How Do Means of Transportation to Work Affect Labor Market

Jiongliang Tao

Apr 14, 2019

Abstract

This paper will review several different factors that contribute to wage differences, and this paper will be mainly focused on how different transportations affect people's wage in labor market. Using data from American Community Survey (ACS), I performed a linear regression with natural log of wage income as the dependent variable and transportation to work, usual hours worked per week, educational attainment, age, and sex as the independent variable. Only people who are employed are focused in this paper. This paper found that the wages of people with bus, bicycle, or walked only as their transportations to work are lower than the wages of people with auto, truck or van at a statistically significant level. This paper also found that the wages of people with streetcar or trolley, subway or elevated, railroad, taxicab, and ferryboat are higher than the wages of people with auto, truck or van at a statistically significant level. Using twitter API, I found that people have positive sentiment to public transportation. The findings were contrary to expectations because people are more likely to think transportations have no effects on people's wages. This might not be a reasonable result, but this result does exist in labor market and more research may be done on this topic.

Introduction

When we talk about wage gaps, people are more likely to think about race discrimination and gender discrimination. Race discrimination and gender discrimination are becoming more and more popular topics in recent years in the United States. However, almost no people would think about that transportations could also affect people's wages. Wage gaps could also result from transportations that people choose. Since transportation has a lot of effects on environments, employments, and economy, policy makers would better pay more attention on it. By using empirical study, I will test whether transportation can affect the wage or not in this paper.

Literature Review

Previous work by Weisbrod and Reno (2009) has shown that public transportation can generate jobs, shape land use, support public policies, enable economic growth, and provide mobility. Public transportations have influences on economy such as business income, wages, and employment. The key findings in Weisbrod and Reno (2009) include public transportation has created a lot of jobs and income by supporting public transportation operation activities, construction and manufacturing; public transportation enables various economic productivity (changes in travel costs, access factors, and time) and efficiency.

The reasons that public transportation can affect economy include: public transportation can reduce traffic congestion for people who are using automobile and truck as their major transportation; vehicle ownership and travel cost saving for people who are switching from automobiles to public transportation passengers; business operating cost saving; reduced traffic congestions enable business productivity from having more diverse-skill labors from broader labor markets; business growth due to the indirect effects of business growth on supplies (Weisbrod and Reno, 2009).

Haynes (1997) states that transportation can affect labor markets in different levels. For example, transportation affects micro level search behavior of job seekers and meso level trade-offs between labor migration and commuting on the supply side. Demand labor can have broader labor market to select from with more specialized or targeted fit between employees and jobs on the demand side (Haynes, 1997).

Hirota (1984) has shown an example of Shinkansen (Japan), which is a kind of public transportation in Japan, and Hirota (1984) presented that employment growth in wholesaling, industry, retail, and construction was 16 to 34 percent higher in station than non-station locations. This result corroborated to the result that presented by Amano and Nakagawa (1990). Amano and Nakagawa (1990) got a result that employment growth level was 26 percent higher in station than non-station locations.

Data

The primary data source for this paper is from American Community Survey (ACS). The data set used in this paper includes ACS data from 2015 to 2017. SEX, AGE, EDUC, EMPSTAT, UHRSWORK, INCWAGE, and TRANWORK are the main variables included in this data set. It is obvious that male and female are included in SEX. Educational attainment are represented by EDUC. In EDUC, N/A or no schooling, Nursery school to grade 4, Grade 5, 6, 7, or 8, Grade 9, Grade 10, Grade 11, Grade 12, 1 year of college, 2 years of college, 3 years of college,4 years of college, and 5+ years of college have been included in this variable. EMPSTAT represents employment status, and this variable includes N/A, Employed, Unemployed, and Not in labor force. UHRSWORK means usual hours worked per week, and INCWAGE shows wage and salary income. This paper will be mainly focused on TRANWORK, which is the means of transportation to work. TRANWORK has mainly contained private motorized vehicle such as auto, truck, or van, and motorcycle; public

transport such as bus or trolley bus, streetcar or trolley car, subway or elevated, railroad, taxicab, and ferryboat; other way to work such as bicycle, walked only, and worked at home. TRANTIME means travel time to work, and the unit for it is minute. MET2013 represent Metropolitan area 2013 in the United States(Ruggles et al., 2019).

