and nicknames and acronyms are used (e.g L.I.C. for Long Island City which is part of Queens in New York City, NY). To make sure geographies are correctly assigned to a metropolitan area two crosswalks were used.

The first crosswalk was created using the United States Postal Service (USPS) Address Information System (AIS) City State Product. This AIS database contains detailed zip code information for all towns and cities, including their abbreviations and alternative names. This database was merged with the U.S. Department of Housing and Urban Developments (HUD) USPS ZIP Code Crosswalk, which matches zip codes to Core Based Statistical Areas (CBSA) such as a metropolitan or micropolitan area.[1] By combining these two data sources we can match a work location by city and state (e.g. Boston, MA) to a zip code and metropolitan/micropolitan area. An additional step was needed for New England geographies, as the OES provides data by New England City and Town Areas (NETCAs) which do not conform to CBSA boundaries. To match zip codes to NECTAs, a crosswalk was generated between five digit zip codes and NECTAs from the Missouri Census Data Center's "Geocorr12: Geographic Correspondence Engine." This identified nearly all major zip codes in the USPS AIS system, but some manual additions were required to complete the crosswalk. This crosswalk provides the primary option to identify geographies, but will only match to correctly spelled city and state names.

The second crosswalk was created from the LCA database itself. The LCA data provides the city, state, and zip code of a requesting employer. As the city name is prone to misspelling in the administrative data, the zip code tends to be a more reliable geographic identifier. The city of the

[1] The HUD USPS Zip Code Crosswalk provides a list of CBSAs and which zip codes are associated with it. In some cases a zip code fell within more than one CBSA. The HUD database gives the concentration of business addresses within each CBSA, along with residential, other, and total address concentrations. After merging the USPS AIS data with the HUD USPS Zip Code Crosswalk, all towns that fell into more than one CBSA were identified and assigned to the CBSA with the largest concentration of business addresses. Alternative assignments of these zip codes were run and had no notable impact on the intensity of demand measures.

requesting employer is as prone to misspelling as the work location, so the addition of the zip code provides an additional geographic identifier by which to correctly identify misspelled cities, less commonly used names, or acronyms. From the LCA we collapsed the data set into a list of city and state names (e.g. Boston, MA) with the mode zip code within the data. The above steps were repeated for matching the zip codes to CBSAs and NECTAs. This crosswalk provides a secondary option by which to identify any geography that is not identified using the USPS crosswalk.

To improve the matching of geographies the most obvious misspellings of cities (e.g. Wocester, MA to Worcester, MA) were manually edited. The USPS based crosswalk was then merged with the work city variable in the LCA, yielding a match of over 99 percent. The LCA based crosswalk was then merged to any work city observations that were not matched in the first merge. In a number of instances where the geography could not be identified the remaining issue appeared to be a missing or incorrect state. To try and correct for this, the missing or potentially incorrect state initials of the work location were replaced with those of the employer location. For example, a request from a Boston, MA employer for someone to work in Boston, MT, which does not exist, would be changed to Boston, MA. The above iteration of merges is completed again to correct for any remaining mistakes. As a result the geography of the work location was identified for 99.5 percent of LCA requests in any given year.

Occupations

The inclusion of detailed occupational information in the LCA allows for further dissection at a local level. Starting in 2010, the LCA begin to provide 6-digit Standard Occupational Classification (SOC) codes. Unfortunately, prior to 2010 the LCA used Dictionary of Occupation Titles (DOT) to classify occupations. Multiple DOT codes map to multiple SOC code making comparisons over time difficult. To avoid incomplete comparisons of occupational groups' overtime analysis of occupational level data was limited to 2010, 2011, and 2012.

The 2010 LCA occupational coding is based on 2000 SOC codes while 2011 and 2012 codes are based on 2010 SOC codes. To be consistent all codes were cross walked into 2010 codes. In numerous instances SOC codes were missing or incorrectly specified. When possible, codes were imputed based on the SOC title that was provided. A limited number of observations were dropped when the occupation could not be identified.

