glassdoor | ECONOMIC RESEARCH

April 2019

Methodology: Glassdoor Job Market Report

By
Andrew Chamberlain
Chief Economist

Daniel Zhao
Senior Economist

I. Introducton

The Glassdoor Job Market Report provides a real-time view of job and hiring trends and wage growth in the U.S., including several metro areas, based on millions of online jobs and salaries on Glassdoor.

As one of the world's largest job sites, Glassdoor collects millions of job postings from a wide variety of online sources each month. In addition, Glassdoor is a leading source for real-time salary information, based on millions of crowd-sourced salary reports submitted online, voluntarily and anonymously, by current and former employees.

Leveraging the same technology behind Glassdoor's powerful job search and Know Your Worth products, the Job Market Report applies proprietary data science and machine learning algorithms to Glassdoor data to aggregate and analyze online job openings and estimate wage growth across the nation, including in several U.S. metro areas.

The following provides a detailed methodology behind our estimates of open jobs and median base pay, leveraging Glassdoor data. Section II explains how we estimate the number of open jobs online. Section III outlines the methodology behind our estimates of median base pay. For more information, please visit our "frequently asked questions" page at https://www.glassdoor.com/research/studies/glassdoor-job-market-report-faq/.

II. Open Jobs

A. How It Works

Glassdoor's technology gives us the power to be one of the world's largest and most advanced jobs aggregators. Every day, our algorithms crawl the web to find every online job posting available on company career sites, job boards, and applicant tracking systems used by employers in the United States and around the world. The goal of Glassdoor's jobs ecosystem is to compile millions of online job listings and make them available free of charge for job seekers.

Glassdoor's online job openings data represents a significant fraction of the labor market. As of January 2019, the BLS estimates there were roughly 7.49 million job openings in the U.S. On Glassdoor,[1] we estimate there were approximately 5.85 million online open jobs at that time – 78 percent of the BLS estimate. Previous studies estimate that between 60-70 percent of all actual job openings, as of 2014, appeared as online job ads (See Carnevale et al. 2014). Since then, even more hiring has moved online – making online jobs data from Glassdoor a powerful real-time measure of labor demand by American employers.

B. Sources for Jobs Data

There are three main sources of online jobs aggregated by Glassdoor:

- Jobs that appear on job boards (such as Monster, Jobg8, and Jobs2Careers);
- Jobs listed by companies in their applicant tracking systems (such as Taleo, iCIMS, and BrassRing); and
- Jobs listed on companies' own career sites.

Jobs from Job Boards: Although Glassdoor collects data from job boards, the Job Market Report does not use data from job boards in our count of open jobs online. Job board data can suffer from duplication issues (e.g., an employer posting the same job to multiple job boards). Many job board postings contain "spam" advertising text rather than actual job titles. And many job postings from job boards are outdated, remaining active after actual positions have been filled.

Jobs from Applicant Tracking Systems: Most employers today use some form of applicant tracking system (ATS) to manage their hiring processes. Applicant tracking systems like Greenhouse, iCIMS and BrassRing allow company HR professionals to list open jobs, manage incoming applicants and organize the company's interview and hiring process. Most of these applicant tracking systems today feature open-source application programming interfaces (APIs) which allow Glassdoor to collect company-listed open jobs directly. Job listings from ATSs do not suffer from the systematic job duplication problems present in job board data. They reflect actual job requisitions created by company HR professionals, which companies have listed online to attract hires. For this reason, job listings from ATSs are a credible source of information about actual hiring intent by companies.

1. Source: Not seasonally adjusted job openings from BLS JOLTS, available at https://data.bls.gov/timeseries/JTU00000000JOL

II. Open Jobs *(continued)*

Jobs from Company Career Sites: Some employers do not list their open jobs on an applicant tracking system, or use an ATS without a public API. However, these employers still typically list their open jobs on their own company career site. Using thousands of automated crawls of individual company career sites, Glassdoor collects open job information from companies themselves and makes it available to job seekers on Glassdoor. Because open jobs listed directly on company career sites are carefully maintained and curated by companies' HR teams, they are a very high quality source of job listings with little duplication. As with job postings from applicant tracking systems, online jobs listed on individual company career sites are a highly credible source of information about hiring intent by firms.

