H-1B Prevailing Wage Claims vs. Reality: What They Say Versus What They Pay

Robert T. Hill

CUNY School of Professional Studies

Murphy Institute for Worker Education

M. A. Labor Studies Program

Labor 699 Capstone Project

May 2011

Glossary of Acronyms

Acronym	Explication
DOL	United States Department of Labor
IMMACT90	Immigration Act of 1990
OES	Occupational Employment Statistics. The Occupational Employment Statistics (OES) program produces employment and wage estimates for over 800 occupations. These are estimates of the number of people employed in certain occupations, and estimates of the wages paid to them. Self-employed persons are not included in the estimates. http://www.bls.gov/oes/
OFLC	Office of Foreign Labor Certification. A division of the Department of Labor. http://www.foreignlaborcert.doleta.gov/
SOC	Standard Occupational Classification codes. The 2010 Standard Occupational Classification (SOC) system is used by Federal statistical agencies to classify workers into occupational categories for the purpose of collecting, calculating, or disseminating data. All workers are classified into one of 840 detailed occupations according to their occupational definition. http://www.bls.gov/soc/
USCIS	United States Citizenship and Immigration Services. http://www.uscis.gov

Table of Contents

1. Introduction

The H-1B visa program allows U.S. employers to hire temporary, foreign "guest workers" to fill certain kinds of jobs when they can't find enough properly trained U.S. citizens. The program attempts to strike a balance between employer and employee concerns. Employers can fill the jobs and get their work done in the U.S. rather than ship the jobs overseas, and employees are protected from exploitation by a set of rules and procedures. For example, employers are required to formally declare that they have been unable to find and hire qualified U.S. citizens. Further, they must declare that they will pay these guest workers "at or above prevailing wages for the same jobs in the same work locations." Many other qualifications and protections are also applicable.

Hundreds of thousands of H-1B guest workers have been processed since the program started in 1990, and the government has long provided data and statistics about their job titles, "prevailing wages," and "actual wages." Research attempts have been made to compare the claimed prevailing wages to actual prevailing wages, but problems have been identified with the accuracy and relevancy of the data. For instance, an employer may have reported that they hired a H-1B "Computer Consultant" at a claimed prevailing wage of $45,000 per year, but there was no way to compare this job title to official statistics from the Bureau of Labor Statistics because the job title "Computer Consultant" does not exist. Further, the employer was allowed to claim the prevailing wage from a variety of sources, some of which were quite questionable. Neither job titles nor prevailing wage claims were standardized, so accurate analysis was difficult, if not impossible. Apples were being compared to oranges.

A recent change was made in the way that employers file their claims. The new, web-based iCert system solves many of the previously identified problems by requiring that job titles be drawn from the Standard Occupational Classification (SOC) codes used by the Department of Labor (DOL), and it also draws prevailing wage data directly from the DOL's Occupational Employment Statistics (OES) wage estimate library. By requiring that H-1B jobs be reported using standard SOC and OES codes, it is now possible to directly compare the claimed H-1B "prevailing" and "actual" wages to those of the same U.S. citizen workers in the same locations. With the new system, it is possible to compare apples to apples.

H-1B job and wage data is reported by the Foreign Labor Certification Data Center, and fiscal year 2010 is the first year that this report contains ONLY data collected by the new iCert system.

This study isolates the claimed prevailing and actual wages that employers reported for the sub-set of computer technology worker Labor Condition Applications (LCAs), some 89,585 approved LCAs, representing 232,487 requests for workers, granted for fiscal year 2010. It groups these LCAs by the set of 10 SOC codes associated with computer technology workers, e.g. programmers, analysts, and customer support workers. It then compares the LCA wages to OES wage data for the same SOC codes, as reported for domestic workers. Results indicate that LCA wages for H-1B workers are typically much lower than those for domestic workers, but that these wages are still within the legal requirements because the prevailing wage estimates, although relatively low, are provided by the Department of Labor.

2. *Brief History and Overview of the H-1B Visa Program*

The predecessor to the H-1B program was the H-1 program. The H-1 nonimmigrant category was created under the Immigration and Nationality Act of 1952, also known as the McCarran-Walter Act. Created during the so-called Cold War Era, this act continued the national origins quota system established by the Immigration Act of 1924, but it augmented the system in two important ways. First, it extended immigration limitations on citizens from countries that had been or were at war with the United States, such as Japan, Germany, and the Soviet Union. Second, it established a new precedent of giving immigration priority to "individuals with special skills."[1]

One of the first widespread applications of this new "special skills" provision was observed in the sheep herding industry in the western U.S. Michael Yatsko (1997) describes how the shortage of sheep herders in the American West during World War II fueled the illegal immigration of thousands of Basque sheep herders during the war.

> The migration of Basques to the United States continued due largely to the shortage of herders in the American West during the Second World War. The labor shortage became so acute that representatives from several western states introduced a series of bills, aptly called the Sheepherder Laws, which granted permanent residency to those Basques who had entered the country illegally to become sheep herders. Despite the Sheepherder Laws remaining in effect after World War II, the shortage of herders continued which prompted a number of American sheep organizations, such as the Western Range Association and the California Woolgrowers Association, to pressure the federal government for relief. The response of the federal government was the passage of the

[1] Department of State, US Government (2011). Milestones, 1945-1952: The Immigration and Nationality Act of 1952 (The McCarran-Walter Act). (Last accessed online http://history.state.gov/milestones/1945-1952/ImmigrationAct May 2011).

> McCarran-Walter Act to allow these sheep organizations to recruit foreign herders, exempted from the immigration quotas, for three-year contracts. While the McCarran-Walter Act, specifically, did not cite the Basques as the only ethnic group eligible for this program, the principal applicants were Basques. Nearly 5,500 Basques applied to the program between 1957 and 1970.[2]

This new practice of identifying a class of workers "of distinguished merit and ability" was to become a significant feature of the next major revision of immigration law. According to a study published by the United States General Accounting Office in 2000, this shift of attention away from the national origin quota system toward the needs of employers was a primary motivator behind the Immigration Act of 1990.[3]

The Immigration Act of 1990,[4] commonly known as "IMMACT90," was designed to greatly increase the number of foreign technology workers working in the United States for U.S.-based businesses. Its passage was urged by the National Science Foundation (NSF), among others, which had predicted that there would be a sustained and debilitating shortage of skilled technology workers in the United States unless guest workers were attracted into the country.

---

[2] Yatsko, Michael S. (2007) Ethnicity in Festival Landscapes: An Analysis of the Landscape of Jaildi '95 as a Spatial Expression of Basque Ethnicity, Chapter 3. Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science in Geography. (Last accessed online at http://scholar.lib.vt.edu/theses/available/etd-2230102449761431/unrestricted/etd7_chap3.pdf May 2011).

[3] General Accounting Office, US Government (2000). H-1B Foreign Workers: Better Controls Needed To Help Employers and Protect Workers. *United States General Accounting Office, Publication GAO/HEHS-00-157, September 2000.* (Last accessed online http://www.gao.gov/archive/2000/he00157.pdf May 2011).

[4] Immigration Act of 1990, Pub. L. No. 101-649, 104 Stat. 4978 (1990).

The IMMACT90 replaced the H-1 visa system with the revised H-1B system. The H-1B visa program was designed to correct some perceived shortcomings of the previous H-1 visa program in three ways. First, it was widely believed that while the old H-1 visa program was intended to attract the world's "best and brightest" workers, termed "Aliens of Distinguished Merit and Ability," in actuality it had devolved into simply requiring that H-1 workers hold a Bachelor's degree or higher. Second, in response to concerns over the misuse of the H-1 program to favor foreign workers over domestic ones, the H-1B system was designed to make conditions for granting the visas more precise, add some protections for domestic workers, and allow "dual-intent" status so that guest workers could pursue standard immigration pathways while they worked in the U.S. using their H-1B visas.[5] Third, it imposed limits to the number of visas that could be awarded each year to further protect the domestic workforce from loss of domestic jobs.

According to a report published by the United States General Accounting Office in 2000, the H-1B visas are available to people in so-called "specialty occupations."

> "Under the H-1B program specialty occupations are those requiring theoretical and practical application of a body of specialized knowledge and the attainment of a bachelor's or higher degree (or its equivalent) in the specific specialty.

[5] Matloff, Norman (2003). On The Need For Reform Of The H-1B Non-Immigrant Work Visa In Computer-Related Occupations. University of Michigan Journal of Law Reform, Volume 36:4. (Last accessed via http://heather.cs.ucdavis.edu/Mich.pdf May 2011).

These can be in a range of fields from architecture, engineering, and mathematics to medicine, education, theology, and the arts." [6]

While the IMMACT90 placed a limit of 65,000 H-1B visas annually, this number is subject to Congressional adjustment every year. The initial limit was not reached until fiscal year 1996, but demand for H-1Bs has exceeded that limit every year since. Accordingly, Congress has increased the number several times in successive years. For instance, the limit was increased to 115,000 in FY 1999 and 2000 and 107,500 in FY 2001. It was further increased to 195,000 for FY2001 through FY2003, and has remained at 65,000 every year since FY2004.[7]

Competition among international workers who are seeking to secure one of the 65,000 annual H-1B visas is fierce. For FY2011, The Office of U.S. Citizenship and Immigration Services (USCIS) announced on January 26, 2011, that so many applications had already been received that additional applications after that date would be rejected. [8] Those H-1B applications that were received were to be subjected to a computer-generated random selection process until the 65,000 limit was reached. Those applications not selected were to be returned, along with the filing fees.

---

[6] General Accounting Office, US Government (2000). H-1B Foreign Workers: Better Controls Needed To Help Employers and Protect Workers. *United States General Accounting Office, Publication GAO/HEHS-00-157, September 2000.* (Last accessed online http://www.gao.gov/archive/2000/he00157.pdf May 2011).

[7] Ibid.

[8] U.S. Citizenship and Immigration Services. H-1B Fiscal Year (FY) 2011 Cap Season. (Last Accessed online at http://www.uscis.gov May 2011).

3. The LCA and H-1B Process Overview

It is important to understand that one Labor Condition Application (LCA) does not equal one proposed job, nor one H-1B visa granted. Employers file LCAs to indicate their needs for employees for specific jobs, and there is no limit to the number of jobs that employers can file LCAs for. Additionally, while most LCAs represent the need for one employee, it is common for LCAs to claim the need for multiple positions. In the present study, one single LCA for computer technical support workers claimed the need for 250 workers. In total, the 89,585 computer technology LCAs represent 232,487 open job positions.

For fiscal year 2010, there were a total of 335,328 LCAs filed. Of those LCAs, 89,585 were designated as "computer technology" LCAs.[9] Therefore, computer technology jobs represent about 26.7% of all LCAs filed for fiscal year 2010. However, the Congressionally mandated limit of 65,000 H-1B visas is applied to the total number of LCAs, not just the computer technology workers. It is not possible for each of the LCA positions to be filled; only about 1 in 5 of the FY2011 LCAs can possibly be filled, across all SOC codes.

A more detailed overview of the end-to-end process for employing a foreign worker under the H-1B visa program can be found at the Department of Labor's website for the Office of Foreign Labor Certification (OFLC) (http://www.foreignlaborcert.doleta.gov). Briefly, the process is outlined as follows.

[9] "Computer Technology LCAs" means they were assigned SOC codes in the "15-10**" series.

1. **Logging in to the iCERT System.** As of April 2009, the Department of Labor's Employment and Training Administration has made available an online tool for filing Labor Condition Applications (LCAs). This tool, called the iCert Visa Portal System, is used by employers and their agents to create and track LCAs. The employer or employer agent establishes an account in the iCert system and logs in to create a new LCA as well as to manage and track previously submitted LCAs.

2. **Determination of the relevant Prevailing Wage (PW).** As part of the LCA creation process, employers must first determine the Prevailing Wage (PW) for each job that they wish to fill, which varies depending on the location where the work is to be performed. For instance, a Database Administrator working in New York City will have a different PW than someone doing the same job in Houston, Texas. Additionally, the PW is dependent on the worker's level of experience.

   There are several methods allowed for determining the appropriate PW, but the vast majority of LCAs now use the iCert's integration with the Bureau of Labor Statistics' Online Wage Library. This allows for the direct insertion of Occupational Employment Statistics (OES) estimated wage data to the LCA as a method of assuring that LCA wages are "at or above" the true prevailing wages. However, in situations where there is a collective bargaining agreement that establishes the prevailing wage, that agreement takes precedence. Figure 1 shows a sample prevailing wage search for a Computer Programmer working in the New York City metropolitan area for the period of 2010-11. Note that it displays four estimated wage levels, reflecting different prevailing wages for different levels of experience.

Figure 1: iCert Prevailing Wage search tool.

3. **Filing the Labor Condition Application (LCA).** As part of the filing process, the employer must provide specific information about the nature and location of the work to be performed as well as the steps they have taken to fill the job with a domestic worker. According to the OFLC "by completing and signing the LCA, the employer agrees to several attestations regarding an employer's responsibilities, including the wages and benefits and working conditions provided to US workers and the nonimmigrant workers."

4. **Posting the approved LCA for public inspection.** Employers are required to "make the LCA and necessary supporting documentation available for public examination at the employer's principal place of business in the U.S. and/or the place of employment within one working day after the date on which the LCA is filed with ETA." Employers must post notices in "conspicuous public places" in the workplace, and they may also send out communications to other employees via email. In addition, in situations where there is a collective bargaining agreement, they must provide this information to the bargaining representative associated with the workplace.

5. **Submitting the forms to USCIS.** Employers must then submit the approved LCAs to the USCIS, along with the I-129 forms that identify the workers that they wish to hire.

6. **Beginning work.** The employer must not allow the H-1B guest worker to begin work until after the certified LCA and USCIS I-129 forms are filed and approved. H-1Bs are approved for specific fiscal years, with limits on how long they are to remain active. For example, workers approved for FY2010 can submit their applications starting in April 2010, but they may not begin work until the fiscal year starts in October 2010.

4. Arguments For and Against the H-1B Program

There is a long-running argument about whether the technology business community in the United States is justified in its practice of hiring "foreign guest-workers" via the H-1B program instead of domestic, "citizen" workers. Businesses have long claimed that they have jobs that must be filled immediately (or else, the U.S. will suffer significant economic demise), but they cannot find U.S. citizens with the required training and skills to fill these jobs. Therefore, they argue, Congress must allow business to bring in increasingly greater numbers of foreign "guest-workers" to work in the United States, or else the jobs will leave the U.S. and go to where the workers are. An article from the Wall Street Journal in 2007 is typical of the press coverage of the issue and refers to the Congressional testimony given by Bill Gates, Chairman of Microsoft, one of the largest consumers of H-1B visas.

> "Congress is under pressure from employers' groups to vastly expand the number of visas available each year, and is generally in favor of the idea. Employers say there aren't enough visas to meet their needs, even though the visas are renewable, and Congress added 20,000 visas this year for foreigners who have at least a master's degree from a U.S. college. Microsoft chairman Bill Gates testified on behalf of the program on Capitol Hill, warning of dangers to the economy if employers can't import skilled workers to fill job gaps."[10]

A more recent article in Business Week magazine reports that the need for additional H-1B workers is growing ever-more acute and that the consequences for failing to increase the H-1B limits include the permanent loss of technology jobs to other countries. Writing in January 2011, Business Week columnist Frank Aquila claims:

---

[10] Wall Street Journal (2007). Visa Window Opens; Scramble Is About To Begin. Wall Street Journal, March 28, 2007. (Last accessed online at http://blogs.wsj.com/washwire/2007/03/28/visa-window-opens-scramble-is-about-to-begin May 2011).