Requests and Positions

On the LCA employers are not required to identify an individual they hope to employ. As a result a single request can be for multiple positions to be filled by H-1B workers. Between 2008 and 2012, 93.4 percent of requests were for 1 position, with another 4.4 percent for between 2 and 10 positions. Only 2.2 percent of requests were for more than 10 positions and came largely from a small subset of employers.[2] To be conservative in estimates of demand this report treats each LCA as requesting one worker to avoid having a single employer, request, or error in total positions requested from overstating the demand for H-1B workers in a regional labor market.

[2] Only 137 employers averaged more than 10 positions per requests over this time period. Out of these 137 employers with requests averaging more than 10 positions 68 averaged less than 1 request a year and only 10 averaged more than 100 requests per year. Alternative intensity measures were created using multiple positions per request. Such measures were comparable in the geographic and occupational distribution of measures treating each request as being for one position, but notably larger in high intensity labor markets.

Employers

Analysis of employer level demand presented similar challenges to the geographic analysis as there is no unique identifier for employers. Instead a requesting employer's name and location (street, city, state, and zip) were provided. Such information was as error prone as the spelling of cities and difficult to analyze in a systematic way. To simplify analysis of employer level requests, we focused only on the top 100 requesting employers in New England and the United States. Even within this small subset of employers there was significant variation of employer names due to the inclusion of commas, periods, ampersands, and abbreviations (e.g. Corporation, Corp, Inc., ... etc.). All names were assigned to match the version with the most requests. To the extent possible, requests from subsidiaries of the top requesting employers were assigned to the parent company.

One concern raised by critics of the H-1B visa program, for example see Hira (2007 and 2010), is that the visa is heavily used by a number of "offshore outsourcing" firms. The concern is that these firms use H-1B workers as a source of cheap temporary labor to learn a job function on-site at a contracting employer and then transfer the knowledge required to perform this job to a cheaper offsite location where work can be done remotely. If these firms play a significant role in the demand for H-1B visas, particularly in regional labor markets, then such demand may not necessarily reflect local labor market tightness but instead serves a specific industry business model. To gauge the role of such firms at the regional level we reviewed the services and business activities of the top H-1B visa requesting firms nationally and regionally and identified 20 employers where outsourcing is a significant component of business activity. All 20 firms identified as potential outsources were primarily in the computer systems design and related services industry, provided IT and business process outsourcing (BPO) services, obtained a large portion of annual revenues from such services, requested workers for locations in multiple states and Census divisions, and requested mostly H-1B visas in computer and mathematical occupations. The 20 firms identified as potential outsourcers are: Capgemini, Cognizant, CSC Covansys, Dotcom Team, Fujitsu (Consulting), HCL America, Hexaware Technologies, IBM, Infosys, KPIT Infosystems, Larsen & Toubro Infotech, MphasiS, Patni Americas/iGate, Satyam Computer Services, Synechron, Syntel, Tata Consultancy, Tech Mahindra (Americas), UST Global, and Wipro. It should be noted that all of these firms are also involved in business activities outside of outsourcing and could be requesting H-1B visas for purposes other than outsourcing related activities.

Lastly, in an effort to quantify employment of the top requesting employer's in the New England states, the names of the top 20 firms by requests for H-1B workers in New England for FY 2010-2012 were searched in the U.S. Department of Labor's Employment and Training Administration CareerOneStop Employer Locator database, using the keyword search function and checking for each employer in the six New England states. This database provides ranges of approximate employer sizes by location, along with industry classifications and contact information. Information on employers is compiled from public sources with follow on phone calls to identified businesses to verify the accuracy of the information.[3] All employment ranges across the New England states were totaled to obtain a potential range of the firm's New England employment. Ranges were standardized to the ranges listed on the Employer Locator tool (e.g. 100-249, 250-499, 500-999). Firm's regional employment ranges were cross referenced with Hoovers' company database.[4] Firms with smaller regional employment totals were more variable between sources so the smallest employment estimates that could confidently be reported was fewer than 100 employees across the six New England states.

