



December 2021

Final Project: Amazon's Human Resources



Company & Industry

Amazon is an American technology company based in Seattle, which operates in online retailer, cloud computing, digital streaming, and artificial intelligence industries [1]. It is most recognized as the world's largest e-commerce platform. Recently, Amazon made \$3.2 billion in profit, about half as much as last year. To keep up with its growth, Amazon has hired 133,000 new employees, bringing its workforce to 1.47 million, up 30 percent from a year ago. In September, it has announced plans to hire 125,000 hourly workers and 150,000 seasonal workers in the United States in advance of the holiday shopping rush [2].

One factor that drives the company's growth is undeniably related to human resources management. Specifically, employee retention is one of the competitive edges of a successful business. While employee turnover can lead to some notable drawbacks, retaining the workforce comes with a number of advantages such as saving recruiting costs, maintaining a solid market knowledge base, as well as bettering customer service and product quality. These factors are overall help increase the company's efficiency and play an essential part in leading to higher profits. Therefore, one possible end goal of Amazon is to meet its long-term human resources strategic objective of bettering employment longevity.

Analysis Goal

It is commonly believed that one possible thing that impacts employment longevity is employee's monetary compensation. Assuming that the corporation is interested in the sensitivity of its employees' salary to employment longevity, this report contains an analysis for Amazon's human resources, focusing on the causal relationship between salary changes and longevity. This analysis is based on observational data that contains information about average employee longevity, average salary, and other variables for multiple groups of employees of Amazon at one point in time.

The Effect of Salary Changes on Longevity

To find the effect of salary changes on employment longevity, the estimation that comes from the simple regression model of average salary and average longevity is not a good estimate of the causal effect of salary changes on employment longevity in the population. This happens because salary is a strategic variable that does not

[1] "<https://fortune.com/company/amazon-com/>". Fortune. Oct 12, 2021.

<https://www.businessinfocusmagazine.com/2014/04/chargepoint/>

[2] "Amazon sales growth slows and costs rise". Sept 2021. <https://www.nytimes.com/2021/10/28/business/amazon-stock.html>.



have the feature of a randomly assigned treatment. What this means is the treatment (higher salary) probably was given to observations (employees) where the effect was greater. For example, giving a higher salary to older employees might make them stay at the company longer; different groups of employees have a different response to the treatment of higher or lower salaries.

This causes selection bias to exceed zero, which means employment longevity is not about the salary alone. This concept connects to the fact that the model is endogenous. Endogenous means that there are other possible omitted factors outside the model that correlates with salary changes and directly affect longevity. These factors are confounding factors that mess up an ability to estimate the causal effect of the X variable (salary changes) on Y variable (employment longevity). With this, it is hard to say that a higher salary causes longer employment longevity.

Confounding Factors

To find potential confounding factors which arise because of different entities, the first step is finding the differences between a group of high salary workforce (treated group) and a group of low salary workforce (untreated group) at Amazon. The possible differences include position, age, education years, location of work, possibly race, and gender.

Different Factors	Correlated with Salary Changes	Directly Impacts Longevity
Position	✓	✓
Age	✓	✓
Education Years	✓	✓
Location of work	✓	✓
Race	✓	✗
Gender	✓	✗

The factors mentioned above are automatically correlated with salary, given that they are the reasons why two groups are different. However, factors such as race and gender do not directly impact longevity. With this, the major confounding factors are position, age, education years, and location of work. The list of confounding factors also confirm that there is an endogeneity issue posed.

Tools to Reduce Endogeneity Issue

What is needed to estimate a causal relationship is an exogenous model, meaning that a model should not have any endogenous X variables. Because the data is cross-sectional where there is no variation in time, attacking an endogeneity problem by adding control variables can reduce the endogeneity issues. A control variable is any independent variable that has been confounding factors, but can be added to a regression model to reduce an endogeneity problem.

Adding Control Variables

The first model is a regression of average salary on employment longevity, controlling for age, education years, and location of work, utilizing the code reg on Stata. The control variables include average age, average education years, unemployment rate, and dummy variable coast, which is a binary variable where 1 means the office location is near to coast and 0 means the office location is not near to the coast. Locations that are near to the the coast and have low unemployment rate tend to have stronger labor market and more job availability that may refrain employees from retaining at Amazon for a long period.

. reg AvgLongevity AvgSalary000s AvgAge AvgEducationYrs UnemploymentRate Coast						
Source	SS	df	MS	Number of obs	=	454
Model	1988.21675	5	397.643349	F(5, 448)	=	5837.96
Residual	30.5148317	448	.068113464	Prob > F	=	0.0000
Total	2018.73158	453	4.4563611	R-squared	=	0.9849
				Adj R-squared	=	0.9847
				Root MSE	=	.26099

AvgLongevity	Coefficient	Std. err.	t	P> t	[95% conf. interval]
AvgSalary000s	.0264189	.0068767	3.84	0.000	.0129044 .0399334
AvgAge	.1208882	.0007816	154.66	0.000	.1193521 .1224243
AvgEducationYrs	.062198	.0686991	0.91	0.366	-.0728144 .1972105
UnemploymentRate	.0273099	.2084573	0.13	0.896	-.3823657 .4369855
Coast	-.2331361	.0245753	-9.49	0.000	-.2814333 -.1848389
_cons	-1.696652	.0941537	-18.02	0.000	-1.88169 -1.511614

