A more human resource.℠

The ADP National Employment Report

Introduction

Payroll employment released by the Bureau of Labor Statistics on the first Friday of each month is one of the first indicators that the U.S. government publishes after the end of each month. As such, it garners considerable attention as an early barometer of the health of the economy, specifically the labor market. The business community keenly follows the report, and unexpected movements in the report can move U.S. and global financial markets significantly.

Given Automatic Data Processing Inc.'s large client base, which includes about one-fifth of U.S. private payroll employment, it is possible to get a first glimpse at the BLS payroll employment report (Current Employment Statistics) prior to the official release by the BLS. To this end, ADP in partnership with Moody's Analytics has developed a method of predicting the private-sector employment two days before the release of the CES report. This collaboration of ADP with Moody's Analytics has produced the ADP National Employment Report.

Although ADP's client base does not precisely match the distribution of employment as reported by the BLS, Moody's Analytics adjusts the ADP data to match the CES distribution. The methodology used by Moody's Analytics is described below.

A more human resource.℠

Data processing

Moody's Analytics gets access to ADP payroll files, which are compiled into time series beginning in 2001. The data processing involves a number of steps, including the removal of outliers, identifying clients by industry and company size in order to aggregate the individual client data into industries and company size, the creation of matched pairs, seasonal adjustment, and adjustments to match the industry and size distribution of the BLS' employment data.

Much of the data processing mimics the methodology used by the BLS to calculate payroll employment for the CES report.

The files are monthly through 2006, biweekly thereafter until July 2009, and weekly since August 2009. Moody's Analytics uses the following data fields from ADP's weekly files: number of employees, frequency of payment, processing date of payment, industry and location of establishment (where payrolls are processed), and whether the establishment is active or terminated. A company may process the payrolls of employees that work in a number of locations in one central location or a company may process payrolls in several locations.

Unlike the BLS, which tracks employees paid during the month, the ADP methodology includes employees who are active. These are chosen over paid because employees may receive more than one paycheck for incentive pay or bonuses. This would result in multiple records and, therefore, the inclusion of such records would overstate the actual number of employees on a company's payroll.

Moody's Analytics checks the raw data for outliers, anomalies and inconsistencies. This may mean adjusting records for which there are missing or incorrect data fields. If the correct data field (such as the industry or location) cannot be determined, the record is excluded.

Moody's Analytics also assigns to each record an industry from the North American Industrial Classification System used by the BLS. Since some ADP clients still classify their businesses using the Standard Industrial Classification, which was in effect prior to the introduction of the NAICS in 1997, Moody's Analytics uses a NAICS-SIC mapping to reclassify those establishments into NAICS industries. Companies are classified into the 13 NAICS private industry sectors: (1) natural resources and mining, (2) construction, (3) manufacturing, (4) trade, transportation and utilities, (5) information, (6) financial activities, (7) professional, scientific and technical services, (8) management of companies and enterprises, (9) administrative and support services, (10) education , (11) healthcare, (12) leisure and hospitality, and (13) other services.

Matched pairs

Moody's Analytics then creates matched pairs of establishments that have reported employment in two consecutive months. There are more than 400,000 such matched pairs, representing more than 23 million employees in the U.S. Each month's data include only the matched pairs available in that month. Matched pairs are aggregated and matched-pair growth rates of employment are computed into cells made up of nine size classes and the 13 NAICS super sectors.

The CES measures the number of people on payrolls during the pay period that includes the 12th of the month (the reference period). A pay period can be any length of time; the most common pay frequency is weekly. But a pay period can also cover two weeks; it can be bimonthly, monthly, etc. Since the ADP records provide pay dates rather than pay periods, matched pairs must be constructed using interpolation. By combining the pay date and the frequency of pay, Moody's Analytics matches the BLS pay period concept as closely as possible.

Most of the ADP clients do not pay their employees on the 12th. Rather, they may pay workers before or after that date. If there is no recorded employment for the pay period that includes the 12th, but a record exists for either a later or earlier pay period during the month, Moody's Analytics estimates employment for the reference period by linearly interpolating between the level of employment on the prior record and the record for the later pay period. The maximum time range for linear interpolation to capture missing employment on the 12th depends on the payment frequency of an establishment. If the previous record falls out of this range, the recorded employment in the same month immediately after the 12th is used for the employment on the 12th.

Seasonal adjustment

After calculating the matched employment growth in every industry by size class, Moody's Analytics applies the X-13ARIMA-SEATS method to remove outliers and seasonally adjust these employment growth series. Deseasonalized trends for the industrial cells are recalculated with each new month of data. The X-13ARIMA SEATS is the seasonal adjustment method developed by the U. S. Census Bureau in collaboration with the Bank of Spain that integrates an enhanced version of X-12-ARIMA with an enhanced version of SEATS to provide both X-11 method seasonal adjustments and ARIMA model-based seasonal adjustments and diagnostics. This is the seasonal adjustment software the Census Bureau uses.