In total, Glassdoor collects online job postings from hundreds of applicant tracking systems and tens of thousands of automated crawls of company career pages covering, as of April 2019, more than 93,000 unique U.S. employers. This coverage provides a broad sample of open jobs online in America.

C. How We Count Open Jobs Online

The definition of "open jobs" in the Glassdoor Job Market Report is conceptually similar to the definition of U.S. job openings in the Bureau of Labor Statistics (BLS) Job Openings and Labor Turnover Summary (JOLTS) survey.[2] The JOLTS survey asks employers how many open jobs they have available and are ready to hire for as of the last business day of each month. Our definition of online open jobs counts the number of unique open jobs on Glassdoor that were listed as active on each Monday of the week. This is a "snapshot" definition of open jobs: The counts represent how many unique jobs were open on Glassdoor at a single point in time.[3] In order to smooth out random week-to-week variation in job counts, the Glassdoor Job Market Report shows 4-week moving averages of open online jobs by metro, industry and company size.

D. Comparison to BLS JOLTS

There are many differences in coverage and definitions between counts of open jobs online from the Glassdoor Job Market Report and counts of job openings from BLS JOLTS. They include:

- **Survey vs. Crowdsourced:** The Glassdoor Job Market Report is based on an exhaustive aggregation of every online job posting we're able to find, from tens of thousands of automated crawls of U.S. employer career sites and job openings from hundreds of applicant tracking systems. As of April 2019, we count job listings from 93,000 employers whereas JOLTS figures are based on a nationally representative sample of roughly 16,000 U.S. employers each month.
- **Timing:** JOLTS figures report estimates of job openings as of the last business day of each month. The Glassdoor Job Market Report creates snapshots of unique open jobs online as of each Monday on Glassdoor. In our published report, we show a 4-week moving average of open online jobs by metro, industry and company size.

2. See https://www.bls.gov/news.release/jolts.tn.htm for more details on the BLS JOLTS methodology.

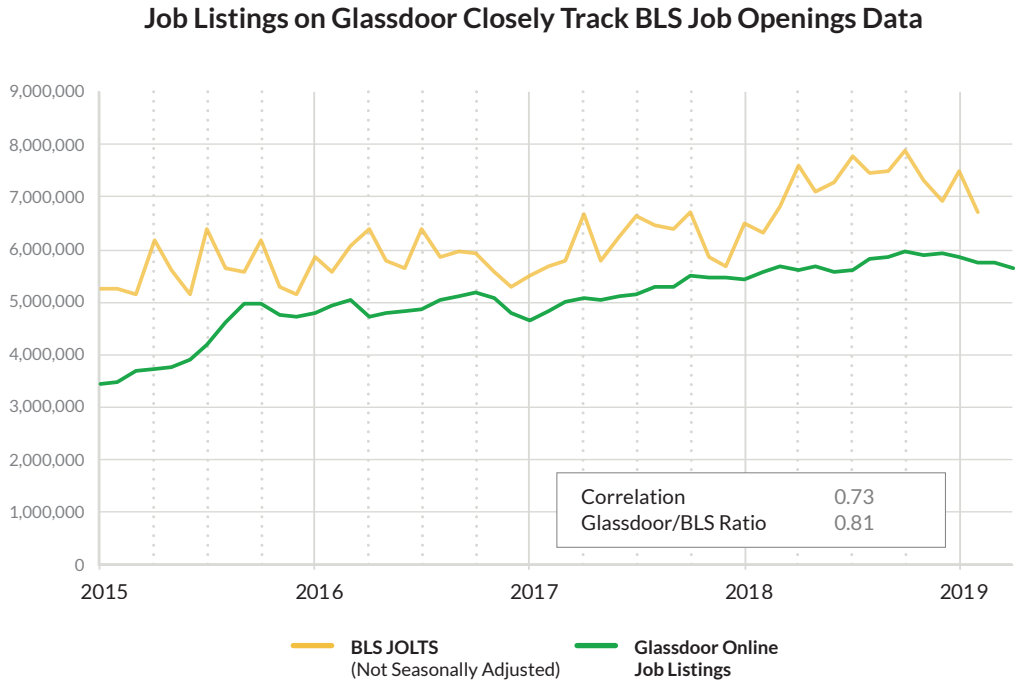
3. Note this "snapshot" definition of open jobs online differs from the definition of online job ads used by the Conference Board's Help Wanted Online ® (HWOL) series, which counts the sum of all job ads posted during the course of each month (some of which may be active or closed at any particular point in time during that monthly period). See https://www.conference-board.org/pdf_free/press/HWOL_Technical_Note_201711.pdf for more details about the methodology behind the HWOL series.