> The U.S. Labor Department estimates that by 2014, 2 million high-tech jobs will go unfilled simply because the cap on H-1B visas has not been raised. Tech giants such as Google and Apple will no doubt move significant development projects out of the U.S. to places where these skilled workers are available. Smaller high-tech businesses, historically the engine of U.S. job growth and creation, will simply never get off the ground. The consequences are clear: the next generation of innovative companies will not likely be founded here. Instead, due to U.S. policy, these companies will most probably be created in places such as India, China, and Singapore.[11]

Critics of the practice have long claimed that domestic workers are readily available, but that businesses routinely hire foreign H-1B workers because they can and will do the same work for significantly lower wages than the "prevailing wages" that domestic workers expect. To protect against this possibility, H-1B visa legislation includes statutory requirements that compel businesses that hire H-1B guest workers to demonstrate that they are a) actively seeking (and not displacing) domestic workers, and b) paying H-1B guest workers "at or above" the so-called "prevailing wages" for domestic workers.

However, there is widespread suspicion that these statutory requirements are either ignored or manipulated so that their net effect is essentially meaningless. Writing in The Wall Street Journal, columnist Paul Donnelly (2002) even quoted Nobel Prize winning economist Milton Friedman as starting unequivocally that the H-1B system was a "subsidy."

> "During the last H-1B debate, Harris Miller, the head of the ITAA, told an interviewer at the Chicago Tribune that the H-1B is a kind of 'minor league,' a farm team for the IT industry. This spring, I wrote to Nobel prize winning economist Milton Friedman, practically the patron saint of the free market, citing

[11] 10. Aquila, Frank (2011). H-1B Visas: A Modest Proposal for Immigration Reform. *BusinessWeek.com, January 20, 2011.* (Last accessed via http://www.businessweek.com/investor/content/jan2011/pi20110118_876603.htm May 2011).

> Miller's quote and asking: "What is a subsidy?" This is his reply: "'The majority of H-1B immigrants do manage by hook or crook to get permanent residence and become citizens, so as a factual matter they are not a 'farm team' of indefinitely temporary workers. Yet, there is no doubt that the program is a benefit to their employers, enabling them to get workers at a lower wage and to that extent is a subsidy.'"[12]

Presumably, the late Milton Friedman knew a subsidy when he saw one.

[12] Donnelly, Paul. (2002). The H-1B Is A Barrier To The IT Industry's Recovery: Even Milton Friedman Says "there is no doubt" H-1B Is A Subsidy. *Immigration Daily, August 16, 2002.* (Last accessed via http://www.ilw.com/articles/2002,0816-Donnelly.shtm May 2011).

5. Previous Research

In 2003, after the third consecutive year when the H-1B cap had been raised to 195,000 annually, Professor Norman Matloff [13] published an exhaustive meta-study of previously published research on matters concerning reform of the H-1B program.[14] His 100-page article, published in the University of Michigan Journal of Law Reform, was essentially an argument for significant reforms in the H-1B program to address several glaring problems.

The annual caps of 195,000 workers for fiscal years 2001 through 2003 were granted by Congress in large part because the software industry complained aggressively that there were dire shortages of adequately trained domestic software workers. Matloff consulted a wide variety of studies that found that there was indeed no shortage of domestic computer technology labor, but rather an unwillingness of businesses to hire them.

In addition to consulting the peer-reviewed literature, though, his study was bolstered by observations of actual data. Figure 2 below is a table taken from Matloff (2003) that demonstrates the extremely low hiring rates by the software industry's top employers.

[13] Professor of Computer Science, University of California, Davis; B.S. 1970, California State Polytechnic University; Ph.D. (pure mathematics) 1975, University of California, Loa Angeles.

[14] Matloff, Norman (2003). On The Need For Reform Of The H-1B Non-Immigrant Work Visa In Computer-Related Occupations. University of Michigan Journal of Law Reform, Volume 36:4. (Last accessed via http://heather.cs.ucdavis.edu/Mich.pdf May 2011).

TABLE 2
PERCENT OF SOFTWARE APPLICANTS HIRED

American Management Systems	2%
Broderbund Software	1%
Cisco	5%
Cohesive	2%
Datascan	5%
Deltanet	4%
Ecbridges	2%
Flashpoint Technology	2 to 5%
R.D. Raab	1%
H.L. Yoh	4%
Inktomi	less than 5%
Microsoft	2%
Net Perceptions	2%
New England firm	1%
Qualcomm	4.5%
Radiant Systems	under 1%
Red Hat Linux	under 1%
Tangis	under 1%

Figure 2: Table 2 from Matloff (2003) showing extremely low hiring rates.

These numbers are percentages of applicants hired. For instance, while Microsoft Chair Bill Gates was repeatedly testifying before Congress that his corporation was unable to find trained domestic workers, his company was actually hiring only about 2% of the people who submitted their resumes for the jobs. Matloff points out that far from being unable to find employees, Microsoft was able to dismiss 98% of the people that sought employment with it. Further, of the 18 top computer technology employers in Matloff's survey, not one of them hired more than 5% of their applicants. This factual finding was inconsistent with the claim that qualified employees were unavailable. As well, Matloff argues that bringing in an additional 195,000 workers per year would only make these percentages smaller, not larger.

The real reason, Matloff argues, that the software industry was seeking ever more H-1B workers is that such workers represent cheap, compliant labor. He cites the Cappelli Principle: workers are available, but not at the price that employers want to pay. The Cappelli Principle is especially relevant to computer technology workers, he argues, because as employers demand highly-specialized skills, accepted market and labor philosophies agree that such skills demand higher wages. Figure 3 below demonstrates the high salary premiums that employers were forced to pay to attract people with these highly-specific skillsets.

TABLE 4
SALARY PREMIUMS FOR SPECIFIC SOFTWARE SKILLS

Oracle DBMS tool	24%
SAP	24%
HP UNIX	20%
Visual C++	20%
Java	16%

Figure 3: Table 4 from Matloff (2003), showing elevated wages for high-demand skills.

Matloff points out that foreign, H-1B workers are not only able to provide these skills, but by increasing the supply of labor employers are relying on the well-accepted market principle that as supply goes up, cost goes down. Thus, the argument is made that the true goal of increasing the supply of H-1B workers is empirically demonstrated to be driven by finances, not supply.

In addition to this evidence, Matloff draws from a number of studies documenting the differences between H-1B and domestic wages. Figure 4 below highlights a few comparisons from one of the studies that Matloff surveys.

TABLE 11
FOREIGN-NATIONAL VS. LABOR MARKET SALARIES

JOB	STATE	MEAN FOREIGN-NATIONAL WAGE	MEAN MARKET WAGE
Comp. Sys. Analysts and Scientists	NJ	$15.24	$21.64
Comp. Sys. Analysts and Scientists	NY	$16.28	$20.57
Comp. Programmers	NJ	$15.65	$19.74
Comp. Programmers	TX	$14.12	$15.76

Figure 4: Table 11 from Matloff (2003), demonstrating typical differences in H-1B vs domestic salaries.

In each of the four instances he identifies in his Table 11 (Figure 4 above), mean foreign-national wages were significantly lower than mean market wages for similar work performed in the same work places.

Additionally, Matloff identifies two types of savings in labor costs that are consequences employers' ever-increasing use of H-1B labor. He calls these Type 1 savings and Type 2 savings. Type 1 savings to the employer represent the direct savings accrued by simply paying the H-1B workers less than they would pay domestic workers. For instance, rather than paying a premium of 24% for a domestic Oracle DBMS developer they might be able to pay much less for

an H-1B worker with the same skills. Type 2 savings, though, represent another widely suspected financial motivation for employers: the perception that older workers cost more to hire than do younger workers, and that H-1B workers are categorically younger people in addition to being categorically willing to work for lower wages. By favoring H-1B workers whenever possible, employers can increase savings on both salary and age dimensions.

Finally, Matloff calls on a variety of governmental and academic studies that indicate that H-1B workers are not only categorically less expensive than domestic workers, they are also much more compliant.

> "Since an H-1B is typically in no position to seek other employment (due to sponsorship requirements), the employer need not worry that the worker will suddenly leave the employer in the middle of a pressing project. In addition, the employer can force the H-1B to work long hours. To many employers, this "loyalty" aspect is the prime motivation for hiring H-1Bs, whether or not they are saving salary costs in doing so."

Matloff refers to H-1B workers as "de facto indentured servants" because most H-1B workers are hoping to convert their temporary visa status to permanent status via their employer.

> "This is a multi-year process. Toward the end of the 1990s, the processing time for the two largest H-1B nationalities, Indian and Chinese, was approaching six years. During the time an H-1B's green card application is being processed, he/she is essentially immobile; switching employers during this time would necessitate starting the green card process all over again, an unthinkable prospect for most.

That is to say, despite the attempted protections against exploitation of the H-1B workers themselves, Matloff argues that most if not virtually all of the "loyalty," expressed as working

14-hour days for low salaries, is actually a modern form of indentured servitude that translates to economic benefits to employers.

For all the strengths of Matloff's arguments, and in spite of the thoroughness of his 100-page, peer-reviewed journal article, Matloff points out that the inadequate record keeping practices of the LCA and H-1B administrators and employers lead to an inability to determine accurate job titles and prevailing wages. The final part of his article argues for needed reforms. While his list of needed reforms does not call for the adoption of standardized SOC job codes and standardized OES prevailing wage data, he does acknowledge that these record keeping problems make it inherently difficult to accurately assess the available data.

In a follow-up effort to determine whether the prevailing wages that were reported on the 2004 and 2005 LCAs was an accurate reflection of the true prevailing wages, Miano (2007) undertook an exhaustive study of the data that was available at that time for FY2005.[15] Miano's key findings included the following:

- Very few of the H-1B workers were classified as "highly skilled," as argued by the entry level wages that the employers reported. "Employers who used the Department of Labor's skill-based prevailing wage system (as available at that time) classified most workers (56%) as being at the lowest skill level (Level 1)."
- For FY2005, the employer prevailing wage claims averaged $16,000/yr below the median wage for U.S. computer workers in the same location and occupation.

[15] Miano, John (2007). Low Salaries For Low Skills: Wages and Skill Levels for H-1B Computer Workers, 2005. Center for Immigration Studies, Backgrounders and Reports, April 2007. (Last accessed via http://www.cis.org/articles/2007/back407.pdf and http://www.cis.org/LowSalariesforLowSkills-H1B May 2011).

- 89% of the H-1B employer prevailing wage claims for programming occupations were below the median U.S. wage for the same occupation and location, with 62% of the wage claims in the bottom 25th percentile of U.S. wages.

- While higher than the prevailing wage claims, the actual wages reported for H-1B workers were significantly less than those of their American counterparts. That is, while employers were paying "at or above" what they claimed to be the prevailing wage, it was still significantly lower than the prevailing wages found by examining OES data.

- Many employers apparently made prevailing wage claims using wage sources that were not valid under the law. The Department of Labor routinely approved wage claims based on these invalid sources.

There were several problems that constrained Miano's research. One is that the data for FY2005 was filed using a combination of two older systems, each of which stored data in differently constructed data structures. The "e-File" system stored one set of information with its naming structure or so-called "schema," and the "e-FAX" system stored a different set of information using a different schema. Miano included both sources in his calculations, but reported that as the e-FAX system represented a small percentage of the total records it could be disregarded in future research. One of the benefits of the transition to the newer iCert system in April 2009 is that this problem is no longer present for current data: all records are stored in a unified database schema.

Another problem that Miano found was that there were various sources of prevailing wages used for the LCAs, and some of them were apparently invalid under the law. For instance, some employers based prevailing wages on surveys provided by private colleges reporting the starting salaries of their graduates. If anything, this data would reflect starting salaries for new graduates, not prevailing wages for experienced domestic workers. And, such data would not

necessarily relate to the national or regional prevailing wages, contrary to the requirements set forth in the IMMACT90.

While 70% of the LCAs (representing about 50% of the workers) claimed to have used OES data as the basis for the prevailing wages, the FY2005 data was influenced by a change in the law made in 2004.

> "As discussed, in 2004 Congress added a new prevailing-wage option for employers. It mandated that the Department of Labor provide employers with four skill-based prevailing wages. To comply with this change, the Foreign Labor Certification Data Center took the OES data produced by the Bureau of Labor Statistics and used them to create four skill-based prevailing wages. This created a mechanism for employers to justify low wages. Regardless of the actual skills of an H-1B worker, employees need only assert that a worker is in the Level One category for entry level, trainee, or intern employees, and pay according to that prevailing wage."

Miano calculated that 89% of the H-1B workers were assigned prevailing wages that were below the median prevailing wage for their job code at their location. He deduced, therefore, that the workers were given the prevailing wages for the lower skill and experience levels. Figure 5 below, from Miano (2007), demonstrates this skew graphically.

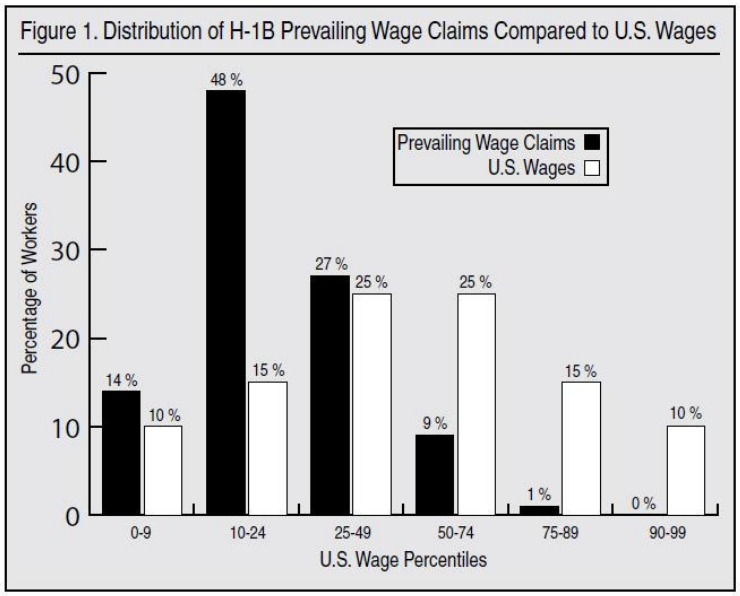


Figure 5: Figure 1 from Miano (2007) showing skewed distribution of wages.

The two bar charts in Figure 5 refer to prevailing wage claims. The black bars represent the prevailing wages claimed for H-1B workers in FY 2005, and the white bars represent the OES wages for domestic workers in the same field. As the chart demonstrates, both the black distribution and white distribution show a roughly "normal" Bell curve shape.[16] However, the H-1B distribution is heavily skewed to the left, indicating categorically lower prevailing wage claims. Further, 62% (14% plus 48%) of the H-1B prevailing wage claims fell below the 25th percentile rank for domestic workers in similar work categories.

The biggest problem that Miano faced was that the LCAs in the FY2005 data used a unique 3-digit occupational code rather than Standard Occupational Classification (SOC) codes. In

[16] An overview of Normal Distribution and the Bell curve .http://en.wikipedia.org/wiki/Normal_distribution

order to compare the jobs in the LCAs to the Occupational Employment Statistics (OES) data, each LCA had to be associated with the appropriate SOC code. The problem was, the 3-digit codes used by the FY2005 LCAs were much broader than the SOCs. Managers of programmers would be assigned the same 3-digit code as the programmers themselves, making it difficult to compare the prevailing wage claims accurately.

Compounding this problem with the data was the fact that job titles were unreliable, insufficiently descriptive, or both. Therefore, Miano faced a huge problem of working to assign LCAs to their appropriate SOC codes so that the data could be reliably compared to the OES data. Problems included instances where the job title was simply "Consultant." Or, the job titles might reflect "Computer Programmer" but fail to distinguish between an "Application" programmer and "Systems" programmer. OES data are different for those two categories of programmer.