Each observation in an industrial cell is then compared with the trend value, and outlier observations are removed. Matched employment growth for the 90 size classes is calculated for the second time using the cleaned data. Each cell's employment growth is seasonally adjusted again using the same X-13ARIMA method.

Complicating this process is the fact that in some months there are five weeks, rather than four, between the survey weeks used by the BLS. The bureau employs an adjustment to account for the longer period. A similar adjustment is made using the ADP payroll data: regressing the growth rate in each cell on a dummy variable which, if significant, is used to eliminate the long-month effect.

Adjusting the ADP data to align with the BLS sample

Although ADP processes the payrolls of companies representing 20% of all U.S. employment, the composition of those firms does not exactly match the size and industrial composition of the firms sampled by BLS in the CES. Therefore the Quarterly Census of Employment and Wages, which is a complete count of payroll employment derived from unemployment insurance tax records, is used to adjust the ADP data. The QCEW data by the nine size classes are available only for March of every year. This is the month to which BLS benchmarks its CES estimates against the QCEW data. The seasonally adjusted matched-sample growth rates by industry are computed by taking a weighted average of the matched sample growth rates by size within each industry. The weights are based on monthly

interpolations of the QCEW March values. For the time period after the last benchmark, the weights are extrapolated forward to the latest month using the ADP matched-pair growth rates within each size class.

Regression and results

Moody's Analytics uses a restricted structural vector autoregression model to estimate the monthly change in private payrolls by regressing the most recent payroll growth reported by the BLS for each super sector on a constant term and (1) ADP matched-pair growth rates by industry, (2) lagged values of BLS estimates of growth of employment by industry with industry specific restrictions, (3) unemployment insurance claims (UNI_US), (4) oil prices, (5) the Michigan Consumer Sentiment Index (CSENT_US), (6) and the Composite Index of Leading Indicators (LEAD_US) . Unemployment insurance claims are used as a common driver across all industries. They have replaced the ADS index, which has been used as a common external driver in previous specifications, since they have shown superior predictive power in out-of-sample tests and capture the monthly volatility in employment changes more efficiently. The driver enters the equations with a negative sign, which signals lower expected employment growth rates when unemployment insurance claim levels increase. The remaining three external drivers are restricted to affect selected industries. Oil prices drive natural resources and mining industry. The Composite Index of Leading Indicators drives the manufacturing and administrative and support services and has a positive impact on employment projections. The Michigan Consumer Sentiment Index enters the construction equation with a positive sign. The equations for all industries are estimated simultaneously. Thus, the model is made up of 13 endogenous variables—the 13 BLS super sector industries—and 17 exogenous variables—the 13 ADP super sector industries, unemployment insurance claims, oil prices and the composite index of leading indicators.

The equation for each industry is constructed as follows:

For example,
% change in construction = -1.17 +0.47 * % change in size-weighted ADP construction +0.203 one-month lagged growth in natural resource and mining -0.206 * two-months lagged growth in natural resources and mining +0.08 * three-months lagged growth in natural resources and mining +1.09 * one-month lagged growth in manufacturing -0.35 * two-months lagged growth in manufacturing -0.40 * three-months lagged growth in manufacturing +0.25 * one-month lagged growth in construction -0.24 * two-months lagged growth in construction +0.13 * three-months lagged growth in construction -0.29 * one-month lagged growth in transportation and utilities +1.16 * two-months lagged growth in transportation and utilities -0.27 * three-months lagged growth in transportation and utilities -0.19 * one-month lagged growth in information -0.10 * two-months lagged growth in information +0.087 * three-months lagged growth in information -0.30 * one-month lagged growth in finance and insurance -0.99 * two-months lagged growth in finance and insurance +1.03 * three-months lagged growth in finance and insurance +0.12 * one-month lagged growth in education -0.21 * two-months lagged growth in education +0.001 * three-months lagged growth in education -1.02 * one-month lagged growth in healthcare +1.85 * two-months lagged growth in healthcare -0.54 * three-months lagged growth in healthcare +0.62 * one-month lagged growth in leisure and hospitality +0.24 * two-months lagged growth in leisure and hospitality +0.76 * three-months lagged growth in leisure and hospitality -0.37 * one-month lagged growth in professional, scientific and technical services +0.48 * two-months lagged growth in professional, scientific and technical services -0.63 * three-months lagged growth in

professional, scientific and technical services +0.34 * one-month lagged growth in management of companies and enterprises +0.53 * two-months lagged growth in management of companies and enterprises -0.52 * three-months lagged growth in management of companies and enterprises -0.41 * one-month lagged growth in administrative and support services -0.15 * two-months lagged growth in administrative and support services -0.41 * three-months lagged growth in administrative and support services +0.15 * one-month lagged growth in other services -1.31 * two-months lagged growth in other services -0.22 * three-months lagged growth in other services -0.0136 * Michigan Consumer Sentiment Index -0.0216 * percentage change in unemployment insurance claims.