II. Open Jobs *(continued)*

- **Definition of Job Openings:** JOLTS defines a job opening as follows: "1) a specific position exists and there is work available for that position, 2) work could start within 30 days whether or not the employer found a suitable candidate, and 3) the employer is actively recruiting from outside the establishment to fill the position."[4] On Glassdoor, one online job ad may correspond to one open position, many open positions (for example, one new retail location hiring for many cashiers with a single job ad) or no open positions (for example, an ad may be posted online with no intention to actually hire for the role). We rely on the judgment of the HR professionals themselves – who maintain job ads on employer websites and applicant tracking systems – for our measure of hiring intent based on counts of open jobs online.
- **Online vs. Offline Jobs:** The BLS JOLTS collects information from employers about job openings, regardless of whether they are advertised online, offline or not at all. By contrast, the Glassdoor Job Market Report only reports jobs advertised online. Some jobs are not routinely posted online – for example, many union-represented jobs are allocated through local labor unions, and may never appear online – and are not included in Glassdoor totals.

The figure below shows a comparison of BLS JOLTS job openings and open jobs online from Glassdoor. The figures are not directly comparable, but offer a high-level perspective on coverage and trends in Glassdoor data compared to survey-based open job counts from BLS. Note that JOLTS figures are as of the last business day of the month, while Glassdoor figures are monthly averages of Monday snapshots during each month.

Figure 1. Comparison of Glassdoor and BLS JOLTS Job Openings



Source: Glassdoor.com/research; BLS JOLTS (Series: JTU00000000JOL)

4. Source: https://www.bls.gov/news.release/jolts.tn.htm.

II. Open Jobs *(continued)*

Overall, the correlation between Glassdoor job counts and BLS JOLTS job openings (non-seasonally adjusted) is +0.73 from January 2015 through February 2019. The ratio of Glassdoor open jobs online to BLS job openings ranges from 60 percent to 97 percent from January 2015 to February 2019, with an average of 81 percent. Put differently, the open job counts in the Glassdoor Job Market Report likely capture approximately 81 percent of actual U.S. job openings.

A 2014 study from researchers at Georgetown University concluded that, as of that time, roughly 60-70 percent of real world job openings appeared online as a job posting (Carnevale et al. 2014). More recently, Glassdoor job listings were 81 percent of the BLS JOLTS figure, broadly consistent with those findings. As more job listings move online, Glassdoor's job listings coverage of actual U.S. job openings improves as well.

E. Metro and Industry Definitions

The metropolitan ("metro") areas in the Job Market Report are based on "core based statistical areas" or CBSAs defined by the U.S. Census Bureau. These areas contain both "metro" (population of 50,000 or above) and "micro" (population of between 10,000 and 50,000) areas. They include the county of the core urban area, as well as all nearby counties that are closely related as defined by patterns in commuting-to-work data.[5]

The industry groupings on Glassdoor are based on our own internal classification of firms into industries based on similarity of firms from the perspective of job seekers. These industries do not correspond to NAICS or other public-sector industry mappings, and are designed from the perspective of job seekers rather than statistical researchers in federal government agencies who are responsible for designing NAICS and other industry classification systems.