Additionally, because of the sheer number of records (over 300,000 LCAs for FY 2005, covering about 700,000 workers – some LCAs are for multiple people), pattern matching techniques were used to try to group similar jobs together. For instance, sorting with a "wildcard" string like "soft*eng*" would return job titles and descriptions that were similar yet not exactly the same. Miano made an effort to discern from the other available data in that LCA, but eventually had to assign some SOC code to every one of them and he acknowledges that this is a weakness in the design of his study.

Miano concludes his report with a set of recommendations, including the following.

> Compel employers to use a standard wage source produced by the federal government when making prevailing wage claims for LCAs. This report has shown that employer prevailing wage claims are much lower than the actual prevailing wages, and that some employers make mistakes or deliberately provide bogus prevailing wage claims.

As will be shown in the next section of this study, this recommendation was implemented subsequent to Miano's report.

6. Recent Changes in Data Collection – the iCert System

This study replicates previously published research (Miano, 2007) on the question of whether employers are categorically paying H-1B visa "guest workers" lower wages than they would have to pay domestic workers in specified jobs. Previous research was plagued by systemic problems, however, that made it both difficult and arbitrary to compare H-1B data to Occupational Employment Statistics (OES) data. Recent changes in what data the government collects and publishes obviate these problems and make such comparisons both more accurate and more relevant.

Specifically, in the old system of accounting job descriptions and wage claims were both arbitrary and unrelated to official OES statistics. In the new system, virtually all Labor Condition Applications (LCA's) reference specific Occupational Employment Statistics (OES) data in their declarations of LCA wages and prevailing wages. This is because the new iCert system for processing LCAs has a feature that integrates an OES database search for prevailing wages into the LCA creation process. At the same time, the new iCert system requires that LCAs be referenced to the Standard Occupational Classification (SOC) codes, which eliminates the previous confusion about which job titles should be associated with which salary information.

7. *Methodology*

Special Note: A companion website to this paper is available at http://roberthill.org/h1b. This website provides screen capture videos explaining and displaying the methods used to retrieve the data and reveal patterns within it.

Acquisition of the Data

The source of the data is the Foreign Labor Certification Data Center website at http://www.flcdatacenter.com/CaseH1B.aspx. This website is the public distribution site for the current and archived LCAs going back to 2001. This study uses the data for FY2010 as published in Microsoft Access ".mdb" format.

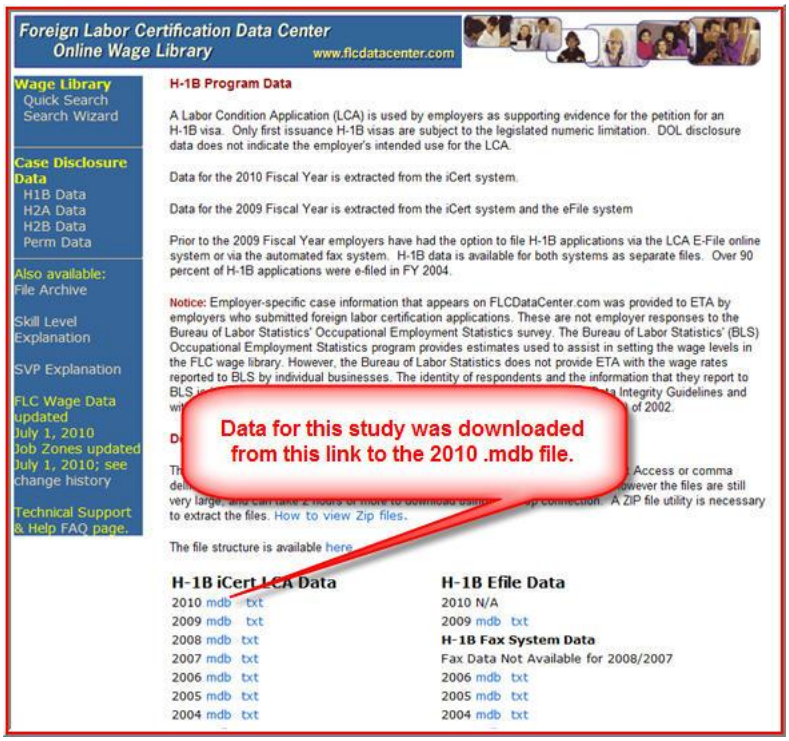
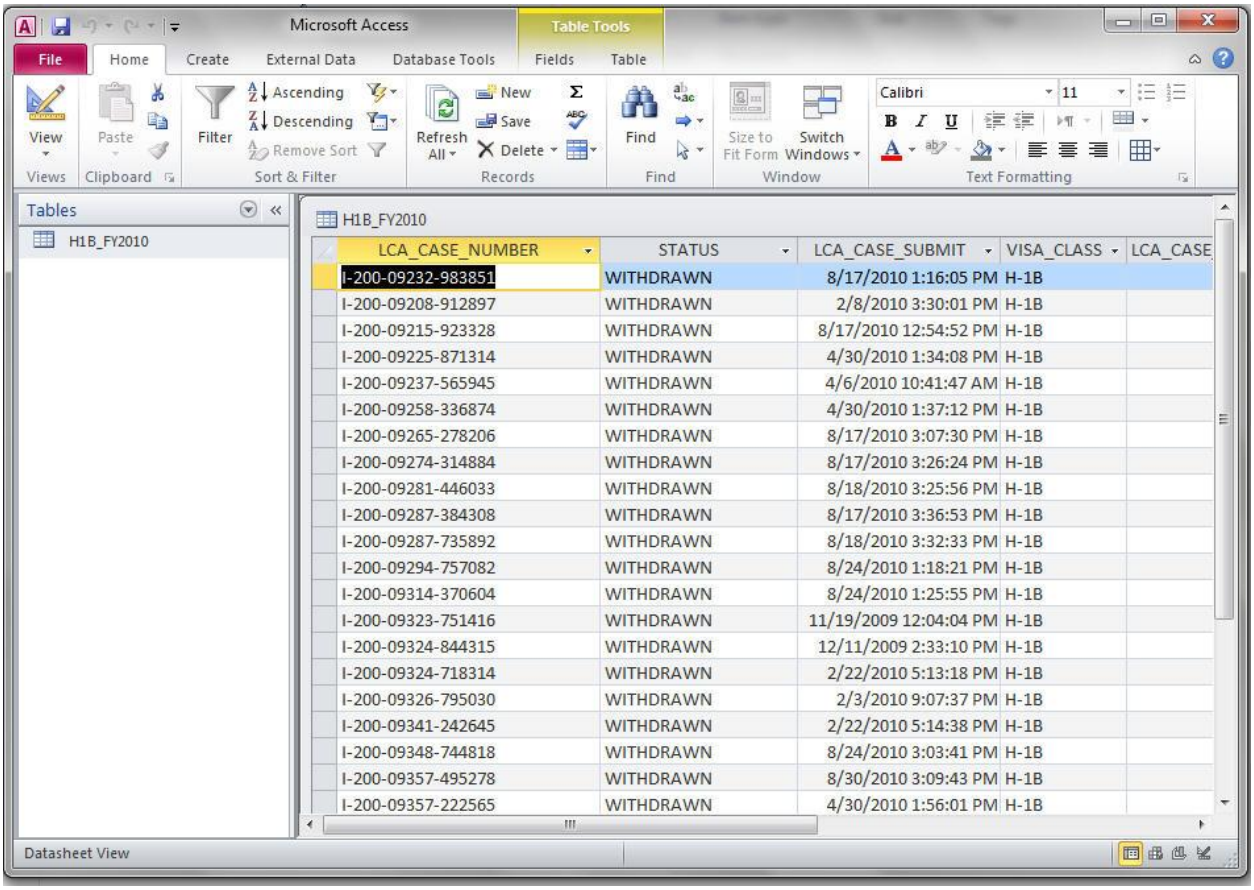


Figure 6: FLCDC Online Wage Library iCert data link.

When opened in Microsoft Access, the downloaded file proves to be a record of 335,328 LCAs stored as a single table named "H1B_FY2010."



LCA_CASE_NUMBER	STATUS	LCA_CASE_SUBMIT	VISA_CLASS	LCA_CASE
I-200-09232-983851	WITHDRAWN	8/17/2010 1:16:05 PM	H-1B	
I-200-09208-912897	WITHDRAWN	2/8/2010 3:30:01 PM	H-1B	
I-200-09215-923328	WITHDRAWN	8/17/2010 12:54:52 PM	H-1B	
I-200-09225-871314	WITHDRAWN	4/30/2010 1:34:08 PM	H-1B	
I-200-09237-565945	WITHDRAWN	4/6/2010 10:41:47 AM	H-1B	
I-200-09258-336874	WITHDRAWN	4/30/2010 1:37:12 PM	H-1B	
I-200-09265-278206	WITHDRAWN	8/17/2010 3:07:30 PM	H-1B	
I-200-09274-314884	WITHDRAWN	8/17/2010 3:26:24 PM	H-1B	
I-200-09281-446033	WITHDRAWN	8/18/2010 3:25:56 PM	H-1B	
I-200-09287-384308	WITHDRAWN	8/17/2010 3:36:53 PM	H-1B	
I-200-09287-735892	WITHDRAWN	8/18/2010 3:32:33 PM	H-1B	
I-200-09294-757082	WITHDRAWN	8/24/2010 1:18:21 PM	H-1B	
I-200-09314-370604	WITHDRAWN	8/24/2010 1:25:55 PM	H-1B	
I-200-09323-751416	WITHDRAWN	11/19/2009 12:04:04 PM	H-1B	
I-200-09324-844315	WITHDRAWN	12/11/2009 2:33:10 PM	H-1B	
I-200-09324-718314	WITHDRAWN	2/22/2010 5:13:18 PM	H-1B	
I-200-09326-795030	WITHDRAWN	2/3/2010 9:07:37 PM	H-1B	
I-200-09341-242645	WITHDRAWN	2/22/2010 5:14:38 PM	H-1B	
I-200-09348-744818	WITHDRAWN	8/24/2010 3:03:41 PM	H-1B	
I-200-09357-495278	WITHDRAWN	8/30/2010 3:09:43 PM	H-1B	
I-200-09357-222565	WITHDRAWN	4/30/2010 1:56:01 PM	H-1B	

Figure 7: Example of the full data table for the FY 2010 H-1B data.

This table of records represents all LCAs from all categories of job for fiscal year 2010, and it also includes those that were denied or withdrawn. It also includes fields that are not of interest. To sort through the data and produce sub-sets of data that could be used to replicate Miano's 2007 research, a set of queries was written to extract the data and export it to Microsoft Excel format.

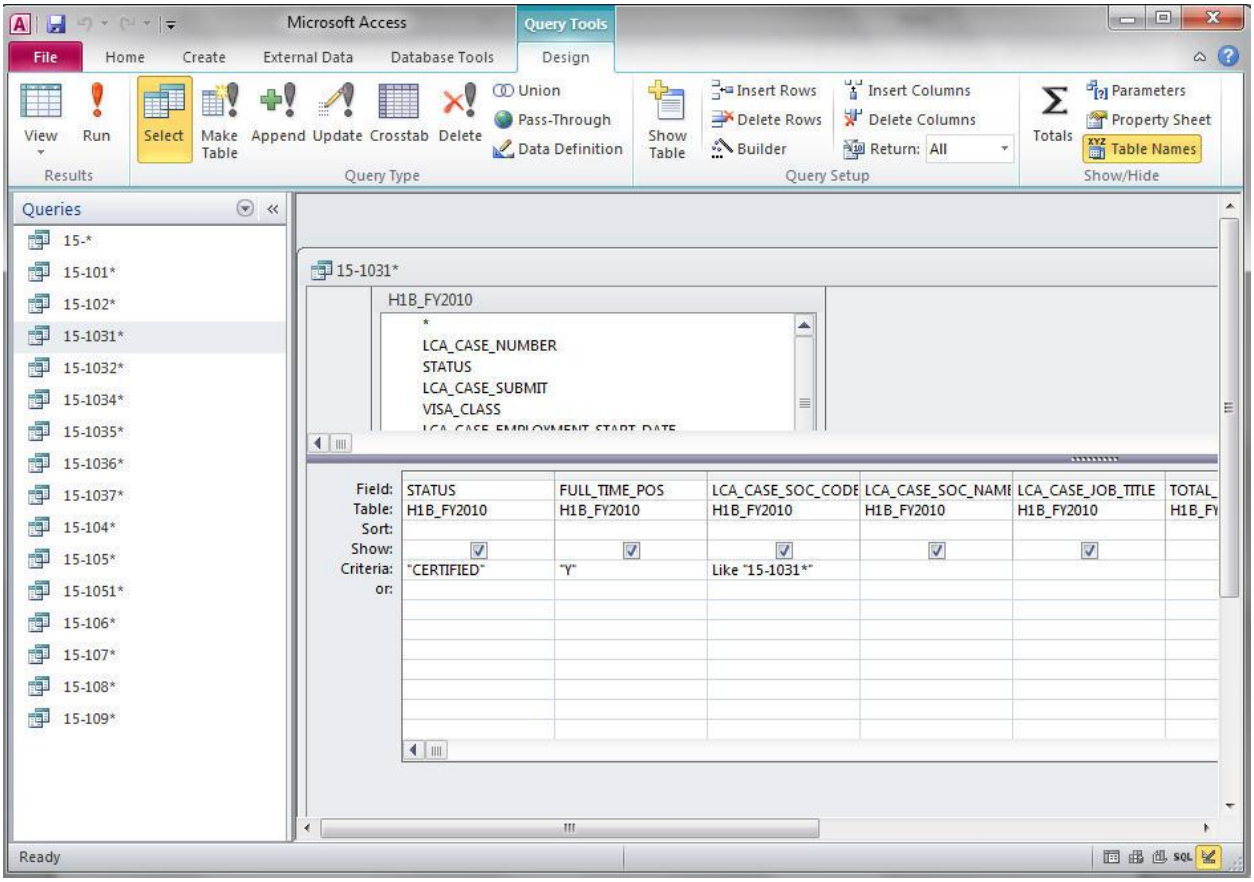


Figure 8: Example of an Access query to export data to Excel.

Miano's 2007 study was drawn from LCA data that used the older 3-digit LCA occupational codes. For his purposes, he focused his attention on the "Computer Related Occupation" codes, as follows:

- **030:** Occupations in Systems Analysis and Programming
- **031:** Occupations in Data Communications and Networks
- **032:** Occupations in Computer System User Support
- **033:** Occupations in Computer System Technical Support
- **039:** Other Computer-Related Occupations

In the corresponding SOC code system, the relevant codes are those in the "15-xxxx.xx" series, as shown in the table below. For the purposes of this study, the ten SOC codes listed below are used for comparison because these are the jobs codes that directly tie the LCAs to the OES data.

- **15-1011:** Computer and Information Scientists, Research
- **15-1021:** Computer Programmers
- **15-1031:** Computer Software Engineers, Applications
- **15-1032:** Computer Software Engineers, Systems Software
- **15-1041:** Computer Support Specialists
- **15-1051:** Computer Systems Analysts
- **15-1061:** Database Administrators
- **15-1071:** Network and Computer Systems Administrators
- **15-1081:** Network Systems and Data Communications Analysts
- **15-1099:** Computer Specialists, All Other

Further, these are the SOC codes for which the Bureau of Labor Statistics maintains wage estimate data, as available on their website.