This method allows different trends of employment by size of payroll within industries, while assuming that the other industry-wide relationships implied by the regressions hold for all size classes within an industry. Inclusion of lagged employment in the regressions controls for shifting differences between the BLS sample and the ADP customer base, and inclusion of other employment-related variables controls for differences in the definitions of employment used by BLS and ADP and the difference in sampled firms. The industry-level data are then aggregated to total private employment using industry weights from the BLS' QCEW.

The resulting growth rate estimates are then converted into differences. The simple correlation between the monthly percentage change in total nonfarm private employment as generated by Moody's in-sample forecast is 0.78 since January 2010. The regression results track the BLS' CES closely. The table below lists the simple correlation between the growth rates generated from the regression results and the BLS data for each reported super sector.

Industry	Correlation
Goods-producing	
Construction	0.860
Manufacturing	0.805
Natural resources and mining	0.950
Services	
Trade, transportation & utilities	0.783
Finance and insurance	0.851
Healthcare	0.708
Education	0.748
Professional, scientific and technical services (54)	0.661
Management of companies and enterprises (55)	0.637
Administrative and support services (56)	0.691
Leisure and hospitality	0.823
Information services	0.794
Other services	0.745

The final step of the process is to use the estimation results to predict the current month's BLS number by super sector and size class, which can then be aggregated into total private sector employment.

Company-level estimates

ADP data are used to construct employment by five company size classes: small companies (1-19, 20-49), medium companies (50-499), and large companies (500-999, 1,000+).

Each control unit in the cleaned ADP datasets with outliers removed is matched to a company, identified by the employer identification number. The size of the control unit may coincide with the size of the company, or the company can comprise many control units. The employment of all the control units with a common EIN is summed up to generate the company-level employment.

The companies in each super sector are then grouped into nine cells according to their sizes, and the matched pair employment growth for each of the 90 industry-company cells is calculated. Using BLS methodology, if a company includes control units from more than one industry, each control unit is allocated to the appropriate industry. However, the company size to which the control unit is allocated is the company size of the combined control keys. For example, if a company is made up of two control units, each with 15 employees in two industries, they are allocated to company size 20-49, but the number of employees allocated in this company size to each industry is still 15. The matched employment growth for each industry-company cell is then seasonally adjusted and the long-week effects are removed.

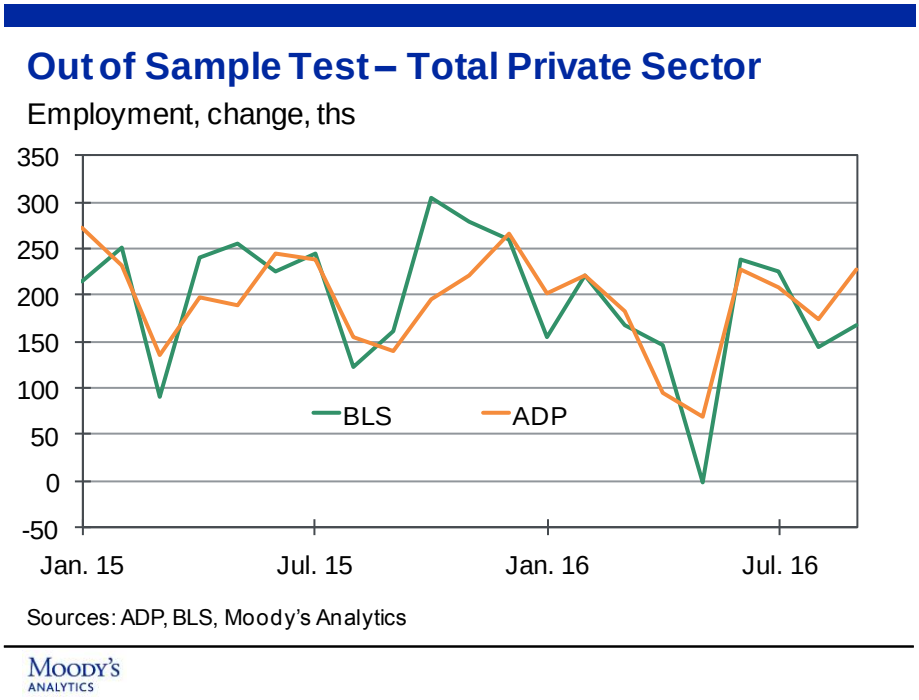
The QCEW establishment-based industry cell sizes and shares of each company cell within every establishment cell are used to construct the company-based industry cell size in the each year's March benchmark month. The benchmarked company cells are then extrapolated to grow at the matched-cell employment growth rates. Monthly distributions of company cells in every super sector are calculated from the extrapolated company cell employment levels.

