# The Conference Board Help Wanted OnLine Data Series™

# Technical Notes

## Background

The Conference Board began publication of the Help Wanted Online Data Series™ (HWOL) in July 2005 as a developmental series and initiated a major expansion of the program in October 2006. This new developmental program of online job advertised provides a rich view of labor demand and complemented The Conference Board's long-standing Help Wanted Advertising Index™ of newspaper print advertising which is currently available for research purposes but is no longer published monthly.

The Help Wanted Online Data Series™ 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 1st 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 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, with the August 2008 data Direct Employers Association's *Job Central* job board officially replaced the U.S. Department of Labor's America's Job Bank (AJB) in the HWOL estimates. *Job Central* provides a centralized, high-quality source of current ads which appear in cooperation with many of the large U.S. corporate job boards as well as most of the State-sponsored job boards. In addition, a number of job boards scrape corporate websites and these ads may also appear in the HWOL data count.

## Concepts and Definitions

**Survey Reference Period.** The HWOL program uses a mid-month survey reference period. For example, data for October would be the sum of all posted ads from September 14th through October 13th. 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 all unduplicated ads appearing in the reference period. This figure includes both new ads and reposted ads.

**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.

**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 8-digit SOC level; at present, over 90% of the ads were able to be 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.** Twice each week, Wanted Technologies (WT) 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 clients located in MSAs across the U.S., and sources such as Weddle's Directory of Employment-Related Internet Sites to identify new internet job boards not currently covered. This process results in monthly updates to its coverage. 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 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 weekly 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 13 million ads to over 4 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.** Job ads are coded at the 6-digit SOC (Standard Occupational Classification) level using a job title and job description matching algorithm. With the release of the July 2008 data, historical data for February 2008 thorough June 2008 have been revised to incorporate an improved SOC methodology. The improved mythology assigns occupational coding to around 95% of the ads, and is a substantial improvement from the pervious level of occupational coding for 60-65% of the ads. All remaining months from the beginning of this series in May 2005 forward will also be revised to incorporate this new coding methodology.

**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.

**National Estimation Methodology.**
The HWOL national and regional estimation methodology was update with the August 2008 data based on the results of research by The Conference Board which focused on the existing unduplication procedures. It was determined that ads which appear in more than one state overwhelmingly represent unique job postings in each location. As a result of this evaluation the national and regional totals for the HWOL data series are the direct sum of the individual State totals (including the District of Columbia). The national historical data series was revised to reflect this change. State and MSA data were unaffected by this change.

## 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.