Computer and Mathematical Science Occupations top						
			Wage Estimates			
Occupation Code	Occupation Title (click on the occupation title to view an occupational profile)	Employment (1)	Median Hourly	Mean Hourly	Mean Annual (2)	Mean RSE (3)
15-0000	**Computer and Mathematical Science Occupations**	3,303,690	$35.05	$36.68	$76,290	0.3 %
15-1011	Computer and Information Scientists, Research	26,130	$48.83	$50.66	$105,370	1.1 %
15-1021	Computer Programmers	367,880	$34.10	$35.91	$74,690	0.7 %
15-1031	Computer Software Engineers, Applications	495,500	$42.06	$43.35	$90,170	0.4 %
15-1032	Computer Software Engineers, Systems Software	385,200	$44.94	$46.45	$96,620	0.6 %
15-1041	Computer Support Specialists	540,560	$21.30	$22.77	$47,360	0.3 %
15-1051	Computer Systems Analysts	512,720	$37.06	$38.67	$80,430	0.4 %
15-1061	Database Administrators	108,080	$34.40	$35.72	$74,290	0.7 %
15-1071	Network and Computer Systems Administrators	338,890	$32.55	$34.10	$70,930	0.3 %
15-1081	Network Systems and Data Communications Analysts	226,080	$35.22	$36.81	$76,560	0.4 %
15-1099	Computer Specialists, All Other	195,890	$37.02	$37.50	$78,010	0.4 %
15-2011	Actuaries	17,940	$41.93	$46.85	$97,450	1.0 %
15-2021	Mathematicians	2,770	$44.99	$45.16	$93,920	1.8 %
15-2031	Operations Research Analysts	60,960	$33.69	$36.23	$75,370	0.8 %
15-2041	Statisticians	21,370	$35.01	$36.16	$75,220	1.9 %
15-2091	Mathematical Technicians	1,090	$19.83	$21.27	$44,230	2.3 %
15-2099	Mathematical Science Occupations, All Other	2,610	$23.18	$29.74	$61,850	7.3 %

Figure 9: Listing of SOC codes on the BLS site, with links to data, published May 2009.

There were a few hundred instances where the queries returned SOC codes that fell outside of those listed above. These outlier SOC codes were:

- **15-1022:** Computer Programmers
- **15-1023:** Computer Programmers
- **15-1034:** Computer Software Engineers, Applications, Non R&D
- **15-1035:** Computer Software Engineers, Applications, R&D
- **15-1036:** Computer Software Engineers, Systems Software, Non RD&
- **15-1037:** Computer Software Engineers, Systems Software, R&D

As there were only a few hundred LCAs with these SOC codes, and as there is no OES data available for them, these relatively few records were left outside of the calculations.

Each of the 10 queries from the database file was exported to one of 10 Excel files for further processing.

Details of how the data was acquired and extracted, including the formal Structured Query Language (SQL) statement and definitions of the fields extracted, can be found in ***Appendix A: Data Acquisition Details***.

Transformation and Grooming of the Data

Once exported from Access to Excel format, each SOC code's returned data was groomed with a number of formatting treatments to make it usable for calculations and to cast it in a standardized format.

One problem that had to be overcome was transforming the character type of the wage data. When data is stored in a database file, certain decisions are made about its "metadata" characteristics. For instance, metadata about whether the data is text, or an integer, or a floating point number is stored along with the data itself. It was discovered that the numbers representing the wage data, both LCA and prevailing, were stored in the database file as "text." To the human eye, the numbers looked like numbers, but to the computer, the numbers looked like text and were unavailable for even simple calculations such as addition and multiplication. Because there were hundreds of thousands of data points with this fundamental flaw, a method was required to convert them from "stored-as-text" values to "stored-as-integer"

values. Once this was accomplished, mathematical calculations could be performed on the wage data.

The original numbers representing the wage data were also stored in a variety of decimal formats. A standardized number format was adopted to transform numbers into integers, with no decimal places, and using a comma to separate groups of thousands. E.g., "75600.00" was transformed to "75,600" without the use of a dollar sign. This transformation was done for all wage columns in all ten Excel files.

Details of how the data was transformed and groomed can be found in ***Appendix B: Transformation and Grooming of the Data, Details***.

Removing Outliers By Calculating and Sorting By Parameters

Once calculations were possible, a set of functions was applied to each of the column of wage data so that outliers could be found and removed from the calculations. The following seven functions were calculated.

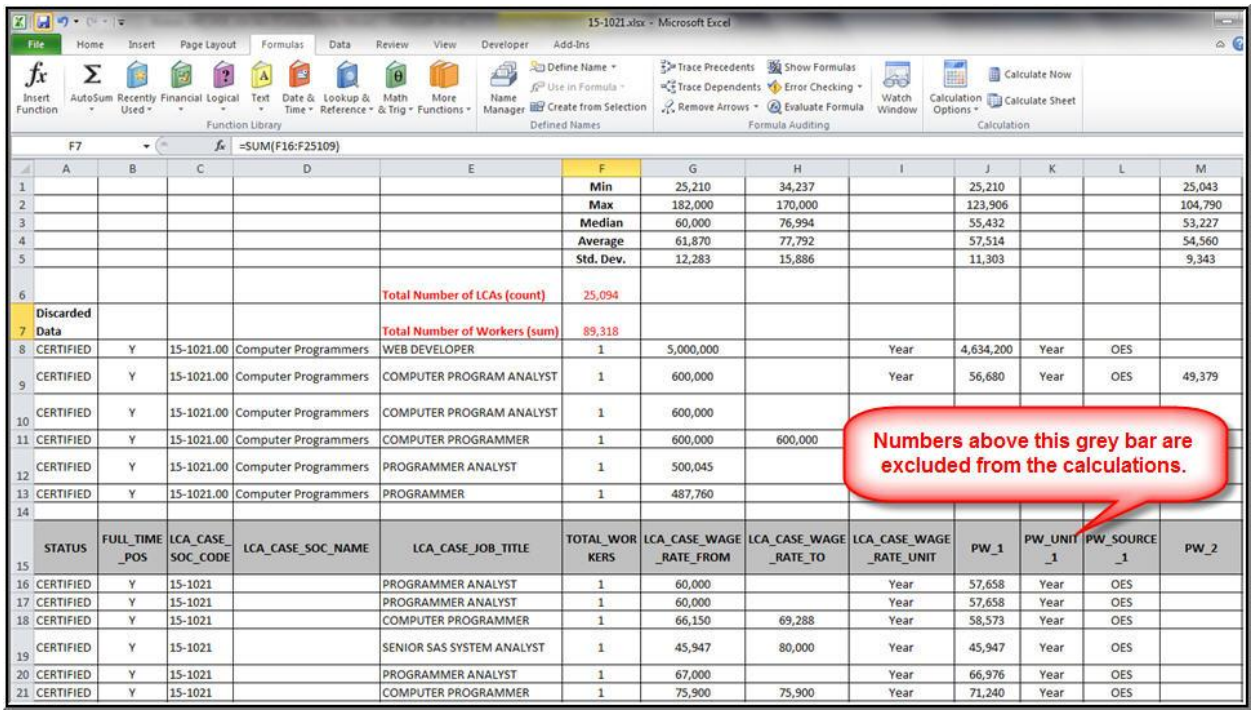
- **Min.** This is the smallest number in the column, used to check for and eliminate negative numbers from the calculations.
- **Max.** This is the largest number in the column, used to check for and eliminate any value over $300,000 as an outlier.
- **Median.** This is the number that represents the midpoint of the set of wage values. It is an actual value in the column. 50% of the values in the column should be larger than the median, and 50% smaller.
- **Mean.** This is the number that represents the average of the set of wage values. It is a calculated value, not necessarily found in the set. If the median and the mean are close to each other in value, then the likelihood that there are significant outliers is low.
- **Standard Deviation.** This is a statistical function that is used to assess how "normal" the distribution is. This value is supplied in the event that others find it useful.

- **Total Number of LCAs (count).** This number was calculated to provide the number of LCAs included in each set of calculations.
- **Total Number of Workers (sum).** This number was calculated to provide the number of workers included in each set of calculations.

These functions were applied to the following four columns of data.

- **LCA_CASE_WAGE_RATE_FROM.** The primary data point used in this calculation. The vast majority of LCAs were for a single worker at a single wage. This value is what the employer proposed to pay the H-1B worker.
- **LCA_CASE_WAGE_RATE_TO.** In some cases, LCAs showed wages as a range, in which case there was a upper limit provided. This data point was not used for any main calculations, but was used to eliminate the row of data from the main calculations in the "From Only" condition.
- **PW_1.** The claimed prevailing wage for the position being filled. This value was used as the primary data point for prevailing wage calculations.
- **PW_2.** In a few cases, LCAs showed prevailing wages as a range. This data point was not used for any main calculations.

Figure 10, below demonstrates a typical screen shot of spreadsheet with its sorting functions.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1						Min	25,210	34,237		25,210			25,043
2						Max	182,000	170,000		123,906			104,790
3						Median	60,000	76,994		55,432			53,227
4						Average	61,870	77,792		57,514			54,560
5						Std. Dev.	12,283	15,886		11,303			9,343
6					Total Number of LCAs (count)	25,094							
7	Discarded Data				Total Number of Workers (sum)	89,318							
8	CERTIFIED	Y	15-1021.00	Computer Programmers	WEB DEVELOPER	1	5,000,000		Year	4,634,200	Year	OES	
9	CERTIFIED	Y	15-1021.00	Computer Programmers	COMPUTER PROGRAM ANALYST	1	600,000		Year	56,680	Year	OES	49,379
10	CERTIFIED	Y	15-1021.00	Computer Programmers	COMPUTER PROGRAM ANALYST	1	600,000						
11	CERTIFIED	Y	15-1021.00	Computer Programmers	COMPUTER PROGRAMMER	1	600,000	600,000					
12	CERTIFIED	Y	15-1021.00	Computer Programmers	PROGRAMMER ANALYST	1	500,045						
13	CERTIFIED	Y	15-1021.00	Computer Programmers	PROGRAMMER	1	487,760						
14													
15	STATUS	FULL_TIME_POS	LCA_CASE_SOC_CODE	LCA_CASE_SOC_NAME	LCA_CASE_JOB_TITLE	TOTAL_WORKERS	LCA_CASE_WAGE_RATE_FROM	LCA_CASE_WAGE_RATE_TO	LCA_CASE_WAGE_RATE_UNIT	PW_1	PW_UNIT_1	PW_SOURCE_1	PW_2
16	CERTIFIED	Y	15-1021		PROGRAMMER ANALYST	1	60,000		Year	57,658	Year	OES	
17	CERTIFIED	Y	15-1021		PROGRAMMER ANALYST	1	60,000		Year	57,658	Year	OES	
18	CERTIFIED	Y	15-1021		COMPUTER PROGRAMMER	1	66,150	69,288	Year	58,573	Year	OES	
19	CERTIFIED	Y	15-1021		SENIOR SAS SYSTEM ANALYST	1	45,947	80,000	Year	45,947	Year	OES	
20	CERTIFIED	Y	15-1021		PROGRAMMER ANALYST	1	67,000		Year	66,976	Year	OES	
21	CERTIFIED	Y	15-1021		COMPUTER PROGRAMMER	1	75,900	75,900	Year	71,240	Year	OES	

Figure 10: Example of wage values excluded from calculations.

Line 8 from Figure 10 demonstrates a LCA claim that was removed from all calculations on the basis of it being an outlier. This LCA for a "Web Developer" claimed that the prevailing wage for the job was $4,634,200 but that they intended to pay the H-1B worker a salary of $5,000,000. While this is possible, it is highly unlikely to be true, to say the least. It is much more likely that this simply represents a clerical error; the correct values should probably be $46,342 and $50,000, respectively. And, since this outlier was two orders of magnitude larger than expected, i.e. 100 times larger, leaving it in the calculations would have significantly skewed the averages.

Rather than applying any corrections to the suspect data, they were simply removed from the calculations. Maximum LCA_WAGE_FROM and PW_1 values that were above $300,000

were sought out and moved to be above the grey bar representing the column headers. Then, calculations were automatically reperformed on all data below the grey header bar. This process was repeated until all Maximum values were brought below $300,000. The data was left in the spreadsheets for review, but removed from the calculations.

Dividing the Data Into "From" and "From Only" Branches

One factor that potentially affected results was whether the LCA provided only a LCA_Wage_From value or also a LCA_Wage_To value. The vast majority of LCAs were for one worker each, and for this worker the employers provided a single wage value. That wage value was recorded in the LCA_Wage_From field, as illustrated in Table 1 below.

Table 1: LCA Wage Examples, From Only vs. From and To

Job Title	Number of Workers	LCA Wage From	LCA Wage To
Java Developer, Entry Level	1	55,000	
Oracle DBMS Developer, Senior	2	96,000	
Systems Architect, Senior	4	70,000	100,000
Customer Support, Entry Level	10	45,000	65,000

Many thousands of the LCAs also included a value in the LCA_WAGE_TO column. This value, if present, indicated an upper limit to the wages that the employers intended to pay the H-1B workers. In the examples above, the wage information for the Java Developer can be reasonably assumed to be $55,000, and the wage information for the two Oracle developers could reasonably be assumed to be $96,000. However, how are we to calculate the salary information for any one of the four Systems Architects or 10 Customer Support workers?

One option considered was to average the LCA_WAGE_FROM and LCA_WAGE_TO valued for use in the calculations, but this was rejected because it represented changing the raw data. Ultimately, the solution chosen was to run the calculations on **two** branches of the data, one "From" and the other "From Only."

- **"From" condition:.** All validated LCAs were included in the calculations.
- **"From " ONLY condition:** Only those validated LCAs that had ONLY a LCA_WAGE_FROM value were included in the calculations. All LCAs that also included a LCA_WAGE_TO value were simply ignored.

The sample size of the "From" ONLY condition is smaller than that for the "From" condition, but it should still be sizeable enough to be valid. At any rate, comparisons can be made after-the-fact to determine whether there is a substantial effect in differentiating the two "From" conditions.

Dividing the Data Into "Non-Weighted" and "Weighted" Branches

Prior to the data analysis, it was not known whether "weighting" the data would reveal significant differences in the patterns revealed, so a decision was made to create both weighted and non-weighted conditions and compare them after analysis.

The term "weighting" refers to converting single LCAs that represent multiple job positions into multiple LCAs that each represent a single job. For instance, consider the following LCA for five entry-level Java developers in Table 2 below. In the non-weighted condition, the salary information for this LCA is counted in the calculations ONCE, even though there are potentially (but not definitely) FIVE hires available.

Table 2: Non-weighted Condition			
Job Title	**Number of Workers**	**LCA Wage**	**Prevailing Wage**
Java Developer, Entry Level	5	55,000	50,000

In the weighted condition, this LCA for five entry-level Java developers was converted so that the wage information was counted in the calculations FIVE TIMES, as in Table 3 below.

Table 3: Weighted Condition			
Job Title	**Number of Workers**	**LCA Wage**	**Prevailing Wage**
Java Developer, Entry Level	1	55,000	50,000
Java Developer, Entry Level	1	55,000	50,000
Java Developer, Entry Level	1	55,000	50,000
Java Developer, Entry Level	1	55,000	50,000
Java Developer, Entry Level	1	55,000	50,000

Essentially, conducting analyses on both non-weighted and weighted conditions allowed for accounting for any POSSIBLE variation from the difference between the two conditions.

Four-Way Matrix Of Data Conditions

Ultimately, there were to be four variations of the data. The same calculations were run on each of the four variations. The four dataset variations are listed and explained as follows:

"From," non-weighted.	**"From" ONLY, non-weighted.**
"From," weighted.	**"From" ONLY, weighted.**

- **“From,” non-weighted:**
    - In this condition, the individual LCAs are all treated as equals, even though some are for more than one worker.
    - An LCA for one 15-1031 worker at a salary of $50,000 is given the same significance as one that is the same in every respect EXCEPT that it is for 120 workers.
    - The value of $50,000 is figured into the calculations only once.
    - Even though some LCAs had only a “From” value while others had both a “From” and a “To” value, all LCAs were included in the calculations. The “To” values were ignored.

- **“From,” weighted:**
    - In this condition, each LCA is each copied by the number of workers called for in the LCA.
    - An LCA for 120 15-1031 workers at a salary of $50,000 is copied into 120 LCAs, (and 8 workers copied 8 times, etc.)
    - The value of $50,000 is figured into the calculations for as many workers as called for.
    - Even though some LCAs had only a “From” value while others had both a “From” and a “To” value, all LCAs were included in the calculations. The “To” values were ignored.