[3] See ReferenceUSAGov for more detail about data collection methods: https://referenceusagov.com/Static/AboutUs
[4] See http://www.hoovers.com/

Defining Science, Technology, Engineering, and Mathematics (STEM) Occupations

A defining feature of H-1B visas is their concentration in STEM fields. However, analysis of H-1B visas and STEM use an array of definitions. For example, Lofstrom and Haynes (2011) estimate that close to 75 percent of all H-1B visas issued in 2009 were in STEM, broadly defined as occupations in engineering, math and sciences, health, and post-secondary education. Ruiz, Wilson, and Choudhury (2012) use the U.S. Department of Commerce's definition that categories 50 detailed occupations as STEM and find 64.3 percent of LCA requests are for STEM occupations. Rothwell and Ruiz (2013), categorize occupations as STEM if they require high levels of STEM knowledge as defined by O*Net, find that 90 percent of LCA requests are for STEM occupations. And these are but a few examples of a broad range of STEM definitions.

Given the variation in STEM definitions any analysis of H-1B visas will be influenced by the definition used. To avoid selecting expansive or limiting definitions of STEM the report looks at three STEM components based on recommendations from the SOC Policy Committee (SOCPC) to the Office of Management and Budget. The SOCPC formed a working group of agencies and organizations from labor and education to review existing definitions of STEM and develop a framework for defining STEM consistent with SOC Classification Principles.[5] The committee developed a definition of STEM composed of two major domains: (1) Science, Engineering, Mathematics, and Information Technology (STEM) Domain, and (2) Science- and Engineering-Related Domain. The STEM domain is composed of two sub domains: (1) Life and Physical Science, Engineering, Mathematics, and Information Technology Occupations, and (2) Social Science Occupations. Similar the Science- and Engineering-Related Domain is composed of two sub-domains: (1) Architecture Occupations and (2) Health Occupations. Within each of these domains there are five types of STEM occupations identified: (1) Research, Development, Design, or Practitioner Occupations; (2) Technologist and Technician Occupations; (3) Postsecondary Teaching Occupations; (4) Managerial Occupations; and (5) Sales Occupations. These domains and occupations create a number of combinations for measuring STEM employment. In an effort to explore the variation in employment concentrations and H-1B visa demand this report breaks STEM into three components: computer and mathematics, science and engineering, and Broad STEM.

At the core of all STEM definitions are the Research, Development, Design, or Practitioner Occupations and Technologist and Technician Occupations in fields such as Computers, Mathematics, Science, and Engineering. The large concentration of requests in computer occupations warranted its own analysis. However, to have adequate employment estimates at the state and metropolitan level we had to group together computer and mathematics to have robust geographic coverage. This likely understates the level of demand for Computer occupations in a number of labor markets, particularly the demand for highly requested occupations such as Systems Analysts and Computer Programmers. However, the demand for computer and mathematics occupations serves as a good proxy for demand for Computer occupations, balancing the tradeoff between geographic and occupational detail. The remaining core STEM occupations fit into either Science or Engineering occupations and were combined to form the science and engineering component. The science and engineering component is a combination of detailed occupations and is much more restrictive geographically.

[5] SOCPC members included the Department of Labor, Bureau of Labor Statistics and Employment Training Administration; the Department of Commerce, Census Bureau; the Department of Defense, Defense Manpower Data Center; the Equal Employment Opportunity Commission; the Department of Health and Human Services, Health Resources and Services Administration; the Department of Education, National Center for Education Statistics; and the National Science Foundation, National Center for Science and Engineering Statistics. For SOCPC STEM classifications see: http://www.bls.gov/soc/Attachment_A_STEM.pdf

Three other occupational groups can be broadly defined as STEM: Social Science, Architecture, and Health Occupations. Further, a number of STEM positions can be in Postsecondary Teaching, Managerial Occupations, and Sales Occupations. To capture this broader classification these occupational groups and STEM related positions were grouped together to form Broad STEM. More detailed analysis was conducted on sub-groups of these components but for brevity in the report these occupations were grouped together as no single Broad STEM component had significant levels of requests comparable to the core STEM components (See Table 2 for the occupational compositions of each STEM component).