Top 12 transportations to work have been shown in Figure 1 by frequency, and they are auto, truck, or van, worked at home, walked only, subway or elevated, bus or trolley bus, railroad, bicycle, other, motorcycle, taxicab, streetcar or trolley car, ferryboat. Missing values have been dropped in the data set. people who are employed, wage income greater than zero, educational attainment greater than 1 years of college, and all transportation to work are focused in this paper. Because sex, educational attainment, Metropolitan area 2013, and transportation to work are dummy variables, they are treated as factors in my model. Variables include SEX, AGE, EDUC, EMPSTAT, UHRSWORK, INCWAGE, TRANWORK, TRANTIME, and Metropolitan area 2013 are kept in my data set.

The measurement of statistics and frequency have been shown in Table 0, and this table gives a summary of data that used in my paper. Frequencies of variables include SEX, EDUC, TRANWORK, and Metropolitan area 2013 are shown as the numbers on the right side of that variable. Since I only focused on educational attainment above 1 year of college, I have dropped the observations below 1 year of college. This is the reason that the frequencies of some variables are zero. The statistical descriptions include minimum, first quantile, median, mean, third quantile, and maximum are also shown in this summary table. We can see that minimum, median, mean, and maximum value of usual hours work per week are 1, 40, 40.17, and 99.

Empirical Methods

The Empirical Method that used in this paper is linear regression. Linear regression is a good model that can easily estimate the correlations between different variables. The first linear regression model that I formulated include motorcycle, Bus or trolley bus, streetcar or trolley car, subway or elevated, railroad, taxicab, ferryboat, bicycle, walked only, other, worked at home, usual hours worked per week, 2 years of college, 4 years of college, 5+ years of college, age, age square and female as independent variables, and natural logarithm of the wage as dependent variable. If dummy variables such as sex, educational attainment, and transportation to work equal to one, it means people are in that classification and zero if otherwise. The first linear regression model that I have performed is showing in the following equation:

$$\begin{aligned}\log(\text{Incwage}) =& \beta_0 + \beta_1 \cdot \text{Motorcycle} + \beta_2 \cdot \text{Bus or Trolley bus} + \beta_3 \cdot \text{Streetcar or Trolley car} + \\ & \beta_4 \cdot \text{Subway or Elevated} + \beta_5 \cdot \text{Railroad} + \beta_6 \cdot \text{Taxicab} + \beta_7 \cdot \text{Ferryboat} + \\ & \beta_8 \cdot \text{Bicycle} + \beta_9 \cdot \text{Walked only} + \beta_{10} \cdot \text{Other} + \beta_{11} \cdot \text{Worked at home} + \\ & \beta_{12} \cdot \text{Usual hours worked per week} + \beta_{13} \cdot \text{2 years of college} + \beta_{14} \cdot \text{4 years of college} + \\ & \beta_{15} \cdot \text{5+ years of college} + \beta_{16} \cdot \text{Age} + \beta_{17} \cdot \text{Age}^2 + \beta_{18} \cdot \text{Female} + u \end{aligned} \tag{1}$$

The second linear regression model that I formulated includes two additional independent variables, and the two additional independent variables are travel time to work and Metropolitan area 2013. The second linear regression model that I have performed is showing in the following equation:

$$
\begin{aligned}
\log(\text{Incwage}) = &\beta_0 + \beta_1 \cdot \text{Motorcycle} + \beta_2 \cdot \text{Bus or Trolley bus} + \beta_3 \cdot \text{Streetcar or Trolley car} + \\
&\beta_4 \cdot \text{Subway or Elevated} + \beta_5 \cdot \text{Railroad} + \beta_6 \cdot \text{Taxicab} + \beta_7 \cdot \text{Ferryboat} + \\
&\beta_8 \cdot \text{Bicycle} + \beta_9 \cdot \text{Walked only} + \beta_{10} \cdot \text{Other} + \beta_{11} \cdot \text{Worked at home} + \\
&\beta_{12} \cdot \text{Usual hours worked per week} + \beta_{13} \cdot \text{2 years of college} + \beta_{14} \cdot \text{4 years of college} + \\
&\beta_{15} \cdot \text{5+ years of college} + \beta_{16} \cdot \text{Age} + \beta_{17} \cdot \text{Age}^2 + \beta_{18} \cdot \text{Female} + \beta_{19} \cdot \text{TRANTIME} + \\
&\beta_{20} \cdot \text{Metropolitan area 2013} + u
\end{aligned}
\tag{2}
$$

β_1 through β_{11} are the parameters of interest. The reason that natural logarithm of the wage is used in this model is because I would like to see wage income changes in percentage rather than changes in dollars. This is helpful because I can easily see percentage changes in wage corresponding to changes in independent variables.