- **“From” ONLY, non-weighted:**
    - In this condition, the individual LCAs are all treated as equals, even though some are for more than one worker.
    - An LCA for one 15-1031 worker at a salary of $50,000 is given the same significance as one that is the same in every respect EXCEPT that it is for 120 workers.
    - The value of $50,000 is figured into the calculations only once.
    - ONLY those LCAs that list ONLY an LCA “From” wage (and leave the LCA To field blank) are factored into the calculations. All LCAs that list a LCA “To” wage are discarded.

- **“From” ONLY, weighted:**
    - In this condition, each LCA is each copied by the number of workers called for in the LCA.
    - An LCA for 120 15-1031 workers at a salary of $50,000 is copied into 120 LCAs, (and 8 workers copied 8 times, etc.)
    - The value of $50,000 is figured into the calculations for as many workers as called for.
    - ONLY those LCAs that list ONLY an LCA “From” wage (and leave the LCA To field blank) are factored into the calculations. All LCAs that list a LCA “To” wage are discarded.

8. Results

Using these sort and filter criteria, this returns 89,585 LCAs covering 232,487 potential H-1B workers for the computer technology SOC codes for fiscal year 2010. These numbers are smaller than Miano's report numbers (300,000 LCAs covering 700,000 workers), but they are perhaps more accurate because there was no need to guess at which SOC codes to compare the LCA jobs to, and there was no need to use wildcard filters to extract relevant job titles. Table 4 below displays the number of LCAs and Worker Requests (the "weighted" LCAs) across the various SOC codes.

Table 4: Number of LCAs for each condition, sorted by SOC codes.

SOC Code	"From," non-weighted	"From," weighted	"From" ONLY, non-weighted	"From" ONLY, weighted
15-1011: Computer and Information Scientists, Research	675	717	488	492
15-1021: Computer Programmers	25,094	89,318	17,989	28,731
15-1031: Computer Software Engineers, Applications	20,647	34,229	15,308	23,007
15-1032: Computer Software Engineers, Systems Software	10,468	14,127	8,220	11,003
15-1041: Computer Support Specialists	533	786	399	413
15-1051: Computer Systems Analysts	26,273	84,362	20,463	51,512
15-1061: Database Administrators	2,513	4,082	1,799	2,234
15-1071: Network and Computer Systems Administrators	2,716	3,943	1,886	1,990
15-1081: Network Systems and Data Communications Analysts	666	923	510	529
15-1099: Computer Specialists, All Other	5,345	7,562	4,240	3,713
Totals	89,585	232,487	67,062	119,911

Across the 10 computer technology worker SOC codes studied, 67,062 LCAs (representing 119,911 potential jobs) reported only LCA_WAGE_FROM data. That is, approximately 74.8% of the total number of LCAs did not have the problem of having to determine whether the difference between LCA_WAGE_FROM and LCA_WAGE_TO was significant.

Figure 11 below shows the distribution of LCAs and the potential number of workers requested across the 10 SOC codes.

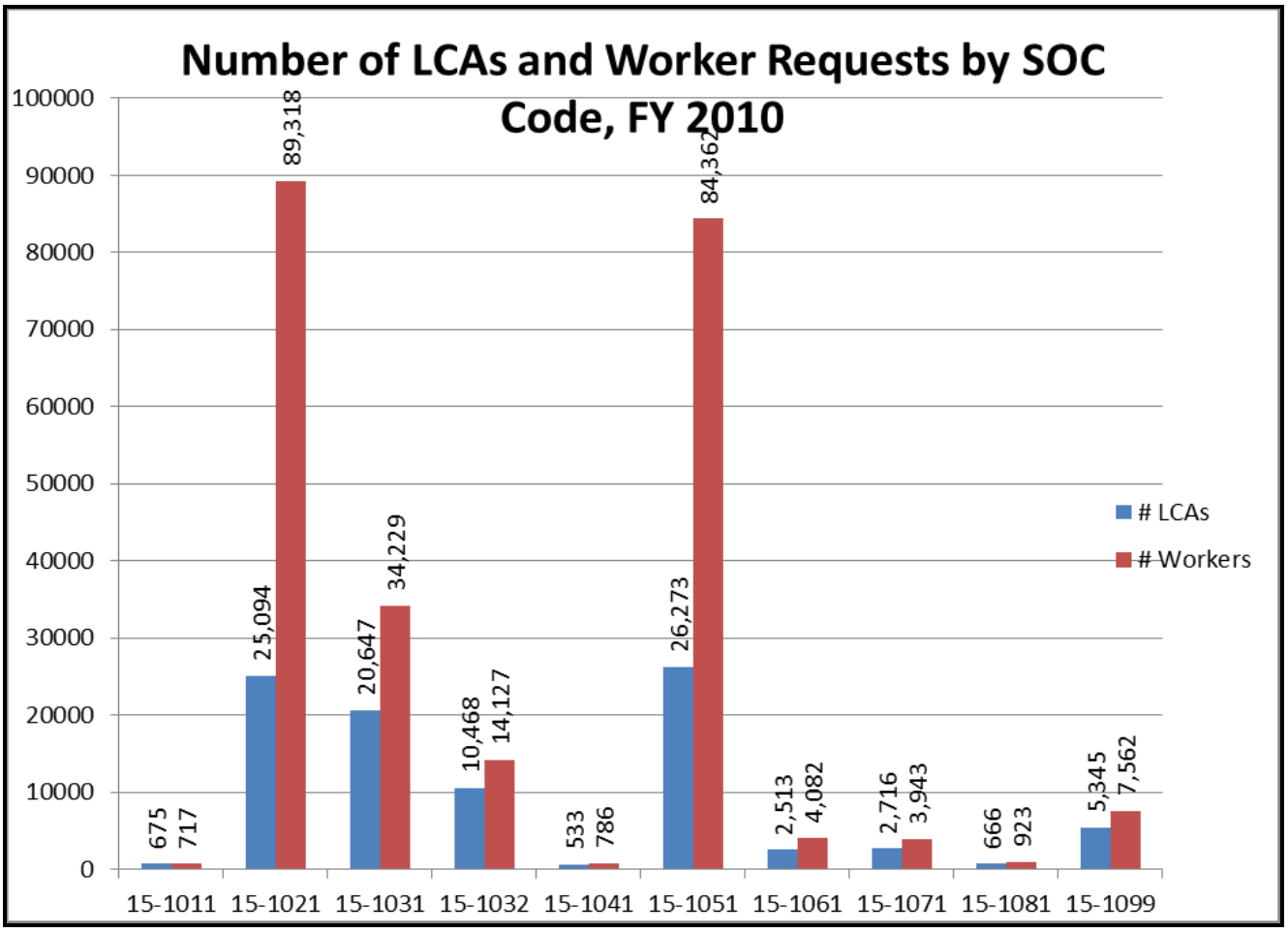
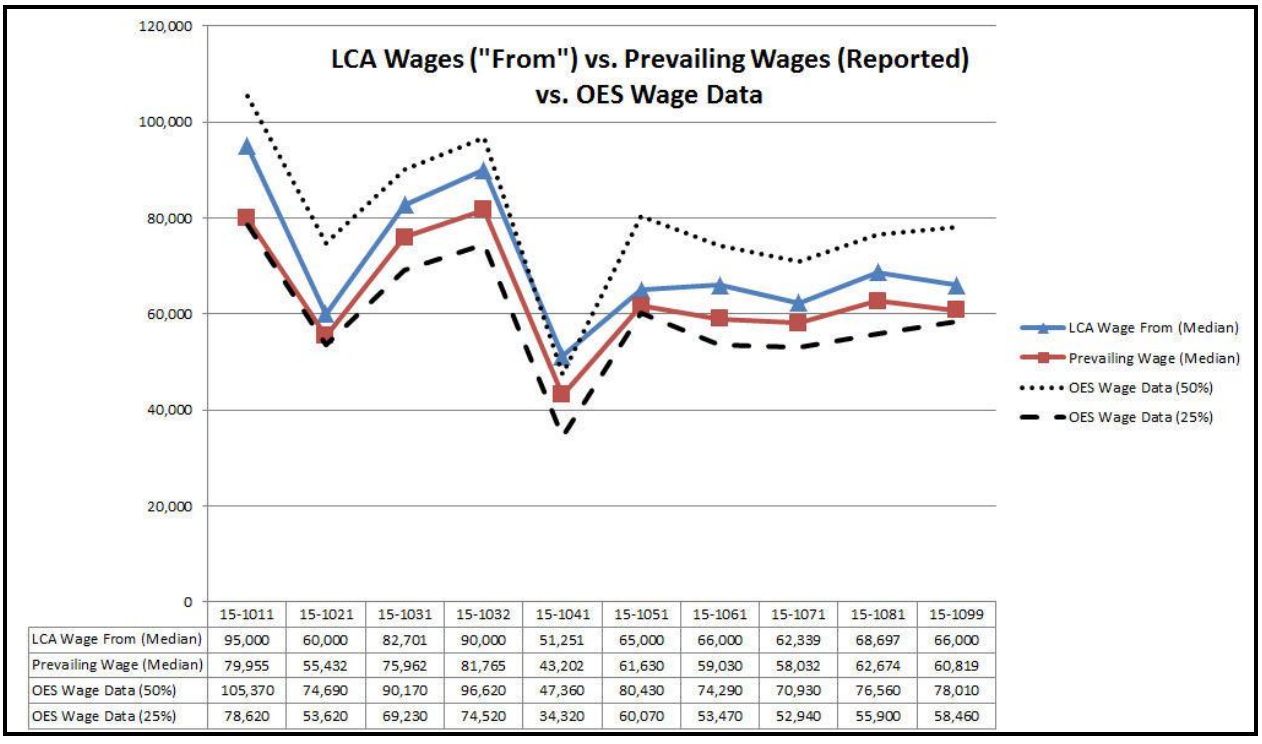


Figure 11: Number of LCAs and Worker Requests by SOC Code, FY 2010

The distribution shows that SOC codes 15-1051 and 15-1021 were most represented in the total distribution of both SLAs and number of workers requested, with codes 15-1031 and 15-1032 trailing closely behind them. These four SOC code groups combined represented 82,482 of the total of 89,585 LCAs studied, or approximately 92%. Similar trends are observed for the number of workers requested. The same four SOC codes represented 222,031 requested workers out of the total of 232,487, or approximately 95.5%.

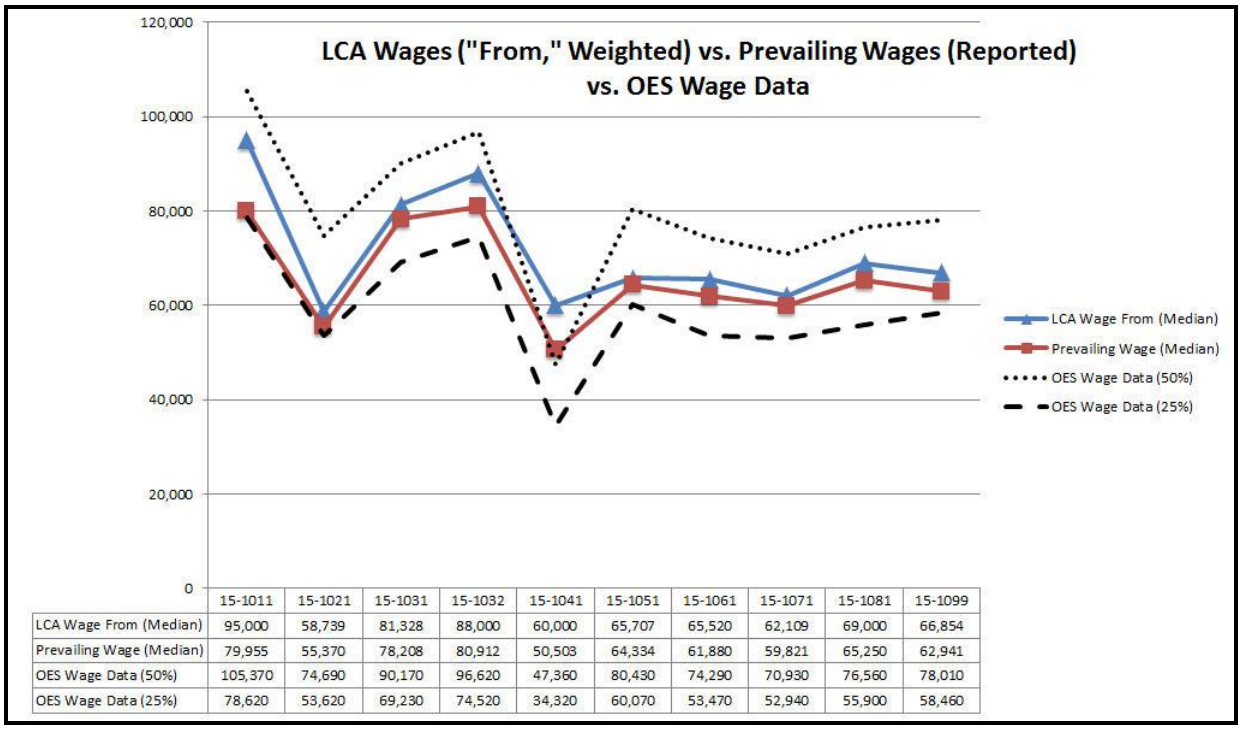
Figures 12, 13, 14, and 15 below represent the same graphic distribution of data across the four identified variations of the data:

Figure 12: "From," non-weighted.	**Figure 14:** "From" ONLY, non-weighted.
Figure 13: "From," weighted.	**Figure 15:** "From" ONLY, weighted.