Occupational Employment Survey

The OES survey is a mail survey measuring occupational employment and wage rates of wage and salary workers in nonfarm establishments in the 50 States and the District of Columbia. Guam, Puerto Rico, and the Virgin Islands are also surveyed, but their data are not included in national estimates. Estimates of occupational employment and occupational wage rates are based on a rolling three-year cycle. OES estimates are constructed from a sample of about 1.2 million establishments for six-panels. Each year, forms are mailed to two semiannual panels of approximately 200,000 sampled establishments, one panel in May and the other in November. For example, May 2012 estimates are based on responses from six semiannual panels collected over a 3-year period: May 2012, November 2011, May 2011, November 2010, May 2010, and November 2009. Estimates are benchmarked to the latest May and November employment totals from the Quarterly Census of Employment and Wages (QCEW) (i.e. May 2012 estimates are benchmarked to November 2011 and May 2012).[6] This report uses occupational employment estimates from the May OES for 2008-2012.

It should also be noted that with the OES estimates based on responses over a three-year period the employment values used in the denominator of the intensity of demand measures do not cover the exact same time period as the LCA totals in the numerator. The benchmarking to the latest employment totals from May and November somewhat mitigates this concern as this more closely aligns estimates with the fiscal year.

The OES provides the most detailed occupational employment information for states and metropolitan areas, but in a number of cases the detailed occupational information is suppressed.[7] However, the OES data indicates when information has been suppressed. In order to confidently measure the share of STEM employment and the intensity of LCA requests we needed to fill in missing information when possible with the best available information. To do this linear interpolation was used when possible to fill in the missing data in-between years (i.e. 2010 suppressed values were replaced with a linear interpolation between 2009 and 2011). Any remaining missing values were imputed based on the closest year (i.e. 2012 values were replaced with 2011 values if 2012 values were suppressed). As the survey is based on a three-year sample such interpolation and imputation is reasonable as there is typically significant overlap in the survey samples between the closest years.

To minimize the impact of interpolation and imputation on estimates, STEM measures were

[6] For a complete description of the Occupational Employment Survey sampling and methodology see http://www.bls.gov/opub/hom/pdf/homch3.pdf.

[7] Generally, information is suppressed when: (1) fewer than three firms responded for an occupation in a particular industry; (2) any one firm represents more than 50 percent of the employment in an occupation; (3) two firms combined represent more than 75 percent of the employment in an occupation.

dropped when the share of interpolated and imputed values exceeded 10 percent of employment or where 90 percent of employment for an estimate did not come from a consistent group of occupations overtime.[8] As a result, the STEM employment estimates provided in the report and appendix are conservative estimates. For example, with the computer and mathematical occupations being a major employment category in the OES only 4 out of 372 metropolitan areas fail the interpolation and imputation test. In comparison, the science and engineering employment estimates rely on detailed occupation codes which are often suppressed, resulting in 103 metropolitan areas failing the imputation and interpolation tests. An additional three metropolitan areas failed these tests for Broad STEM occupations. In an effort to be able to provide accurate estimates of total STEM employment, estimates of Broad STEM employment were suppressed if core STEM components did not yield reliable employment estimates, resulting in 106 metropolitan areas having no Broad STEM employment estimates (Table 3).

Matching the Estimates of STEM Requests and STEM Employment

When STEM LCA requests were combined with STEM employment measures to produce the intensity of demand measures the accuracy of all measures were checked by making sure the detailed occupations that make up each STEM component had matching employment estimates. Given that the administrative data in the LCA is prone to errors, there are instances where SOC codes are entered incorrectly. Further, suppressed OES employment information may result in some STEM intensity measures being artificially higher if employment information for an occupation is absent and there are a large number of requests for such an occupation. To avoid such inflated measures, any intensity measures where more than 10 percent of LCA requests are not matched to occupations with employment information were dropped. For example, in Burlington, VT there were 10 requests for scientists and engineers (largely for Computer Hardware Engineers and Biological Scientists, All Other) that had no employment information in the OES; equivalent to roughly 30 percent of all scientists and engineer requests in the NECTA. Although the misalignment could be a result of occupational coding errors in the LCA or suppressed employment values due to small local employment in such occupations, when such a high degree of mismatch between requests and employment information occurred the measures were dropped. As a result we were able to produce intensity of demand measures for H-1B workers in computer and mathematics for 360 metropolitan areas but were limited to only 164 metropolitan areas for both science and engineering and Broad STEM intensity of demand measures.