The core research question in my paper is whether transportations to work have influence on wage or not. Usual hours worked per week is one of the important factors contribute to wage. It is obvious that people who worked more would get higher wage. This is why a lot of people choose to have more than one job. Educational attainment has been considered in my model, because most people would think that the productivity of more educated people is higher than less educated people. Age is also an important factor that could affect people's wage. Most people would think that older people should have higher wages than younger people because older people have more experience than younger people. It is obvious that gender discrimination does exist in labor market because the wage of female is lower than male in average. So, gender effect has been considered in this research.

There are some dummy variables that have been excluded in this model because they are the basic variables that were compared with other dummy variables. The sign of these dummy

variables were affected by these excluded dummy. For transportation, all transportations are compared to the auto, truck, or van, and there are no expectations for the coefficients of transportations since the effects of transportations on wage really depend. For usual hours worked per week, it is obvious that the expectation of coefficient would be positive because people tend to think that spending more time on working will give them more return on wage. All dummy variables of education attainment are compared to 1 years of college, and the expectations for these dummy variables are positive since education has positive effects on wage. The expected sign on coefficients of age would be positive because older people have more experience than younger people. The expected sign on coefficients of female would be negative since people tend to think men are more productive than women.

There are many other factors that contribute to wage difference, and I could not contain all the variables that may affect wage in the model. So, omitted variable bias might exist in my model. Experience is one of the important factors that could affect wage, but experience is not something that can be easily measured in data. Therefore, experience was excluded from this model. However, experience would have positive effects on wage. In this case, the coefficient of variable in my model would be smaller because of the upward bias.

Because the p-value of white test and Breusch-Pagan test are smaller than 0.05, the null hypothesis of homoskedasticity was rejected. In this situation, the number five Gauss-Markov assumption was violated, and the heteroskedasticity does exist in my model. Since heteroskedasticity is existed in my model, the usual t-statistic, f-statistic, and standard errors are no longer valid in my model. Therefore, heteroskedasticity-robust inference has been done in this research in order to make t-statistic, f-statistic, and standard errors valid.

Research Findings

Table 1 and table 2 report the results of first linear regression model. The results of usual OLS were reported in table 1 while the results of robust OLS were reported in table 2. Heteroskedasticity-robust confidence intervals, heteroskedasticity-robust t statistic, and heteroskedasticity- robust standard errors were included in table 2. The value of t-statistic, f-statistic, and standard errors are valid in table 2, and the regular standard errors are slightly smaller than heteroskedasticity- robust standard errors.

The coefficient of motorcycle is -0.004, and the standard error is small for this variable. However, the t-statistic for motorcycle is very small, and the p-values for it is larger than 0.1. So, the null hypothesis that motorcycle has no effects on people's wage cannot be rejected in this case. We cannot make any conclusions from this coefficient. As we see in figure 1, the counts of motorcycle are very small. The reason that we got this statistically insignificant results might because the observations of this variables are very limited.

The estimate of Bus or trolley bus, Bicycle, and Walked only are -0.072, -0.099, and -0.304. Those number means the wages of people who take Bus or trolley bus, Bicycle, and Walked only as their transportation to work would be 7%, 9.9%, 30.4% lower than the wage of people who take auto, truck, or van as their transportations to work (holding all else constant). The t-statistics and p-values are large enough and small enough to reject the null hypothesis that taking Bus or trolley bus, Bicycle, and Walked only as people's transportations to work have no effects on their wage. So, the results we got from this research are statistically significant.

The reason that people who take Bus or trolley bus, Bicycle, and Walked only as their transportation to work would have low wage might because those people are poor. It is obvious that poor people would get lower wage than rich people since people tend to think the productivity of poor people is less than rich people. Poor people don't have enough money to buy a car, and they would spend more on traveling to work place, which may lower

their productivity. Also, bicycle and walked only are not efficient ways to work since they will spend more time than other transportations.

