

# **The Conference Board Help Wanted OnLine™ Data Series**

## **Technical Notes**

### **Background**

*The Conference Board Help Wanted Online Data Series™* (HWOL) was first published in July 2005 as a developmental series. A major expansion of the program was initiated in October 2006. The HWOL program provides data on online advertised job demand.

HWOL fills a critical gap in the current U.S. economic indicators by providing timely monthly measures of labor demand (advertised vacancies) at the national, regional, State and metropolitan area levels. These monthly measures are comparable in timing and geographic detail to the Bureau of Labor Statistics (BLS) monthly measures of labor supply (unemployment).

The Conference Board HWOL program is one of the earliest published monthly indicators of economic activity in the previous month, with data publication centered around the 1<sup>st</sup> of each month. The program provides measures of levels and rates for both Total Online Ads and New Online Ads. The online vacancy program is one of the few economic indicators to provide extensive occupational detail, with national estimates published at the major occupational group level and State and MSA estimates at higher level aggregates. To provide users with a broader analytical view of labor supply and labor demand, the press release includes the most recently available data from the Bureau of Labor Statistics on labor supply (unemployment) along with average wage levels for the occupational detail.

### **Coverage**

The HWOL program is targeted to cover the full universe of all online advertised vacancies which are posted directly on internet job boards or through newspaper online ads. At present, ads on corporate web sites for their own jobs are excluded from coverage. However, since a number of job boards scrape these corporate websites, these ads may also appear in the HWOL data count.

### **HWOL Concepts and Definitions**

**Survey Reference Period.** The HWOL program uses a mid-month survey reference period. Data for October would be the sum of all posted ads from September 14<sup>th</sup> through October 13<sup>th</sup>. This reference period was aligned to the BLS unemployment “job search” time period to provide for a more accurate comparison of labor supply and labor demand in the U.S. economy.

**New ads.** New ads are all unduplicated ads which did not appear in the previous reference period. An ad is counted as “new” only in the month it first appears.

**Total ads.** Total ads are unduplicated ads appearing in the reference period. This figure includes both new ads and ads reposted from the previous month.

**Ad Rates.** Ad rates are the number of advertised vacancies as a percent of the BLS civilian labor force data for a geographic area. Ad rates represent the number of ads per 100 participants in the civilian labor force. The HWOL ads rate definition parallels the construction of the widely understood unemployment rate, i.e. unemployed persons divided by the civilian labor force.

**Supply/Demand Rate.** The supply/demand (S/D) rate is the number of unemployed divided by the number of advertised vacancies (i.e. the number of unemployed per advertised vacancy). The S/D rate provided an indication of the tightness of the labor market and whether overall supply and demand is out of balance. Additionally, because of the economic trend relationship between unemployment and vacancies, this rate should also provide a sensitive indicator of trend changes in an area or occupation.

**Regional data.** Regions are as defined by the U.S. Census Bureau.

**Metropolitan area data.** The Conference Board uses the 2005 Office of Management and Budget (OMB) county-based MSA definitions for its data whereas the BLS programs in this report use the OMB alternative NECTA (New England Cities and Town Area) MSA definition.

**Occupational data.** Occupational data use the 2000 OMB Standard Occupational Classification (SOC) system. Ads are coded to the 6-digit SOC level and 8-digit O\*Net level whenever possible; at present, about 99% of the ads are assigned an occupational code. Data in the monthly release is aggregated to the major occupational group level or higher.

**Unemployment data.** The unemployment and labor force data used in this release come from the BLS Current Population Survey (CPS) and the Local Area Unemployment Statistics (LAUS) program. Taken together, both programs provide a timely and accurate profile of labor force information for the nation and all major levels of geographic detail.

**Wage data.** The average hourly wage data used in this release are the latest available from the BLS Occupational Employment Statistics (OES) program. The OES wage data provide an accurate, comprehensive, point-in-time snapshot of wage levels across all 800 SOC occupations at the national, State and MSA levels.