III. Median Base Pay

A. How It Works

Since 2008, Glassdoor has collected millions of reviews and insights, including salary reports, submitted anonymously by current and former employees and spanning hundreds of thousands of employers. These salary reports are provided for specific job titles and industries, at specific employers, for metro areas.

To isolate changes in the composition of Glassdoor salaries over time from underlying changes in U.S. labor market wages, it's necessary to statistically control for possible composition biases in the types of jobs, companies, industries and areas that are represented over time in our data. By statistically holding those changes in composition steady, Glassdoor salary data can allow us to estimate real underlying economic changes in pay over time.

The pay estimates in Glassdoor's Job Market Report employ a proprietary machine-learning model that estimates the separate impact of many job features on salary — job title, company, industry, seniority level, company size, metro location and more. Each of these features is like a "building block" of a salary. We use millions of Glassdoor salaries to estimate the impact of each building block on pay. We then re-assemble those statistical building blocks into localized pay estimates for metro areas, specific job titles, industries, by company size and at different points in time.

5. U.S. Census Bureau CBSA definitions are available at http://www.census.gov/population/metro/data/def.html. A list of cities and counties contained in U.S. Census Bureau CBSAs is available at http://www.census.gov/population/metro/data/def.html. GIS shape files for U.S. Census Bureau CBSAs are available at https://catalog.data.gov/dataset/core-based-statistical-areas-national.

III. Median Base Pay *(continued)*

By using **all** of the salary data on Glassdoor — and not just the data collected during the previous month — the model allows us to estimate pay for more granular job titles and locations than would be possible with more traditional survey approaches, which by contrast simply total up weighted survey responses from the current month.

B. The Salary Estimates Model

The pay estimates in the Glassdoor Job Market Report rely on a statistical model known as an "elastic net regression."[n6] This approach models the relationship between a dependent variable (in this case, salary) and a set of independent variables or "predictors." For each predictor, the model returns "beta coefficients" that quantify the effect of that predictor on salaries, after controlling for the effect of all other predictors in the model. Statisticians refer to this class of models as "generalized linear models" (GLM), which are widely used in statistical and economic research.

Computationally, the elastic net regression produces a set of estimated beta coefficients by solving the minimization problem in Equation (1). For some parameter α strictly between 0 and 1 and some nonnegative parameter λ, it solves:[7]

$$\hat{\beta} = \underset{\beta_0,\beta}{argmin} \frac{1}{2N} \sum_{i=1}^{N} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \left[\frac{(1-\alpha)}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1 \right] \quad (1)$$

where y_i is individual base pay reported to Glassdoor by person i, and x_i is a vector with the set of K predictors in the model, including job title, industry, metro location and more. As is conventional in regressions on salaries, y_i is estimated as the natural log of base salary.

By summing various combinations of beta coefficients from the model, it's possible to produce salary estimates for a given job title, in a given metro location, in a given industry, for a given employer size, at a specific month and year in time. In this way, the beta coefficients from the above model are the basis for the pay estimates in the Glassdoor Job Market Report. In all cases, predicted median base pay from the Job Market Report should be interpreted as pay for **annual salaried workers,** not hourly wage earners.

By estimating Equation (1) using log-transformed base pay, beta coefficients from the model have a "median" interpretation. Summing beta coefficients from the model provide the predicted value of ln(salary). Exponentiating those predicted values yield the predicted **geometric** mean of salary, not the predicted **arithmetic** mean.[8] Since salaries are approximately lognormally distributed, and because under a lognormal distribution the geometric mean is equal to the median, the model outputs should be interpreted as the predicted **median** salary for job titles, industries, company sizes and locations.