The coefficients of streetcar or trolley, subway or elevated, railroad, taxicab, and ferryboat are 0.048, 0.207, 0.405, 0.034, and 0.400. From these numbers we know that the wage of people who go to work by Subway or elevated, railroad, taxicab, and ferryboat as their transportations will be 4.8%, 20.7%, 40.5%, 3.4%, and 40% higher than the wage of people who go to work by auto, truck, or van as their transportations when controlling all else constant. The results are statistically significant since the t-statistics and p-values are large and low enough to reject the null hypothesis at a significant level of 5 percent.

The reason that people who go to work with streetcar or trolley, subway or elevated, railroad, taxicab, and ferryboat as their transportations would have high wage is because those people would have high productivity. One of the effects of public transportation on economy is that public transportation can reduce traffic congestions for people who are using automobile and truck as their major transportation (Weisbrod and Reno, 2009). Traffic congestion and limited parking lots for people who are using automobile and truck as their major transportations are serious issues in big city. Both traffic congestion and limited parking lots will have negative effects on their sentiment and productivity.

In order to know which direction of people's sentiment will be affected by public transportation to work, I use twitter API to grab data of the words that used by people about public transportation to work. I have transited the words used by people in twitter to words that represent sentiment. From figure 2, we can see top two words that people use about public transportation to work are positive and anticipation. It is obvious that public transportations to work such as streetcar or trolley, subway or elevated, railroad, taxicab, and ferryboat have positive effects on people's sentiment, and those positive sentiment will have positive effects on people's productivity. There is no doubt that the productivity of people with positive sentiment is higher than the productivity of people with negative sentiment. So, this is why

people who go to work with streetcar or trolley, subway or elevated, railroad, taxicab, and ferryboat as their transportations would have higher wage than people who take auto, truck, or van as their transportations to work.

The result of usual hours worked per week shows that one more hour worked per week would increase people's wage by 4.1 percent, and this result is statistically significant at significant level of 1 percent. This result is consistent with my expectation that more hour worked would have positive effects on the wage.

The estimates of educational attainment showed that the wages of people with 2 years of college, 4 years of college, 5+ years of college are 12.3%, 39.8%, 62.7% higher than the wage of people with 1 years of college. This result is reasonable, and it is consistent to my expectation that more educated people would have higher wages.

Just as what we expected, the result of age shows that one more year older would increase people's wage by 10.3 percent. There is no doubt that older people would have more experience than younger people, and age will have positive effects on their wage. This is the reason why most of employers would like to hire older people than younger people.

From the result of female, we can see that gender discrimination is actually existed in labor market. The wage of female is 22.8 percent lower than the wage of male. The reason that we got this answer is because of gender discrimination. Most employer would think that male should have higher productivity than female because the productivity of female will be dropped by Maternity leave and some other family reasons.

Table 3 reports the results of second linear regression model. From the results of second linear regression model, the estimate of Bus or trolley bus, Streetcar or trolley car, Bicycle, and Walked only are -0.205, -0.083, -0.136 and -0.286. Those number means the wages of people who take Bus or trolley bus, Streetcar or trolley car, Bicycle, and Walked only as their transportation to work would be 20.5%, 8.3%, 13.6%, and 28.6% lower than the wage of people who take auto, truck, or van as their transportations to work (holding all else

constant). The results are statistically significant.

The coefficients of subway or elevated, railroad, and ferryboat are 0.013, 0.185, and 0.235. From these numbers we know that the wage of people who go to work by subway or elevated, railroad, and ferryboat as their transportations will be 1.3%, 18.5%, and 23.5% higher than the wage of people who go to work by auto, truck, or van as their transportations when controlling all else constant. Also, the results are statistically significant.

Conclusion

From the research above, I found that the wages of people who take Bus or trolley bus, Streetcar or trolley car, Bicycle, and Walked only as their transportation to work would be 20.5%, 8.3%, 13.6%, and 28.6% lower than the wage of people who take auto, truck, or van as their transportations to work when controlling all else constant. The reason for it may be those transportation is not efficient enough for people to go to work. Spending more time on those types of transportation would have negative effects on their productivity. Because of the omitted variable bias, the actual effects of them would be lower.

I also found that the wage of people who go to work by subway or elevated, railroad, and ferryboat as their transportations will be 1.3%, 18.5%, and 23.5% higher than the wage of people who go to work by auto, truck, or van as their transportations when controlling all else constant. Traffic congestion and limited parking lots became more and more serious issues for people who are using automobile and truck as their major transportation. Those types of issues can let people waste time on road and have negative effects on their sentiment and productivity. Using twitter API, I found that most of people have positive sentiment to public transportation. So, public transportation brings people positive sentiment, and positive sentiment could increase people's productivity, which will finally increase people's wage.