The monthly industry distribution of company cells and the forecasted industry employment growth are then combined to produce the forecasted employment growth for company cells in the NER.

A more human resource.℠

Out-of-sample validation results

Out-of-sample validation tests have been performed for all industries with one-month holdout period. The results indicate that the ADP U.S. National Employment Report follows closely the BLS numbers on a monthly basis.

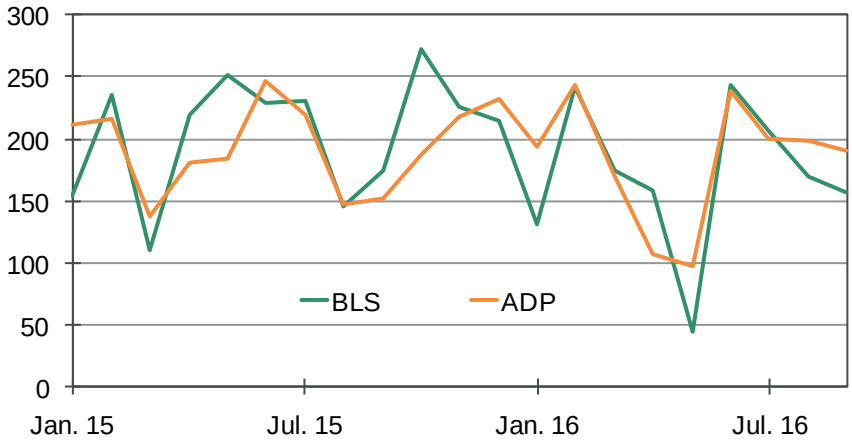


Average Out-of-Sample Forecast Error	Mean Square Error	Mean Absolute Error
-10.30	2,160.87	31.45

A more human resource.℠

Out of Sample Test – Service Industries

Employment, change, ths



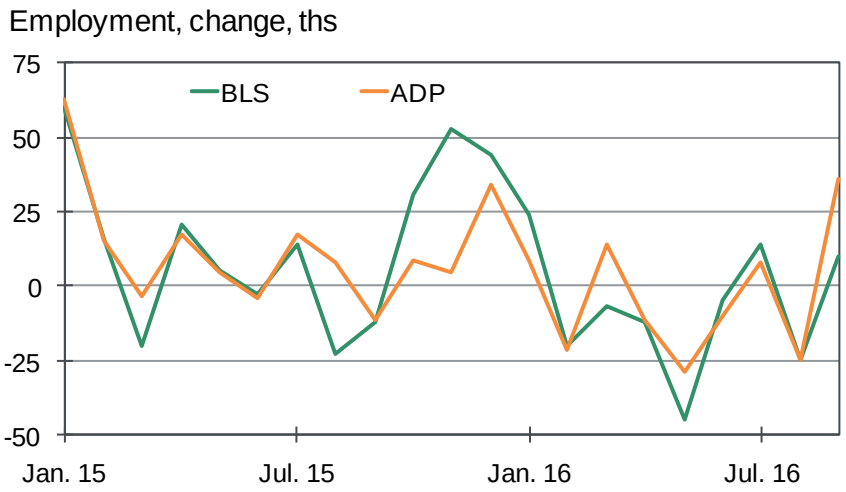
Sources: ADP, BLS, Moody's Analytics

Moody's Analytics

Average Out-of-Sample Forecast Error	Mean Square Error	Mean Absolute Error
-4.504	1,381.43	27.32

A more human resource.℠

Out of Sample Test – Goods Producing Industries



Sources: ADP, BLS, Moody's Analytics

Moody's Analytics

Average Out-of-Sample Forecast Error	Mean Square Error	Mean Absolute Error
-6.39	326.84	13.5

A more human resource.℠

About Moody's Analytics

Moody's Analytics helps capital markets and credit risk management professionals worldwide respond to an evolving marketplace with confidence. With its team of economists, the company offers unique tools and best practices for measuring and managing risk through expertise and experience in credit analysis, economic research, and financial risk management. By offering leading-edge software and advisory services, as well as the proprietary credit research produced by Moody's Investors Service, Moody's Analytics integrates and customizes its offerings to address specific business challenges.