C. Training Data

Each month, we re-estimate the above model based on the latest salary data for full-time workers collected by Glassdoor. Initially, the model was estimated in October 2016 based on a sample of 2.31 million approved Glassdoor salaries for full-time workers submitted on or after January 1, 2011. Each month, this model is re-estimated using the latest approved salaries submitted to Glassdoor. The model is estimated for annual **base pay** only, and excludes all other forms of compensation such as tips, commissions, bonuses and equity. Additionally, the sample omits outlier salaries below $15,000 per year and above $600,000 per year.

6. Zou, Hui, and Trevor Hastie (2005). "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society, Vol. 67, No. 2.*

7. $\|\beta\|$ is the Euclidean norm operator such that $\|\beta\| = \beta 1^2 + \ldots + \beta n^2$.

8. For more background on interpreting regression coefficients from regressions with log-transformed variables, see Cornell University Statistical Consulting Unit (June 2012), "Interpreting Coefficients in Regression with Log-Transformed Variables," StatNews #83. Available at https://www.cscu.cornell.edu/news/statnews/stnews83.pdf

III. Median Base Pay *(continued)*

D. Model Accuracy

Table 1 reports the estimated accuracy of the underlying salary estimates model at the time it was initially developed. To assess accuracy, we examined the "median error" of the model at the job title level when tested on samples of known Glassdoor salary data. In those tests, the model displayed a median error of 10.2 percent, which means that half of the estimates from our model were accurate to within 10.2 percent, while the other half had errors above 10.2 percent.

Additionally, we examined what percentage of Glassdoor salary reports the model was able to accurately predict at the job title level within different levels of error. In other words, we applied the model to actual Glassdoor salary reports and tested how often our model was able to accurately predict those individuals' pay. At the time the model was developed, it predicted 75.3 percent of Glassdoor salaries accurately to within 20 percent of the actual reported base pay, 49.1 percent of salaries to within 10 percent, and 27.2 percent of salaries to within 5 percent.

Table 1. *Estimated Accuracy of Salary Estimates at the Job Title Level*

Model Performance	Within 5%	Within 10%	Within 20%	Median Error
	0.272	0.491	0.753	0.102

Source: Glassdoor Economic Research (glassdoor.com/research)

E. Comparison to BLS Median Pay

As a benchmark of the accuracy of the Job Market Report, we compared estimates for median U.S. base pay from our model to official U.S. median wage and salary estimates from the U.S. Bureau of Labor Statistics. For the comparison, we used figures for "median weekly earnings of full-time wage and salary workers" from the U.S Bureau of Labor Statistics (BLS),[9] which relies on a survey of roughly 60,000 U.S. households per month to estimate pay.

For this comparison, it is important to note that there are large differences in coverage and definitions between Glassdoor and BLS data. Glassdoor's Job Market Report are more current than what BLS offers, and are updated once a month giving a near real-time view of wages at the local level. Additionally, Glassdoor's data is specific to actual job titles, offering more insight than that provided by broad BLS occupational groupings. Another key difference is in the definition of wages and salaries: BLS's definition of median wages and salaries includes tips, commissions and other cash bonuses that workers report usually receiving.[10] By contrast, the Glassdoor Job Market Report reflect only annual base pay and not these other forms of compensation. By benchmarking against BLS data, we hope to provide an accurate picture of how our data compare to what's been considered the norm in terms of measuring U.S. pay growth.

9. Average weekly earnings are converted into annual pay estimates assuming a 52-week work year for full-time workers. Source: U.S. Bureau of Labor Statistics, "Labor Force Statistics from the Current Population Survey: Table 39. Median weekly earnings of full-time wage and salary workers by detailed occupation and sex." Available at http://www.bls.gov/cps/cpsaat39.htm.

10. Source: U.S. Bureau of Labor Statistics, "Usual Weekly Earnings Technical Note," available at http://www.bls.gov/news.release/wkyeng.tn.htm.