**BLS Job Openings Data.** The BLS publishes monthly job openings data from its Job Openings and Labor Turnover Survey (JOLTS) <http://www.bls.gov/jlt/home.htm>. This program provides a broad national picture of hiring activity which includes openings, hires and separations. The Conference Board HWOL data series will complement the data found in the JOLTS program and provide detailed regional, State and MSA data on labor demand along with occupational detail for the various geographic breakouts.

## **Methodology**

**Data Collection.** Wanted Technologies collects online ads from over 1200 job boards in the United States, covering both online newspaper ads and internet job board ads. Internet job boards that are aggregators (i.e., only scrape ads from other boards and provide no unique ads) are identified and removed from active collection in order to eliminate a major source of duplication in counting online ads. While there is no official comprehensive list of all internet job boards, Wanted Technologies uses its own independent research on internet job boards, recommendations from its newspaper and job board clients located in MSAs across the U.S., and other sources to identify new internet job boards not currently covered. This process results in monthly updates to its coverage. Job boards that cover smaller niche markets are included in HWOL, however, smaller local internet job boards in an area with a limited number of ads may not be targeted for collection.

**Duplication.** Data in the HWOL data series reflect unduplicated ads. A major issue in producing estimates of the actual level of advertised vacancies for a geographic area is the elimination of duplicate ads. There is a tremendous amount of ad-scraping within the industry and there are large nationwide job boards that contain only scraped ads. As noted above, Wanted Technologies first identifies the job boards which are only aggregators of ads from other job boards and eliminates them from active

collection. For the remaining 1200+ boards which are under active collection, Wanted Technologies uses its proprietary software to categorize each ad by a number of key variables including company name, job title/description, and location. Ads are then compared across all boards and duplicates are eliminated from the HWOL published estimates. This process significantly limits the level of potential duplicates in the final estimates. The unduplication process reduces the count of overall ads collected from over 14 million ads to over 3 million ads after unduplication – indicating that duplicates represented about 2 out of 3 of the ads prior to unduplication. The resultant unduplicated ad levels for the HWOL program compare favorably to those produced by the BLS Job Openings Labor Turnover (JOLTS) program after allowing for coverage and definitional differences.

**Occupational coding.** Occupational coding is done at the 6-digit Standard Occupational Classification (SOC) level and the 8-digit O\*Net level using autocoder software selected by The Conference Board for its accuracy. The SOC coding used in HWOL is the same definitional coding used for federal employment and unemployment statistics. The Conference Board upgrades the SOC classification of HWOL data on a regular basis and reclassifies the entire database to allow accurate comparisons to be made across time periods. In 2011, 99 percent of all ads are classified at the 6-digit level.

**Area coding.** The area coding for an ad is determined first by the location cited in the text of the ad itself. If this is not present in the ad, the location for a job is determined by checking for the address of the company placing the ad against a universe list of establishments. If this is unsuccessful, the location of the job board is stipulated as the location of the ad. Some jobs are designated simply as nationwide. In this case the ad would appear in the national total but not in any regional, state, or metropolitan area totals.