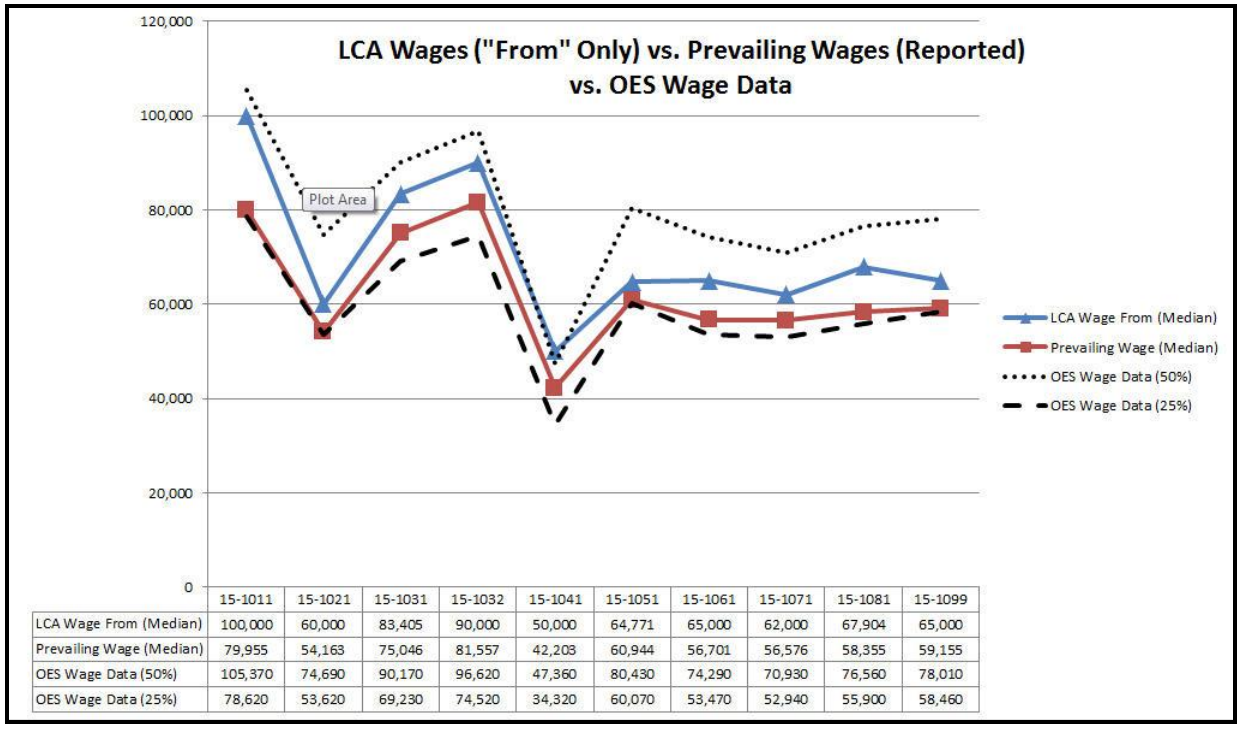
	15-1011	15-1021	15-1031	15-1032	15-1041	15-1051	15-1061	15-1071	15-1081	15-1099
LCA Wage From (Median)	95,000	60,000	82,701	90,000	51,251	65,000	66,000	62,339	68,697	66,000
Prevailing Wage (Median)	79,955	55,432	75,962	81,765	43,202	61,630	59,030	58,032	62,674	60,819
OES Wage Data (50%)	105,370	74,690	90,170	96,620	47,360	80,430	74,290	70,930	76,560	78,010
OES Wage Data (25%)	78,620	53,620	69,230	74,520	34,320	60,070	53,470	52,940	55,900	58,460

Figure 12: LCA "From" wages vs. Prevailing Wages (Reported) vs. OES Wage Data



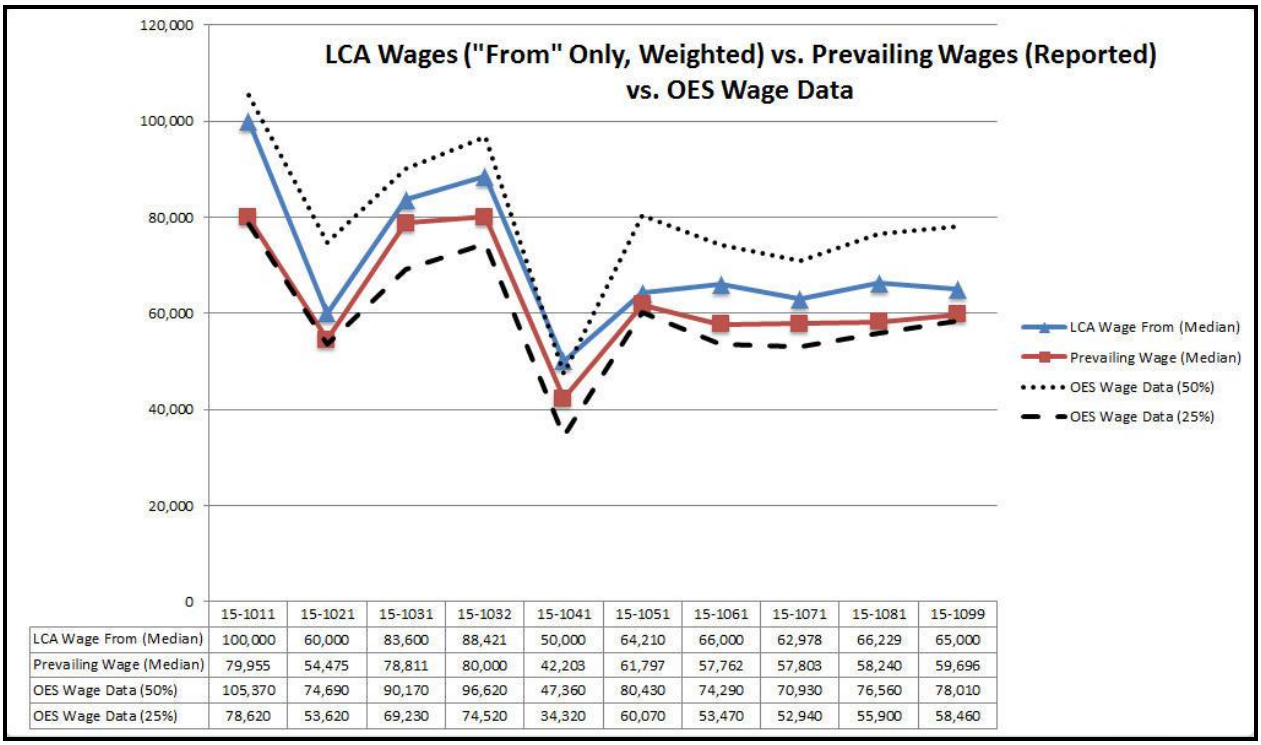
	15-1011	15-1021	15-1031	15-1032	15-1041	15-1051	15-1061	15-1071	15-1081	15-1099
LCA Wage From (Median)	95,000	58,739	81,328	88,000	60,000	65,707	65,520	62,109	69,000	66,854
Prevailing Wage (Median)	79,955	55,370	78,208	80,912	50,503	64,334	61,880	59,821	65,250	62,941
OES Wage Data (50%)	105,370	74,690	90,170	96,620	47,360	80,430	74,290	70,930	76,560	78,010
OES Wage Data (25%)	78,620	53,620	69,230	74,520	34,320	60,070	53,470	52,940	55,900	58,460

Figure 13: LCA "From," Weighted wages vs. Prevailing Wages (Reported) vs. OES Wage Data



	15-1011	15-1021	15-1031	15-1032	15-1041	15-1051	15-1061	15-1071	15-1081	15-1099
LCA Wage From (Median)	100,000	60,000	83,405	90,000	50,000	64,771	65,000	62,000	67,904	65,000
Prevailing Wage (Median)	79,955	54,163	75,046	81,557	42,203	60,944	56,701	56,576	58,355	59,155
OES Wage Data (50%)	105,370	74,690	90,170	96,620	47,360	80,430	74,290	70,930	76,560	78,010
OES Wage Data (25%)	78,620	53,620	69,230	74,520	34,320	60,070	53,470	52,940	55,900	58,460

Figure 14: LCA "From" Only wages vs. Prevailing Wages (Reported) vs. OES Wage Data



	15-1011	15-1021	15-1031	15-1032	15-1041	15-1051	15-1061	15-1071	15-1081	15-1099
LCA Wage From (Median)	100,000	60,000	83,600	88,421	50,000	64,210	66,000	62,978	66,229	65,000
Prevailing Wage (Median)	79,955	54,475	78,811	80,000	42,203	61,797	57,762	57,803	58,240	59,696
OES Wage Data (50%)	105,370	74,690	90,170	96,620	47,360	80,430	74,290	70,930	76,560	78,010
OES Wage Data (25%)	78,620	53,620	69,230	74,520	34,320	60,070	53,470	52,940	55,900	58,460

Figure 15: LCA "From" Only, Weighted wages vs. Prevailing Wages (Reported) vs. OES Wage Data

As was stated earlier, it was not known at the time of outset whether the problems of "From" versus "From ONLY" and non-weighted versus weighted would significantly affect the outcomes of the analysis, so the analysis was done on four branches of the data. Therefore, results for Figures 12 through 15 can be analysed in the same way.

There are four horizontal lines drawn across each of the figures. Each of these four lines has 10 data points, which represent each of the 10 studied SOC codes. The four lines, top to bottom, are described as follows:

- **Top Dotted Line:** The OES Wage Data (50%) for each of the 10 SOC codes.
- **Triangle Line:** The LCA Wage Data (Median) for each of the 10 SOC codes.
- **Square Line:** The Prevailing Wage Data (Median) for each of the 10 SOC codes.
- **Bottom Dashed Line:** The OES Wage Data (25%) for each of the 10 SOC codes.

Specific wage data, expressed in terms of the median values, is included in the tables beneath each of the four figures. For instance, from Figure 12, the least-modified and most-inclusive of the data sets, SOC code 15-1011 shows that the OES median wage (50th percentile rank) for that occupation is $105,370. (Hereafter, this condition is referred to as the "primary data condition.") That is, half of domestic workers earn MORE than that value, and half earn less. The 25th percentile rank wage for the same occupation is $78,620. The median LCA Wage for SOC code 15-1011, represented by the top-left-most triangle, is $95,000, and the LCA Prevailing Wage is reported as $79,955.

Table 5 below displays the OES Wage Division Points for Percentile Rank Divisions for each of the 10 SOC codes. This data was copied directly from the 2009 OES Wage Library website, and it was used in the calculation of the numbers in Table 6.

Table 5: 2009 OES Wage Division Points for Percentile Rank Divisions, All 10 SOC Codes

	1st - 9th	10th - 24th	25th - 49th	50th - 74th	75th - 89th	90th - 100th
15-1011	< 61,480	78,620	105,370	127,600	155,420	> 155,420
15-1021	< 40,640	53,620	70,940	91,000	113,380	> 113,380
15-1031	< 54,840	69,230	87,480	107,710	132,080	> 132,080
15-1032	< 59,600	74,520	93,470	116,510	139,930	> 139,930
15-1041	< 27,200	34,320	44,300	57,290	72,690	> 72,690
15-1051	< 47,130	60,070	77,080	97,200	119,170	> 119,170
15-1061	< 40,780	53,470	71,550	93,260	114,200	> 114,200
15-1071	< 41,940	52,940	67,710	85,830	105,970	> 105,970
15-1081	< 42,880	55,900	73,250	94,320	116,120	> 116,120
15-1099	< 41,680	58,460	77,010	96,890	115,050	> 115,050

For 2009, the OES Wage Library reported that for SOC Code 15-1011, the Median Wage was $105,370. However, the OES data is further broken down into six divisions, three above the median and three below. These six divisions represent the 1^{st} – 9^{th}, 10^{th} – 24^{th}, 25^{th} – 49^{th}, 50^{th} – 74^{th}, 75^{th} – 89^{th}, and 90^{th} – 100^{th} percentile ranks. For instance, the bottom 10% of wages earned for SOC code 15-1011 was $61,480 or less. That is, anyone who earned more than or equal to $1 but less than $61,480 was counted in the 1^{st} through 9^{th} percentile ranks. A worker who earned more than or equal to $64,480 but less than $78,620 was counted in the 10^{th} through 24^{th} percentile ranks. At the top of the scale, a worker who earned anything more than or equal to $155,420 was counted in the 90^{th} to 100^{th} percentile ranks.

Thus, it is possible to compare the LCA Wages that employers claimed to be paying their H-1B workers against the actual OES Wage Library data for the same SOC codes. By defining the percentile rank division points, it is possible to count the numbers and percentages of workers whose wages fall within each of the six divisions. Table 6 below shows this distribution.

Important Note:

Though the data was broken into four branches to control for the possibility of significant influence from LCA Wage From and LCA Wage To and weighting variations, all results from this point forward are calculated using the primary data condition: "From," Non-Weighted.

Table 6: Distribution of LCA Wages Into Standard OES Percentile Divisions, All 10 SOC Codes							
	1st - 9th	10th - 24th	25th - 49th	50th - 74th	75th - 89th	90th - 100th	Re-Total
15-1011	99	80	203	215	65	13	675
	14.67%	11.85%	30.07%	31.85%	9.63%	1.93%	
15-1021	313	5,727	15,220	3,179	544	111	25,094
	1.25%	22.82%	60.65%	12.67%	2.17%	0.44%	
15-1031	498	4,496	7,082	5,761	2,378	432	20,647
	2.41%	21.78%	34.30%	27.90%	11.52%	2.09%	
15-1032	61	816	3,735	3,834	1,680	342	10,468
	0.58%	7.80%	35.68%	36.63%	16.05%	3.27%	
15-1041	2	32	140	157	115	87	533
	0.38%	6.00%	26.27%	29.46%	21.58%	16.32%	
15-1051	457	8,252	10,512	4,744	1753	555	26,273
	1.74%	31.41%	40.01%	18.06%	6.67%	2.11%	
15-1061	40	444	980	619	313	117	2,513
	1.59%	17.67%	39.00%	24.63%	12.46%	4.66%	
15-1071	30	536	1,095	602	319	134	2,716
	1.10%	19.73%	40.32%	22.16%	11.75%	4.93%	
15-1081	31	127	223	173	88	24	666
	4.65%	19.07%	33.48%	25.98%	13.21%	3.60%	
15-1099	98	1,296	2,208	1,228	385	130	5,345
	1.83%	24.25%	41.31%	22.97%	7.20%	2.43%	
Totals	**1,629**	**21,806**	**41,398**	**20,512**	**7,640**	**1,945**	**94,930**
		Below Median	**64,833**	**30,097**	**Above Median**		
		%	**68.30%**	**31.70%**	**%**		

For SOC code 15-1011 shown above, 99 out of the 675 workers, or 14.67%, had reported LCA Wages that were greater than or equal to $1 but less than $61,480 (as drawn from Table 5). Similarly, 80 workers, or 11.85%, had reported LCA Wages that were greater than or equal to $61,480 but less than $78,620. At the high end of the scale, 13 workers, or 1.93%, had LCA reported wages that placed them in the top-tier 90th – 100th percentile rank category.

Figure 16 below graphically demonstrates the number of H-1B workers whose LCA Wages fell into each of the six Standard OES Pecentile Divisions.

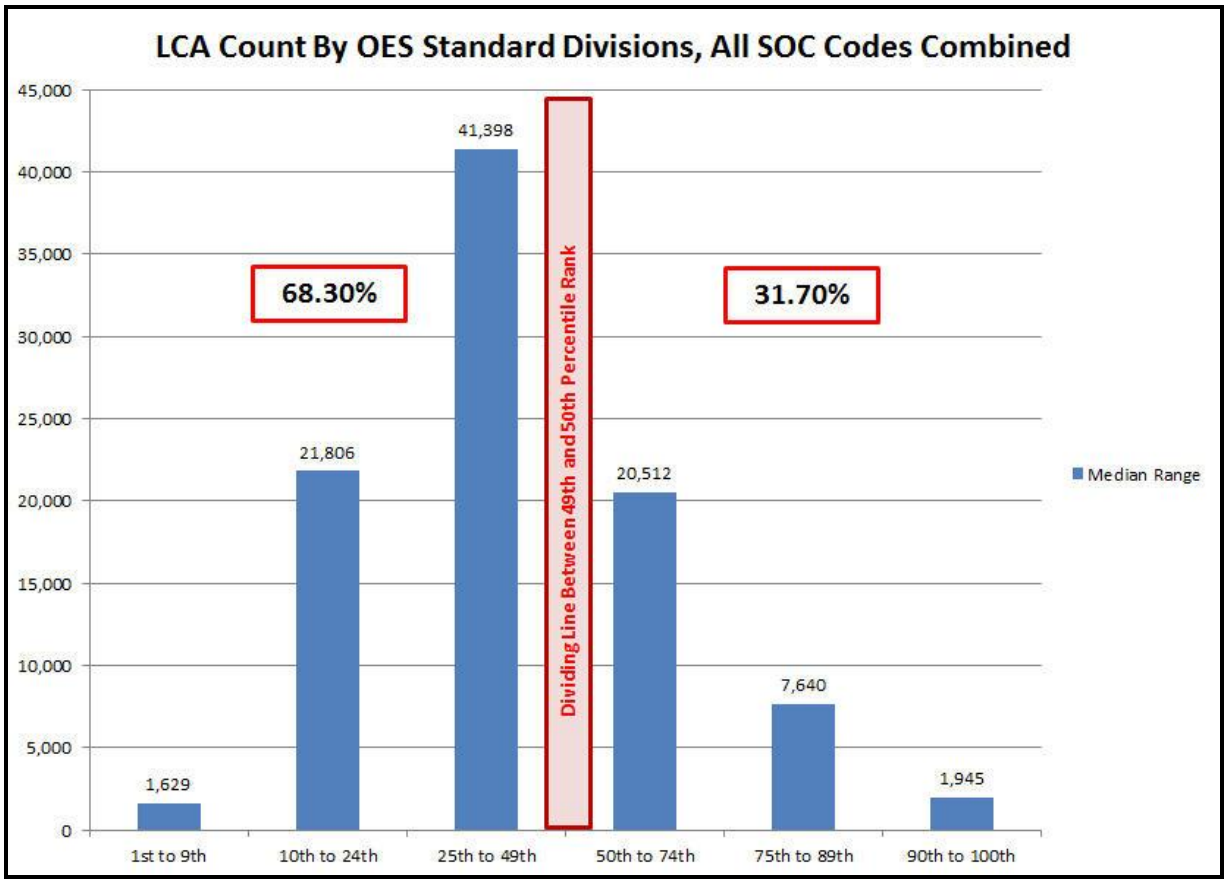


Figure 16: LCA Count By OES Standard Divisions, All 10 SOC Codes

The largest number of H-1B workers earned LCA wages that fell somewhere between the 25th and 49th percentile ranks, being 41,398 workers. Approximately the same numbers of H-1B workers had earnings that fell into the 10th – 24th and 50th – 74th percentile rank divisions, being 21,806 and 20,512, respectively.