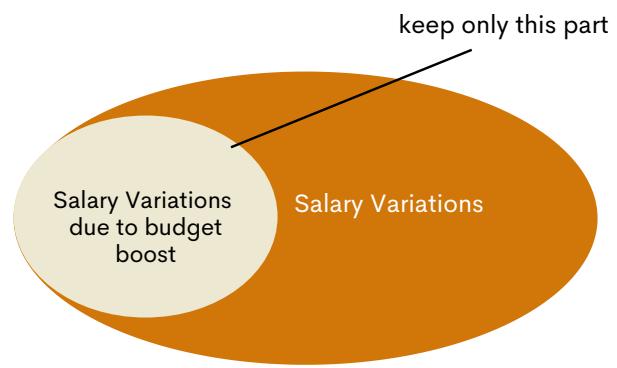
The result of the regression indicates that by holding age, education years, unemployment rate, and coast constant, the coefficient of average longevity indicates that every \$1000 increase in average salary is correlated with 0.026 years increase in employment longevity of employees at Amazon. The t-statistics shows that the observation of 0.026 is 3.85 standard errors above the null hypothesis of the salary having no effect. The p-value is equal to 0, which means that it is statistically significant, so we do reject the null hypothesis of salary has no effect. The insignificant variables are average years of education and unemployment rate. Yet, this method likely does not entirely solve the endogeneity problem because of the unavailability of data which restricts us from including every confounding factors into the model.

Tools to Eliminate Endogeneity Issue

Instrumental Variables

At this point, it is reasonable to believe that X variables is correlated with U or other possible omitted factors outside the model that correlates with salary changes and directly affect longevity. To replicate random assignment and get rid of endogeneity to make model exogenous, utilizing instrumental variable is an appropriate method. Instrumental variable is a variable that varies with salary changes, but does not impact employment longevity.

Suppose budget boost does not impact employment longevity for Amazon, meaning two locations with the same salary, age, and geographic have the same expected longevity, even if their budget boost differ. In the other word, the observed group with budget boost is not special.



To expand, budget boost is a valid instrument for salary because:

- 1) It is exogenous; budget boost does not correlate with anything in U or other possible omitted factors outside the model that correlates with salary changes and directly affect longevity. Budget boost does not help predict employment longevity.
- 2) It is relevant; budget boost correlates with salary changes, but do not directly impact employment longevity. The correlation between salary and budget boost is 0.18. Although the correlation might be weak, it is still useful for the regression since it would preserve more of the good part. Additionally, in a business perspective, this would make sense because budget boost is the available budget which is likely to be passed on to employees through salary.

The second model is a two-staged least squares regression that utilized the code ivreg on Stata. 2SLS can control more than 1 confounding factors, keep good variations, and remove bad variations. Because 2SLS can leave behind some good parts, including significant variables of average age and coast as control variables into the model can retain some good variations. For this regression, supposed average age and coast are exogenous, meaning that those variables are not correlated with other possible omitted factors outside the model.

```
. ivreg AvgLongevity AvgAge Coast ( AvgSalary000s = BudgetBoost)
```

Instrumental variables 2SLS regression

Source	SS	df	MS	Number of obs	=	454
Model	1987.27584	3	662.42528	F(3, 450)	=	8171.27
Residual	31.4557388	450	.069901642	Prob > F	=	0.0000
Total	2018.73158	453	4.4563611	R-squared	=	0.9844
				Adj R-squared	=	0.9843
				Root MSE	=	.26439

AvgLongevity	Coefficient	Std. err.	t	P> t	[95% conf. interval]
AvgSalary000s	.0343215	.0012192	28.15	0.000	.0319254 .0367175
AvgAge	.120877	.0007915	152.72	0.000	.1193216 .1224325
Coast	-.2303522	.0248767	-9.26	0.000	-.2792412 -.1814632
_cons	-1.798942	.126816	-14.19	0.000	-2.048167 -1.549716

Instrumented: AvgSalary000s

Instruments: AvgAge Coast BudgetBoost

From the regression, the coefficient of average salary indicates that \$1000 increase in average salary will result in 0.034 years increase in employment longevity of employees at Amazon. The t-statistics tells us that the observation of 0.034 is 28.15 standard errors above the null hypothesis of the salary having no effect. We can infer the causal relationship because using the budget boost as an instrumental variable, the model is exogenous. All coefficients are significant because the p-values are less than 0.05.

Recommendation

Based on empirical findings of positive causal relationship, salary changes is explicitly link to employment longevity. To meet its long-term human resources strategic objective of increasing employment longevity, Amazon should increase salary to employees. The indirect cost like time and direct financial costs associated with losing employees along with hiring and training replacements for an enterprise like Amazon, given its size and specialization, can be significantly high.

If Amazon fail to offer competitive monetary compensation, Amazon may be at risk of a higher turnover rate. A high turnover rate can negatively impact workplace morale as well as company's reputation, which may ultimately decrease customer loyalty. By increasing salary, Amazon will be able to retain its valued and marketable employees, decrease turnover, and increase employment longevity.