## **Reliability of Estimates**

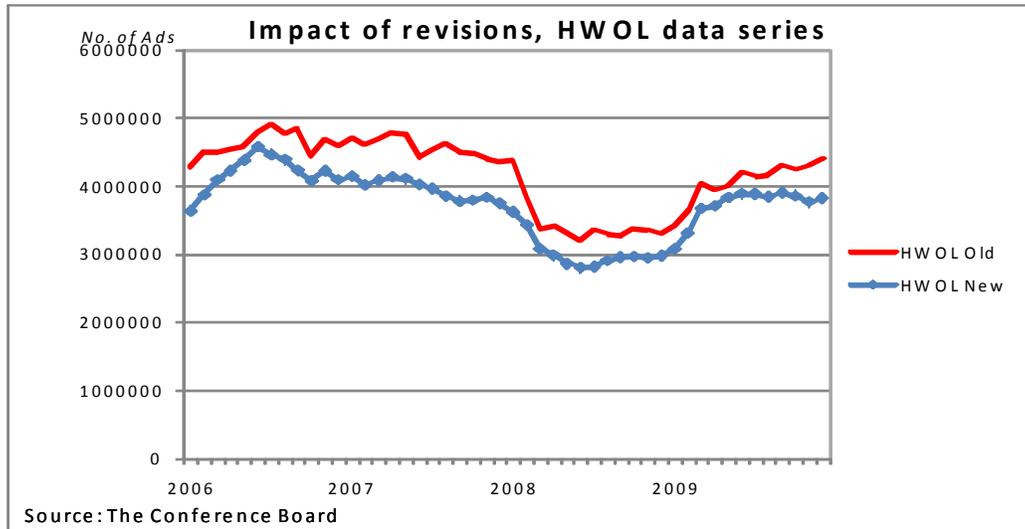
The HWOL program is essentially a universe count and is not subject to the typical sampling error and non-response error components associated with most statistical surveys. The non-sampling error sources for the HWOL program would include population under-coverage due to missing a portion of the targeted population (e.g. a large Internet job board), and over-coverage due to the inability to fully eliminate duplicate ads from survey estimates. Additional potential sources of non-sampling error would include occupational and/or geographic coding errors which could affect the proper classification of individual ads.

## **2011 Annual Revisions**

As a part of the annual HWOL program revision, a number of updates to the historical series were implemented with the release of the January 2011 data. To ensure the consistency and continuity of the time series data, all of the updates described below are applied across the full HWOL time series history, from May 2005 through the current time period. With the January release, revised HWOL time series history from January 2007 forward are being released. The full time series history from May 2005 will be available with the April release on May 2, 2011. This year's revision included: adjustments to the HWOL job board coverage, elimination of "invalid" job ads, improvements to the unduplication methodology, improvements to the occupational coding software, and annual updates to the seasonal adjustment factors.

**Time Series Summary.** The levels in the revised series (see Chart below) are, on average, about 460,000 per month lower than the levels in the prior series; approximately 200,000 of this is due to the elimination of invalid job ads with the remainder attributable to adjustments to the job board coverage and improvements in the unduplication methodology. As can be seen in the chart, the series turning points and the trends have remained essentially unchanged (with the exception on the trend in the last half of 2010). Additionally, the overall losses of 1.7 million job ads during the series downturn (April 2007 through

April 2009) and the gains of 1.0 million job ads during the series early recovery (April 2009 through April 2010) also remained essentially unchanged. Specific details of the revision changes are outlined below.



**Job Board Coverage.** The HWOL program collects data on a daily basis from over 1,200 online job board sources. Each year, new sources are added as they emerge while some existing sources may be dropped, if it is determined that they primarily aggregate their data from other job board sources. This year, a more extensive job board review and analysis was performed for identifying any remaining aggregator job boards; this review has resulted in the elimination of several job boards. In combination with the unduplication improvements, these changes resulted in lowering the series levels by about 260,000 ads per month.

**Occupational Coding.** The HWOL program uses Standard Occupation Classification (SOC) autocoder software which codes over 99 percent of all ads to the 6-digit SOC and 8-digit O'Net level. The latest version of the autocoder incorporated a new feature for identifying and classifying "invalid" job ads into a new Miscellaneous category. This category contains jobs which would generally not be included in the official government employment and job openings data series and are now being classified as "invalid" job ads by the HWOL program. These ads include such categories as: adult ads, get-rich-quick ads, human donors wanted ads, human test participants wanted, products/services-offered ads, job fair ads, and other various types of ads. The Miscellaneous category has now been dropped from the HWOL time series and this change has resulted in lowering the series levels by an average of about 200,000 ads per month.

**Seasonal Adjustment.** The HWOL program is initiating a new practice of calculating and publishing new seasonally adjusted series with the release of each year's January data. New seasonal factors are calculated using historical data from May 2005 through the most current December data.