Intensity of Demand for STEM Components: Three-Year Averages

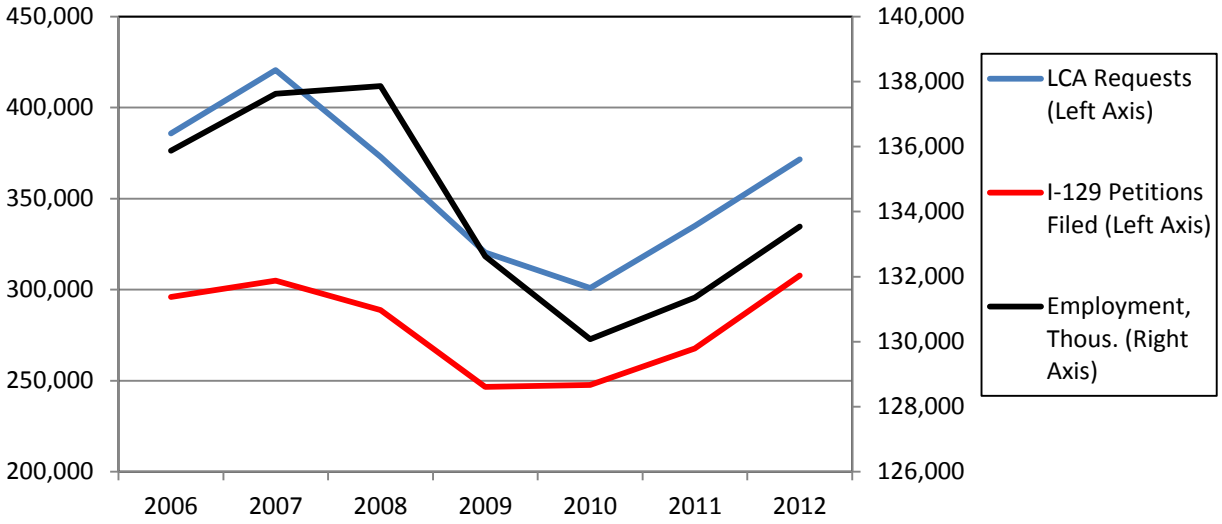
STEM intensity of demand measures could only be produced for 2010 through 2012, due to the limitations with occupational classifications listed above. Given the short-time series, along with missing employment information from the OES and issues with LCA requests not matching OES employment estimates, STEM intensity of demand measures presented in the report are three-year averages between 2010 and 2012 (e.g. average LCA requests 2010-2012 per average 1,000 payroll employees 2010-2012). Averaging over the three-year period helps to simplify the analysis of the intensity of demand measures and improve the geographic coverage by minimizing the number of measures that were dropped due to imputation or unmatched occupational requests and employment.[9]

[8] The selection of these cut offs tries to address the limitations of the OES as it transitioned between the 2000 and 2010 SOC coding systems. The first year the OES produced estimates using 2010 SOC codes was 2012. In 2010 and 2011 they used mostly 2010 SOC codes, but also included hybrid 2000/2010 SOC codes that required the combination of many codes to be combined to be able to crosswalk codes across all years.

[9] In a number of cases, the imputation of employment estimates or matching of LCA requests with employment would lead to the removal of geography due to a single year missing the cut off requirements. Averaging over

three-years preserved many of these geographies where a single missing observation was the issue and only excluded the geographies with the most consistently flawed estimates.

Figure 1. H-1B Applications: LCA and I-129 Forms, FY 2006-2012



Source: U.S. DOL LCA Database 2008-2012, OFLC Annual Performance Reports 2006-2007, U.S. Costums and Immigration Enforcement Speciality Occupation Reports, and U.S. Bureau of Labor Statistics.

Note: Fiscal Years run from October 1st to September 31st. Monthly national employment estimates are averaged over October through September of each respective fiscal year for comparison to the LCA and I-129 data.