III. Median Base Pay *(continued)*

Table 2 shows the comparison between the Glassdoor Job Market Report estimates of median U.S. base pay and BLS estimates of U.S. median wages and salaries for full-time workers for 2014 and 2015. The published Glassdoor Job Market Reports produce estimates of median base pay for **annual salaried** workers only, not hourly wage workers. For this comparison to BLS data, however, we removed that restriction and report median base pay for **all** full-time workers — hourly wage and annual salaried — to make the data series more comparable.

In 2015, the Local Pay Reports (now known as the Job Market Report) predicted a median base pay for U.S. full-time workers of $44,282, compared to the BLS estimate of $42,068, a difference of 5.3 percent. In 2014, the Local Pay Reports predicted a median base pay of $43,808, compared with the BLS estimate of $41,132, a difference of 6.5 percent. In both years, the Local Pay Reports predicted median U.S. pay to within 7 percent of BLS estimates based on large, nationally representative samples of workers.

Table 2. *Comparison of Median U.S. Pay from BLS Estimates and Glassdoor's "Job Market Report" Model*

Year	2015	2014
Glassdoor Local Pay Report Estimate (U.S. Median Base Pay, Full Time Workers)	$44,282	$42,068
BLS Estimate (U.S. Median Wages and Salaries, Full Time Workers)	$42,068	$41,132
Percentage Error	5.30%	6.50%

Source: : U.S. Bureau of Labor Statistics, "Labor Force Statistics from the Current Population Survey: Table 39. Median weekly earnings of full-time wage and salary workers by detailed occupation and sex." Available at http://www.bls.gov/cps/cpsaat39.htm.

F. Comparison to Other Wage Growth Figures

As a second accuracy benchmark, we compared year-over-year wage growth from the Glassdoor Job Market Report over the past several years with two widely used official measures of wage growth: The Employment Cost Index for wages and salaries from the U.S. Bureau of Labor Statistics, and the Wage Growth Tracker from the Federal Reserve Bank of Atlanta.[11]

11. Federal Reserve Bank of Atlanta Wage Growth Tracker is available at https://frbatlanta.org/chcs/wage-growth-tracker/?panel=1).
U.S. Bureau of Labor Statistics Employment Cost Index for wages and salaries is available at http://www.bls.gov/web/eci/echistrynaics.txt.

III. Median Base Pay *(continued)*

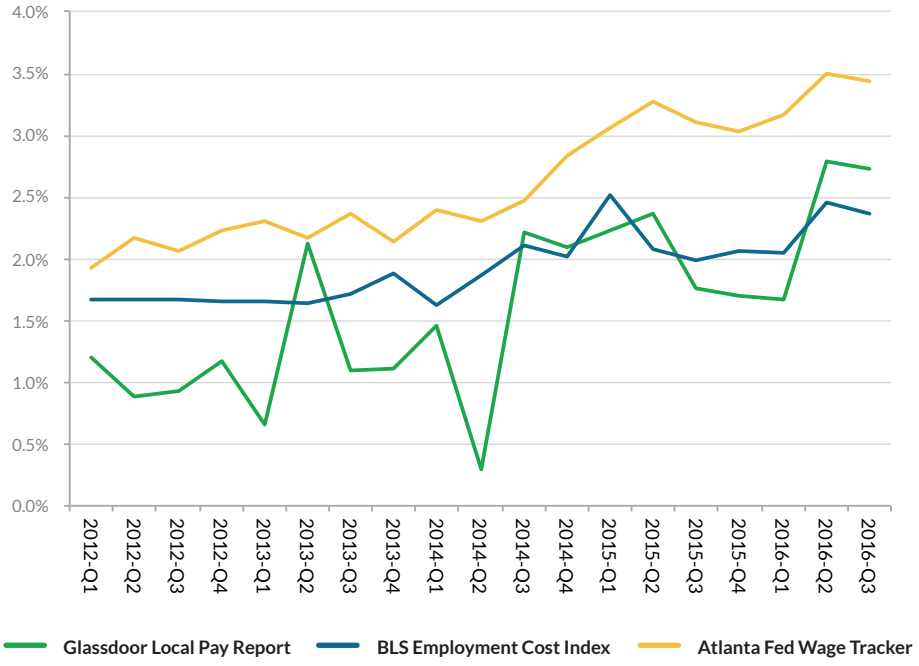
The Employment Cost Index is based on a quarterly survey from BLS of roughly 6,800 business establishments throughout the U.S. The survey asks employers about labor costs in their companies, and the wage and salary component is a widely used measure of pay growth by economists. The Wage Growth Tracker is an analytic data product from the research team at the Atlanta Federal Reserve. It is based on survey data from the U.S. Census Bureau's Current Population Survey, and tracks median hourly pay for workers who've been continuously employed over the past 12 months.