Summing these results, it is possible to determine how many H-1B workers had LCA Wages BELOW the median point, and how many had LCA Wages ABO VE the median. Summing the

numbers 1,629 + 21,806 + 41,398 yields the total of 64,833, representing the number of H-1B workers who earned LCA wages in the three lower percentile divisions. This number represents approximately 68.30% of all H-1B workers earning LCA wages below the median for fiscal year 2009. Likewise, summing 20,512 + 7,640 + 1,945 yields the total of 30,097 H-1B workers (31.70%) earning LCA wages above the median for fiscal year 2009. Figure 17 below demonstrates these percentile rank divisions organized into below-the-median and above-the-median wages across all 10 SOC codes.

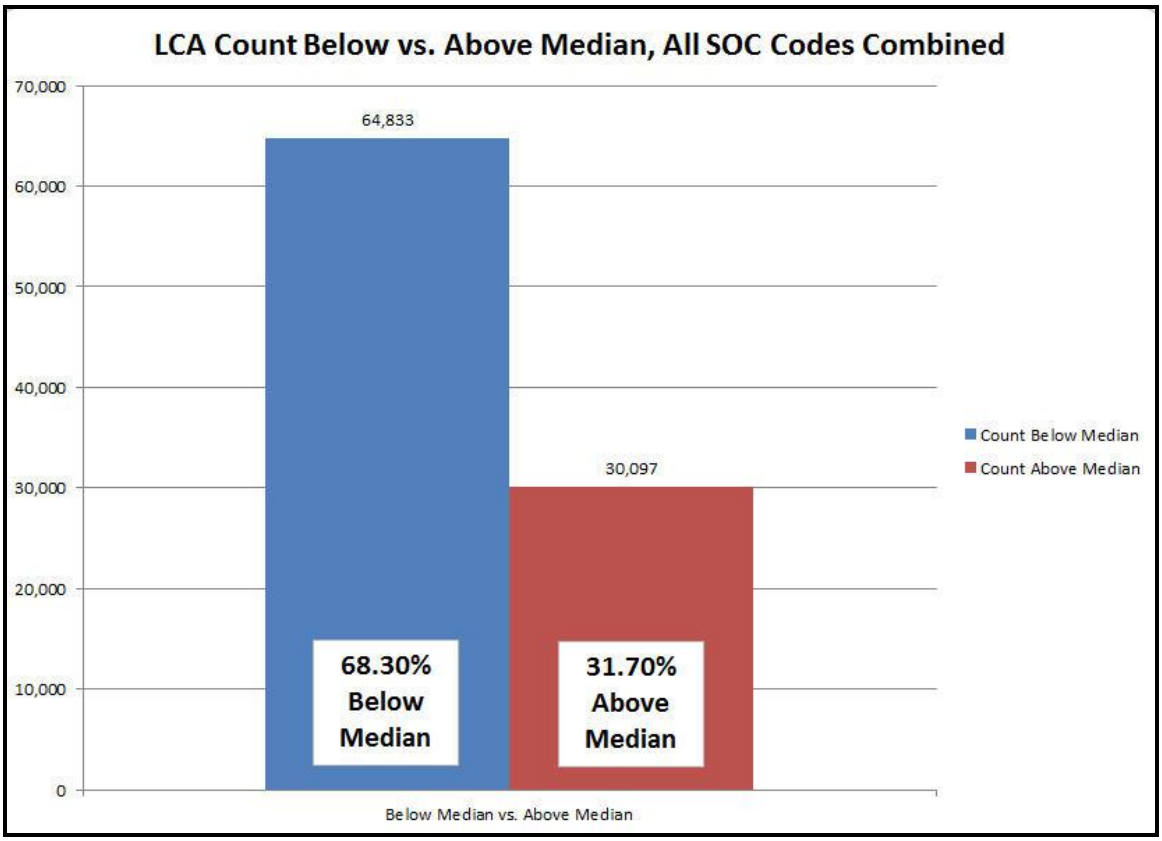


Figure 17: LCA Count, Below vs. Above Median, All 10 SOC Codes

A final presentation of the results is to re-present the data from Figure 12, the primary data condition, as a simple comparison of LCA Median Wages versus Domestic Median Wages to clarify whether any of the 10 SOC codes revealed LCA Wage Means that were above the Domestic Median Wage Means. Figure 18 below represents LCA vs. Domestic Median Wages by SOC Code.

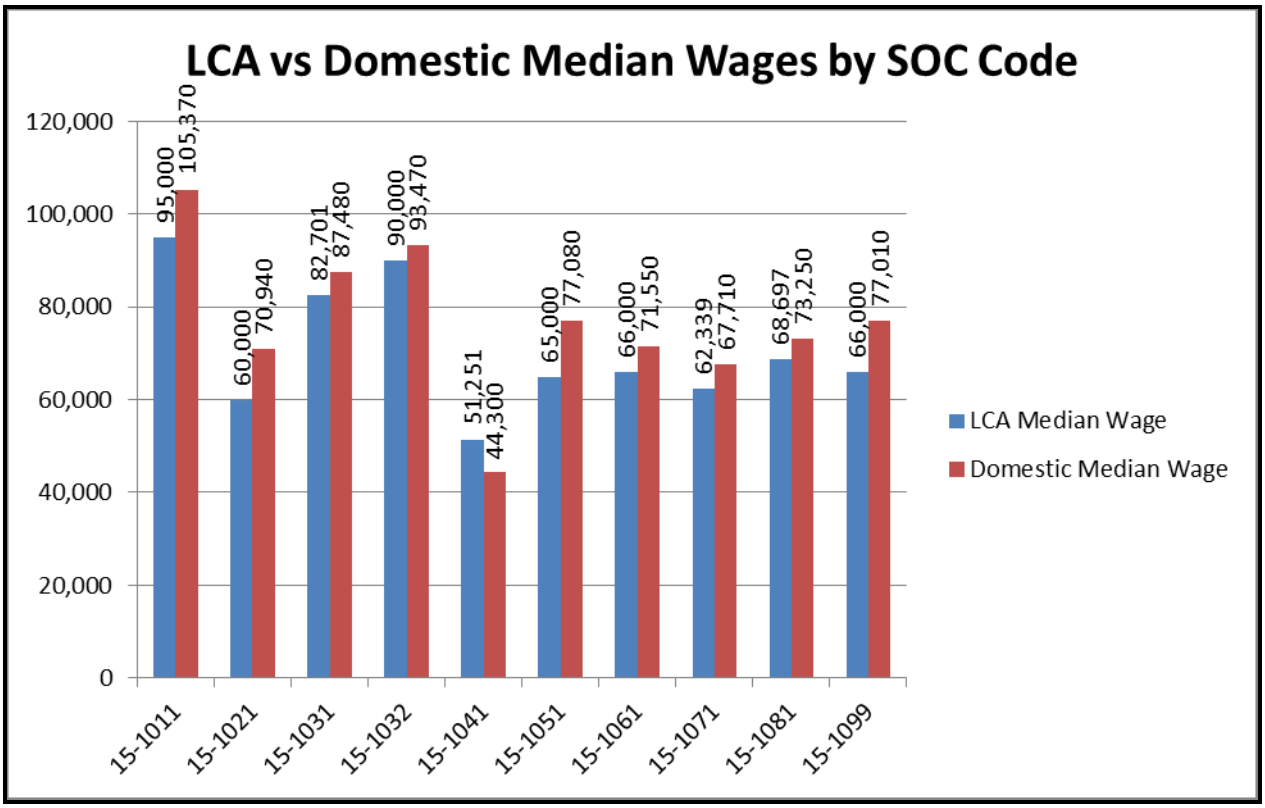


Figure 18: LCA vs. Domestic Median Wages, by SOC Code

For SOC code 15-1011, the median LCA Wage was $95,000 while the median OES Wage was $105,370. Of note, for SOC code 15-1041 the median LCA Wage is actually higher than the median OES Wage, being $51,251 versus $44,300, respectively.

9. Discussion

There are clearly far more computer technology employer requests for H-1B workers via the LCA process than can ever possibly be realized under current law. Congressional caps on the number of H-1Bs available across all SOC codes have remained at the level of 65,000 per year since fiscal year 2004. Yet, computer technology employers alone filed 89,585 LCAs in fiscal year 2010 requesting 232,487 potential workers. Those 89,585 LCAs for computer technology workers represented only about 27% of the total number of 335,328 LCAs filed across all SOC codes for 2010. This study did not count the total number of workers requested across all LCAs, and so it is not possible to calculate the percentage of total requested workers that will ever possibly be hired, but it is less than 1 in 5 if one assumes a ratio of 1 requested worker per LCA (65,000 H-1Bs divided by 335,328 LCAs equals a probability of about 19%).

Four of the 10 computer technology SOC codes represented 92% of all LCAs for the entire computer technology sector.

- **15-1021: Computer Programmers:**
  25,094 LCAs for 89,318 workers.
- **15-1031: Computer Software Engineers, Applications:**
  20,647 LCAs for 34,229 workers.
- **15-1032: Computer Software Engineers, Systems Software:**
  10,468 LCAs for 14,127 workers.
- **15-1051: Computer Systems Analysts:**
  26,273 LCAs for 84,362 workers.

Of these four main SOC code groups, two of them heavily skew the number of requested workers. The 89,318 and 84,362 worker requests (totaling 173,680) for codes 15-1021 and 15-1051, respectively, represent 74.7% of all workers requested. If any effect from weighting was to be seen, it would have been apparent in the results from these two SOC codes. These two

SOC codes are thus examined for the net-effect of weighting the LCAs, and the results are shown in Table 7 below.

Table 7: Median Wages for 15-1021 and 15-1051, Non-Weighted vs. Weighted				
	"From," Non-Weighted	**"From," Weighted**	**Difference ($)**	**Difference (%)**
15-1021	$60,000	$58,739	$1,261	2.10%
15-1051	$65,000	$65,707	$707	1.08%

Examination of the median wages for the two SOC codes determines that weighting the data had no significant effect. For 15-1021, the weighted difference was only $1,261 or about 2.10%. For 15-1051, the weighted difference was only $707 or about 1.08%. Since these two SOC codes had the most extreme potential for weighting being significant, and since analysis shows that it is not in their case, then it is reasonable to conclude that the difference between the weighted condition and the non-weighted condition is negligible.

It was also thought possible that there would be a significant difference between the median wage calculations along the "From" versus "From" ONLY dimension. Table 8 below examines the differences between the median wages for all 10 SOC codes in the "From" and "From" ONLY non-weighted conditions.

Table 8: Median Wages "From" vs. "From" ONLY, All SOC Codes				
	"From"	**"From" ONLY**	**Difference ($)**	**Difference (%)**
15-1011	95,000	100,000	5,000	5.26%
15-1021	60,000	60,000	0	0%
15-1031	82,701	83,405	704	0.85%
15-1032	90,000	90,000	0	0%
15-1041	51,251	50,000	1,251	2.50%
15-1051	65,000	64,771	229	0.35%
15-1061	66,000	65,000	1,000	1.51%
15-1071	62,339	62,000	339	0.54%
15-1081	68,697	67,904	793	1.15%
15-1099	66,000	65,000	1,000	1.51%

Examination of the **DIFFERENCES** between the median wages for each SOC code shows that any difference was negligible. The largest difference was for code 15-1011 at 5.26%. But, that difference was represented in only 675 LCAs out of 89,585 total LCAs, or a mere 0.75% of the sample of all LCAs. Similarly, the next largest difference was seen for code 15-1041, but again that was for a mere 533 LCAs, or a mere 0.59% of the LCAs.

Because it is clear that neither the weighting nor the "From" ONLY conditions showed a significant deviation from the primary data condition of "From," non-weighted, as represented in Figure 12 above, the remainder of the discussion will focus solely on the results from that data condition. Indeed, when one compares the results shown in Figures 12 through 15, one can see the following common patterns that they all four share:

- The LCA Wage Medians (triangles) are always higher than, the claimed Prevailing Wage Medians (squares).
- The claimed Prevailing Wage Medians (squares) track very closely to the 25th percentile OES Wage figures.
- The LCA Wage Medians never exceed the 50th percentile OES Wage figure, with the one exception of code 15-1041.

Because the data points for code 15-1041 are clustered very tightly in Figure 12, the same data was recast in a different form in Figure 18 in an effort to examine it more closely. Figure 18 reveals that LCA Median Wages are indeed less than the Domestic OES Prevailing Wages with that one exception. For 15-1041, the LCA Median Wage was $51,251 while the Domestic OES Median Wage was $44,300. This yields a significant difference between LCA and OES wages, as detailed in Table 9 below.

Table 9: LCA Median Wages vs. Domestic OES Median Wage, Code 15-1041				
	LCA Median Wage	**Domestic Median Wage**	**Difference ($)**	**Difference (%)**
15-1041	$51,251	$44,300	$6,951	15.69%

The LCA Median Wage earned by these H-1B workers is a surprising 15.69% higher than those reported for Domestic OES Wages for the same SOC code. However, it must be remembered that SOC code 15-1041 only had 533 LCAs out of all 89,585 computer technology LCAs, representing a mere 0.59% of the total sample. Therefore, this one single example where the LCA Median Wage was higher than the OES Median Wage is valid, but insignificant.

The OES Wage Division Points listed in Table 5 were used to derive the Standard OES Percentile Divisions listed in Table 6, for all 10 studied SOC codes. It must be understood that the division points for each SOC code were unique, and that there is a direct relationship to the values for each SOC code in Table 5 to Table 6. The critical distinction to be made, however, is that Table 5 represents **OES Wages**, while Table 6 represents **counts of H-1B workers** whose wages fall into the Standard OES Percentile Divisions. These counts were then represented as six bars on Figure 16. The vertical divider in Figure 16 should not be mistaken for additional

data, but rather as a dividing line between below-the-median and above-the-median. The analysis is further carried through to Figure 19, where each side of the below- and above-the-median is aggregated into a binary graph. Figure 19 is re-drawn here for ease of discussion.