Table 1. Labor Condition Application Sample Selection by Fiscal Year

	2008	2009	2010	2011	2012
Total Requests	405,641	356,231	342,600	358,857	415,845
Withdrawn (-)	29,934	33,307	39,458	21,748	42,028
US Territories (-)	751	751	998	804	1,068
Requests with Determinations	374,956	322,173	302,144	336,305	372,749
Unidentified Geography (-)	2,042	1,807	1,184	1,175	1,129
Unidentified Occupation (-)	X	X	45	40	63
Final Data Set	372,914	320,366	300,915	335,090	371,557
as % Determinations	99.5%	99.4%	99.6%	99.6%	99.7%
Determination					
Certified	372,000	300,962	260,975	306,255	350,848
Denied	914	19,403	39,937	28,834	20,709
Geography of H-1B Request:					
Metropolitan Area	363,596	310,580	291,234	325,193	361,101
Micropolitan Area	6,768	7,042	6,914	7,148	7,884
Rural Area	2,550	2,743	2,764	2,748	2,572

Source: Office of Foreign Labor Certification, Labor Condition Application.

Note: Withdrawn LCA applications were incomplete and missing key information in a number of fields. For consistency only complete requests with determinations were included in the final data set. The national employment estimates from the Occupational Employment Survey (OES) are produced by the U.S. Bureau of Labor Statistics using data from the fifty states and the District of Columbia. For consistency in geographic composition requests for H-1B visas in U.S. territories were dropped. Occupation codes in 2008 and 2009 were from the Dictionary of Occupational Titles and were not evaluated for accuracy as all occupation level analysis was done using the Standard Occupational Classification (SOC) system for 2010-2012.

Table 2. Occupational Composition of STEM Components based on SOC Policy Committees STEM Definition

		STEM Domains					
		STEM Domain			Science- and Engineering-Related Domain		
		Life and Physical Science, Engineering, Mathematics, and Information Technology		Social Sciences	Architecture	Health	
		Computer & Mathematics	Science & Engineering	Broad STEM			Total
Occupation Groups	Research, Development, Design, or Practitioners	16 Occupations (e.g. 15-1121 Computer System Analysts)	39 Occupations (e.g. 17-2071 Electrical Engineers or 19-1022 Microbiologists)	11 Occupations (e.g. 19-3011 Economists)	2 Occupations (e.g. 17-1011 Architects)	35 Occupations (e.g. 29-1067 Surgeons)	103
	Technologist and Technician	4 Occupations (e.g. 15-1131 Computer Programmers)	24 Occupations (e.g. 17-3023 Electro-Mechanical Technicians or 19-4021 Biological Technicians)	1 Occupation (19-4061 Social Science Research Assistants)	None	27 Occupations (e.g. 29-2055 Surgical Technologists)	56
	Postsecondary Teaching	2 Occupations (e.g. 25-2021 Computer Science Teachers, Postsecondary)	8 Occupations (e.g. 25-1032 Engineering Teachers, Postsecondary	8 Occupations (e.g. Sociology Teachers, Post Secondary)	1 Occupation (Architecture Teachers, Postsecondary)	2 occupations (e.g. 25-1072 Nursing Instructors and Teachers, Postsecondary)	21
	Managerial Occupations	1 Occupation (11-3021 Computer and Information Systems Managers)	2 Occupations (e.g. 11-9121 Natural Science Mangers)	None	None	1 Occupation (11-9111 Medical and Health Services Managers)	4
	Sales Occupations	None	2 Occupations (e.g. 41-9031 Sales Engineers)	None	None	None	2
Total		23	75	20	3	65	186

Source: SOC Policy Committee recommendation to the Office of Management and Budget for defining STEM.
Note: In total 186 out of a total of 848 SOC Codes can be classified as STEM under the SOC Policy Committee definition. Computer & Mathematics covers all 20 SOC Codes in the major occupation group Computer & Mathematics (15-0000). The Health sub-domain in the Science & Engineering-Related Domain covers all 62 SOC Codes in the major occupation group Healthcare Practitioners and Technical Occupations (29-0000).

Table 3. Geographic Coverage of Intensity of Demand Measures by STEM Component

	Metro Areas Failing Imputation and Interpolation	Metro Areas Failing LCA and OES Merge	Metro Areas with Intensity of Demand Measures for STEM Component
Computer & Mathematics	4	8	360
Science & Engineering	103	105	164
Broad STEM	106	102	164

Note: Out of a total of 372 MSA and NECTAs.