Figure 2 shows the comparison of year-over-year wage growth in quarterly average wages from the three sources since 2012. Pay estimates from the Glassdoor Job Market Report are shown in green, the BLS Employment Cost Index for wages and salaries is shown in blue, while the Atlanta Fed's Wage Growth Tracker is shown in red.

Note that there are important differences in definitions and coverage between the three sources that warrant caution in interpreting this comparison. For example, the Atlanta Fed's Wage Growth Tracker follows only workers who have remained continuously employed over a 12-month period. These workers tend to be older and more educated than the general labor force, which explains the generally higher wage growth in that measure.

Despite many methodological differences, in general estimates of U.S. median base pay from Glassdoor's Job Market Report tracks other measures of wage growth quite closely, particularly in recent years. Table 3 shows the raw correlation between Job Market Report and these two measures, which is positive and large in both cases: 0.75 correlation with the Atlanta Fed Wage Growth Tracker, and 0.74 correlation with the BLS Employment Cost Index.

Figure 2. Comparison of LPR Wage Growth vs. Atlanta Fed Wage Tracker and BLS ECI for Wages and Salaries



Source: Glassdoor Economic Research; Federal Reserve Bank of Atlanta Wage Growth Tracker (available at https://frbatlanta.org/chcs/wage-growth-tracker/?panel=1); U.S. Bureau of Labor Statistics Employment Cost Index, Wages and Salaries (available at http://www.bls.gov/web/eci/echistrynaics.txt).

III. Median Base Pay *(continued)*

Table 3. *Correlation of Job Market Report Wage Growth with Government Figures*

Core-Based Statistical Areas

	Correlation	R-Squared
Atlanta Fed Wage Tracker	0.75	59.8%
BLS Employment Cost Index	0.74	56.7%

Source: Glassdoor Economic Research.

Table 3 also shows the R-squared value from a simple linear regression of each official measure of wage growth against the estimated wage growth from the Glassdoor Job Market Report. This can be interpreted as the percentage of variation in official wage growth measures that is explained by Glassdoor salary data. The R-squared value is 59.8 percent compared to the Atlanta Fed measure, and 56.7 percent compared to the BLS measure. In both cases, wage growth from the Glassdoor Job Market Report explains more than half the variation in these two official wage growth measures.

For More Information

For more information about the Glassdoor Job Market Report, please visit our "frequently asked questions" page at https://www.glassdoor.com/research/studies/glassdoor-job-market-report-faq/ or contact pr@glassdoor.com.

References

Carnevale, Anthony P., Tamara Jayasundera, and Dmitri Repnikov (2014). "Understanding Online Job Ads Data: A Technical Report," Georgetown Center on Education and the Workforce. Available at http://www.workforcedqc.org/sites/default/files/images/Georgetown%20U%20Real-time%20LMI%20Technical%20Report.pdf.

Friedman, Jerome H., Trevor Hastie, and Rob Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent," Journal of Statistical Software, Vol. 33, No. 1. Available at www.jstatsoft.org/article/view/v033i01.

Zou, Hui, and Trevor Hastie (2005). "Regularization and variable selection via the elastic net," Journal of the Royal Statistical Society, Vol. 67, No. 2. Available at https://web.stanford.edu/~hastie/Papers/B67.2%20(2005)%20301-320%20Zou%20&%20Hastie.pdf.