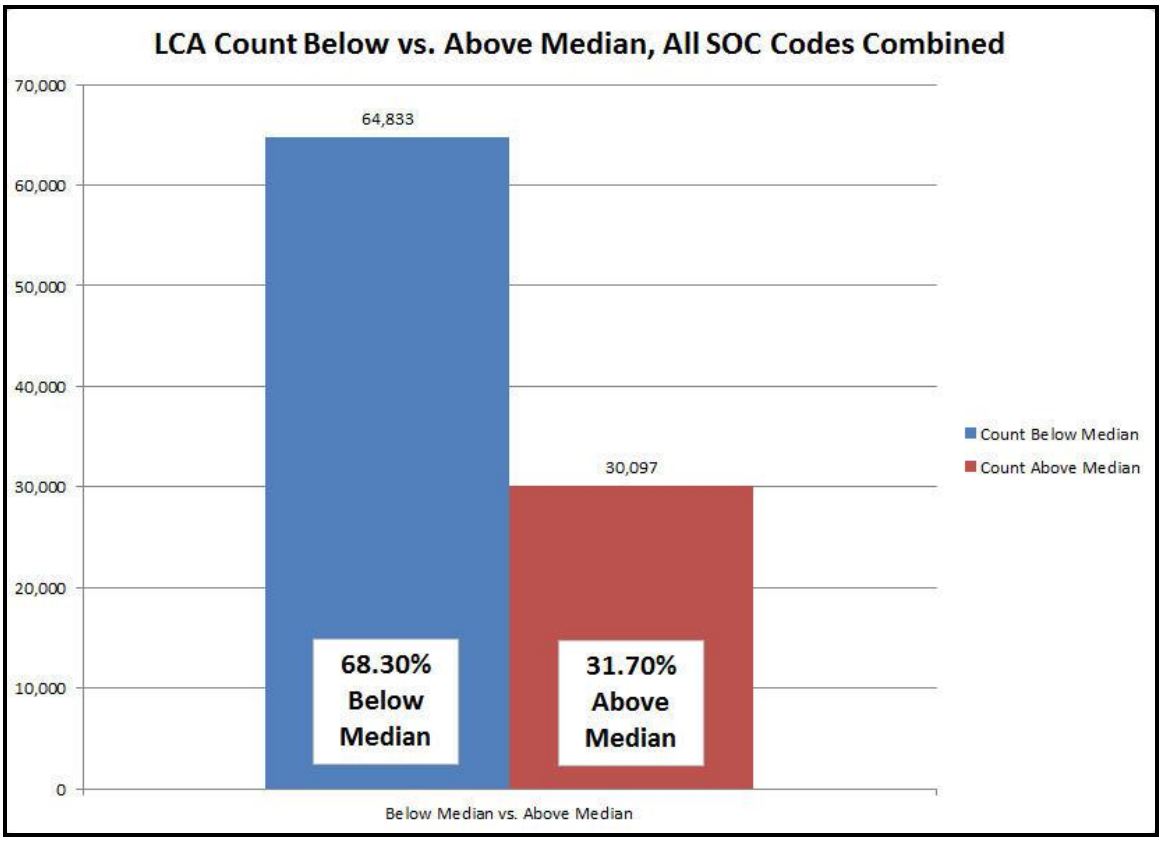


Figure 19: Copy of Figure 17: LCA Count, Below vs. Above Median, All 10 SOC Codes

In any so-called normal Bell curve distribution, one expects to see an even distribution on either side of the median. If we were to imagine a theoretically perfect distribution of H-1B workers below-the-median and above-the-median, it would show an exactly equal number on either side, by definition. The above distribution in Figure 19, however, is anything but "normal." In this distribution, it can be seen that 68.30% of H-1B workers, drawn from fiscal

year 2010 data, are working for wages that are below the OES Median Wages for their SOC codes.

While the same data shows that 31.70% of H-1B workers are working at-or-above the OES Median Wages for their SOC codes, there is still a more than 2:1 ratio of below-the-median versus above-the-median OES Wage earners. Table 10 below further details the breakdown of these OES Standard Division counts, with data drawn from Figure 16.

Table 10: Counts and Percentages of H-1B Median Wages, by OES Percentile Division, All SOC codes.

Row Reference Number	OES Percentile Division	Count of H-1B Workers	Percentage of H-1B Workers
1	1st to 9th	1,629	1.71%
2	10th to 24th	21,806	22.97%
3	25th to 49th	41,398	43.60%
4	50th to 74th	20,512	21.60%
5	75th to 89th	7,640	7.85%
6	90th to 100th	1,945	2.04%
Note: Percentages do not total to 100% due to rounding errors.			

The data shown in both Figure 16 and Table 10 do show a roughly normal distribution, but one that is shifted significantly to the left of the median/mean. If Table 10 showed a normal distribution, it would show that Rows 3 and 4 (the 25th – 49th and 50th – 74th divisions), which both represent 25-point spreads, would have roughly the same counts.

Instead, that relationship exists between Rows 2 and 4. Further, Row 2 represents a 15-point spread (10th to 24th), so that smaller spread of Row 2 has the same count as the larger spread of Row 4. Rows 1 and 6 do show an approximately normal distribution, being very close in value to each other.

Figure 20, below, shows a side-by-side comparison of results of Miano (2007) and the current study, Hill (2011).

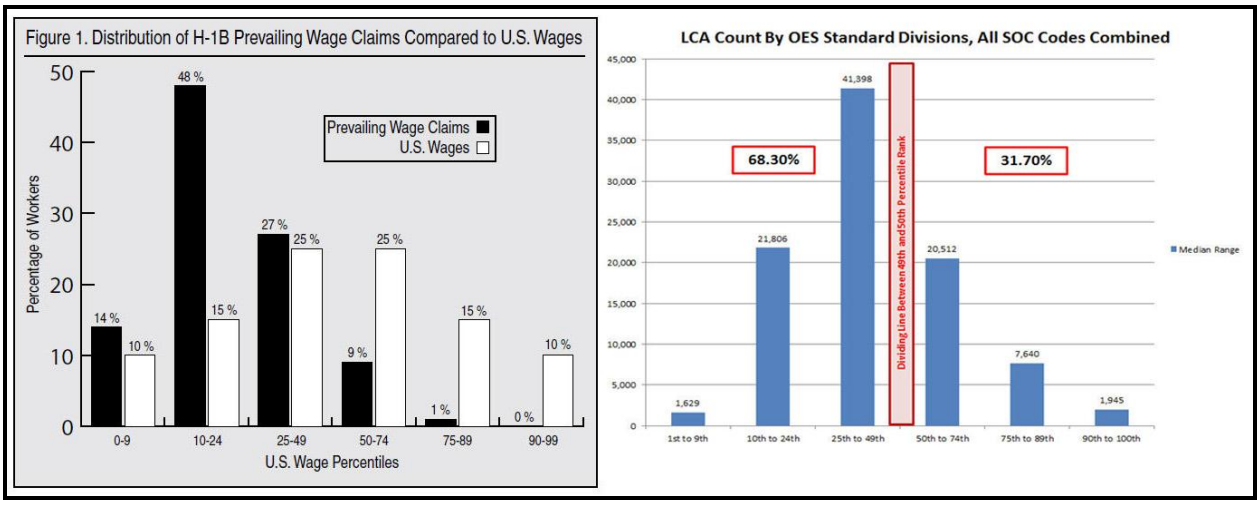


Figure 20: Side By Side Comparison of Fig. 5 (Miano, 2007) and Figure 16 (Hill, 2011) Results

In this side-by-side comparison view, the following conclusions are self-evident:

- Both distributions show a relatively “normal” distribution of LCA Wages.
- Both distributions show a significant skewing to the left of the mean/median.
- The skewing shown in Miano (2007) is more extreme than the skewing shown in Hill (2011).

10. Conclusions

Results indicate that while employers are technically meeting the requirement to pay H-1B workers "at-or-above prevailing wages" for fiscal year 2010, in fact they are choosing prevailing wages that fall primarily between the 25th and 49th percentile ranks.

- H-1B employees are categorically paid less than U.S. workers for the same work, and we can now see this demonstrated with reliable data.
- 68.30% of H-1B workers in the computer technology sector are paid less than the median wages earned by domestic workers.
- Employers are claiming that they must hire H-1B workers because they can't find U.S. workers, but the numbers show that there are powerful financial incentives for choosing H-1B employees over U.S. workers.
- This is perfectly legal, because the iCert system allows employers to choose these artificially low prevailing wages
- These findings are consistent with previously conducted research by Metcalf (2003) and Miano (2007), showing that even though their source data was questionable their conclusions are upheld.

11. Appendix A: Data Acquisition, Details

The text box below provides the standard Structured Query Language (SQL) statement behind the query for a typical SOC code. These queries varied ONLY in the variables in the SOC_CODE names, e.g. “15-1031” vs. “15-1051”. It is provided here to make it possible for others to replicate this work in the future.

```
    SELECT H1B_FY2010.STATUS, H1B_FY2010.FULL_TIME_POS, H1B_FY2010.LCA_CASE_SOC_CODE,
H1B_FY2010.LCA_CASE_SOC_NAME, H1B_FY2010.LCA_CASE_JOB_TITLE, H1B_FY2010.TOTAL_WORKERS,
H1B_FY2010.LCA_CASE_WAGE_RATE_FROM, H1B_FY2010.LCA_CASE_WAGE_RATE_TO,
H1B_FY2010.LCA_CASE_WAGE_RATE_UNIT, H1B_FY2010.PW_1, H1B_FY2010.PW_UNIT_1,
H1B_FY2010.PW_SOURCE_1, H1B_FY2010.PW_2, H1B_FY2010.PW_UNIT_2, H1B_FY2010.PW_SOURCE_2,
H1B_FY2010.LCA_CASE_WORKLOC1_CITY, H1B_FY2010.LCA_CASE_WORKLOC1_STATE

    FROM H1B_FY2010

    WHERE (((H1B_FY2010.STATUS)="CERTIFIED") AND ((H1B_FY2010.FULL_TIME_POS)="Y") AND
((H1B_FY2010.LCA_CASE_SOC_CODE) Like "15-101*") AND
((H1B_FY2010.LCA_CASE_WAGE_RATE_UNIT)="Year") AND ((H1B_FY2010.PW_UNIT_1)="Year") AND
((H1B_FY2010.PW_SOURCE_1)="OES") AND ((H1B_FY2010.PW_SOURCE_2) Is Null Or
(H1B_FY2010.PW_SOURCE_2)="OES"));
```

A layman’s description of this query is as follows.

Display the records in the table H1B_FY2010 with the following specifications:

- Display only the records where the STATUS field is “CERTIFIED,” the FULL_TIME_POS field is “Y” for yes, and the LCA_CASE_SOC_CODE is “Like 15-1031*”. The use of a wildcard “*” character is necessary because some of the records have a shorter form of the SOC code and some have a longer, e.g. “15-1031” and “15-1031.00.” The use of the wildcard character in the query string allows both forms to be returned.

- Also, only display records where the LCA_CASE_WAGE_RATE_UNIT is "Year." This has the effect of filtering out those that are based on "Hourly" or "Weekly" wage rates. The vast majority of the LCAs are reported in terms of yearly wages.

- Similarly, also only display records where the PW_UNIT_1 and PW_UNIT_2 is "Year." If the LCA wage is reported as a yearly unit, then the prevailing wage unit must also be reported in yearly terms. This prevents contamination of the sample where LCA wages are reported yearly but prevailing wages are reported hourly, for instance.

- Similarly, also only display records where the PW_SOURCE_1 is "OES" and the PW_SOURCE_2 is either (null), meaning there's nothing there, or "OES." This prevents contamination of the sample where the prevailing wage source is other than the OES online wage library. The vast majority of the LCA records now use OES as the prevailing wage source.

- For those records that meet all of the above requirements, display the value stored in that field.

- For those records that meet all of the above requirements, also display the values stored in the LCA_CASE_SOC_NAME, LCA_CASE_JOB_TITLE, TOTAL_WORKERS, LCA_CASE_WAGE_RATE_FROM, LCA_CASE_WAGE_RATE_TO, PW_1, PW_2, LCA_CASE_WORKLOC1_CITY, and LCA_CASE_WORKLOC1_STATE.

For further clarification, the following table lists the fields that were drawn from the master table and exported to Excel for manipulation and study. The table lists the formal name of the field, the conditional requirement for selecting the field (if any), and a brief description of what the field is or means.

Table 11: Database Fields from OFLC Source .mdb File

Field Name	Required?	Description
STATUS	"CERTIFIED"	Certified, Denied, Withdrawn, or Pending.
FULL_TIME_POS	"Y"	"Y" for yes, "N" for no.
LCA_CASE_ SOC_CODE	Like "15-1031*"	The SOC code. The formal code includes two digits to the right of the decimal place, but this is often left out. E.g., "15-1031.00" and "15-1031".
LCA_CASE_ SOC_NAME		The official SOC code name.
LCA_CASE_ JOB_TITLE		The Job Title that was entered when the LCA was created. This varies with each LCA.

TOTAL_WORKERS		The total number of workers that are applied for on this single LCA. The vast majority of LCAs are for one worker, but the data includes some LCAs that were for 150+ workers.
LCA_CASE_ WAGE_RATE_FROM		In cases where there are multiple workers per LCA, employers often state a range of proposed salaries. This represents the lowest salary in that range.
LCA_CASE_ WAGE_RATE_TO		This represents the upper range of the salaries for a set of multiple workers on a single LCA. This field may be blank.
LCA_CASE_ WAGE_RATE_UNIT	"Year"	"Year," "Hourly," or "Weekly." Records were only selected if this value was "Year".
PW_1		The lower range of the prevailing wage claimed for this LCA. This is represented as a number, e.g. "86,000."
PW_UNIT_1	"Year"	"Year," "Hourly," or "Weekly." Records were only selected if this value was "Year".
PW_SOURCE_1	"OES"	Indicates the source of the prevailing wage claim. This study is focusing on only those LCAs where the prevailing wage claimed was drawn from the OES data, which is virtually all of the LCAs for 2010.
PW_2		The upper range of the prevailing wage claimed for this LCA, if entered. This is represented as a number, e.g. "86,000."
PW_UNIT_2		"Year," "Hourly," or "Weekly," or (null).
PW_SOURCE_2	(null) or "OES"	For those records that contained a second source, it was possible that that source might be different from that for PW_1. This conditional query statement allows records to be chosen for display if the value of the field is (null) blank or "OES" only to prevent records that contain conflicting wage sources.
LCA_CASE_ WORKLOC1_CITY		The city where the work is to take place. There is a possible field for a second city, but that was not considered in this study.
LCA_CASE_ WORKLOC1_STATE		The state where the work is to take place. There is a possible field for a second state, but that information was not considered in this study.

12. Appendix B: Transforming and Grooming the Data, Details

When data is stored in a database file, certain decisions are made about its "metadata" characteristics. For instance, metadata about whether the data is text, or an integer, or a floating point number is stored along with the data itself.

It was discovered that the numbers representing the wage data, both LCA and prevailing, were stored in the database file as "text." To the human eye, the numbers looked like numbers, but to the computer, the numbers looked like text and were unavailable for even simple calculations, such as addition and multiplication.

Because there were hundreds of thousands of data points with this fundamental flaw, a method was required to convert them from "stored-as-text" values to "stored-as-integer" values. One technique for doing this is to multiply the contents of each cell by the integer "1," which effectively transforms the character of the cell from "stored-as-text" to "stored-as-integer." This was accomplished by picking some empty cell and entering the number "1" in it, then selecting that cell and copying it into the system clipboard.[17] Then, each cell of the column where the "stored-as-text" numbers were found was selected by the technique first selecting the top-most cell, then performing the function CTRL+SHIFT+END to reach the bottom of the data. The focus was redirected solely to the column that required manipulation, and

[17] All manipulation of the data was performed in a Windows environment, with Windows OS and Microsoft Office 2010 conventions. Users wishing to replicate this task on a different platform and/or different applications or versions of Office will need to use the relative techniques for their system.

then a special Paste operation was performed. Using the right-mouse-click command, the Paste Special sub-menu was accessed. On that sub-menu there is the option to perform mathematical operations while pasting the data. By selecting the operation for "multiply," the contents of the cells were effectively multiplied by one. Once this transformation was accomplished by this technique, the LCA and prevailing wage data was available for calculations.

This was repeated for each of the four columns of data that needed this transformation, and for each of the 10 Excel files resulting from the export from the source Access database.