Figures and Tables

Figure 1

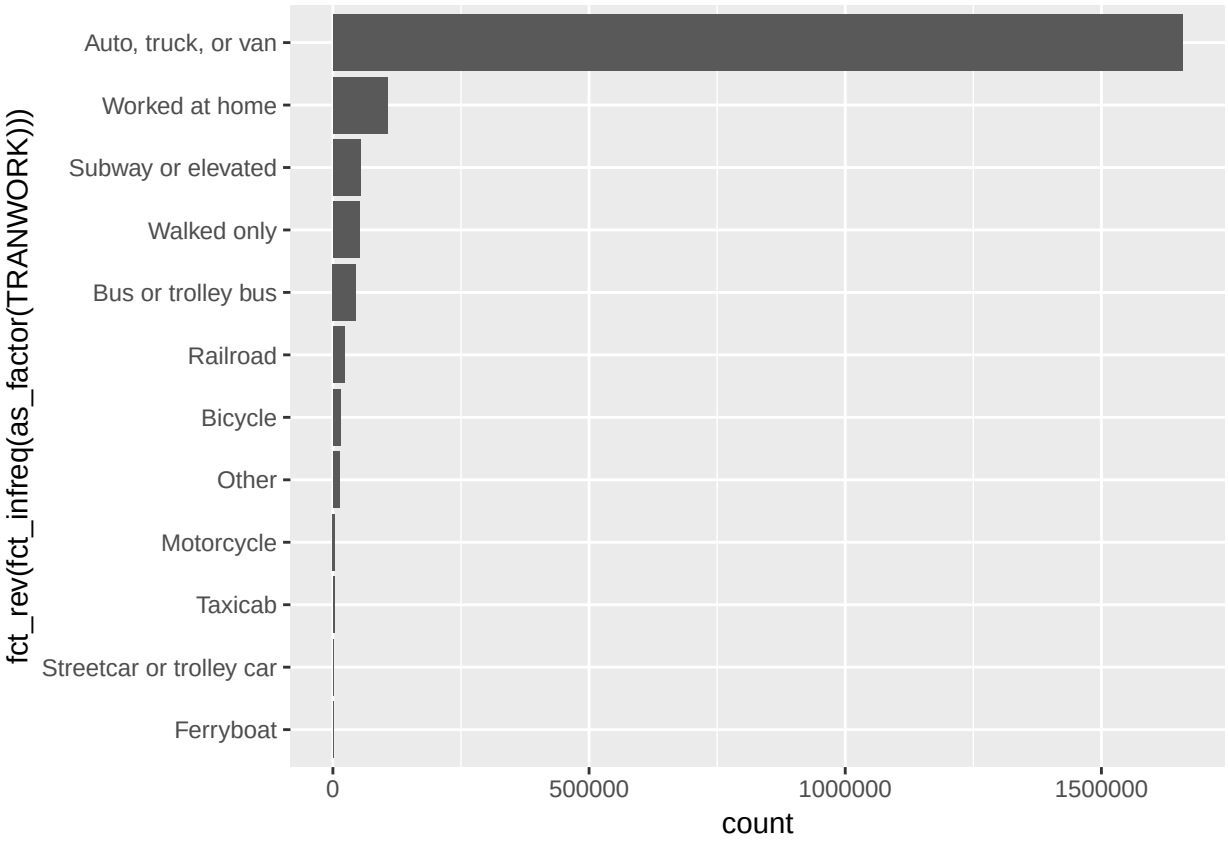

Table 0

```
##      SEX              AGE                                EDUC
##  Male  :979559   Min.   :16.00   4 years of college       :749970
##  Female:994855   1st Qu.:31.00   5+ years of college      :477416
##                  Median :43.00   1 year of college        :474087
##                  Mean   :43.09   2 years of college       :272941
##                  3rd Qu.:54.00   N/A or no schooling      :     0
```

```
##                  Max.   :97.00   Nursery school to grade 4:     0
##                                    (Other)                  :     0
##     EMPSTAT     UHRSWORK        INCWAGE
##  Min.   :1   Min.   : 1.00   Min.   :     4
##  1st Qu.:1   1st Qu.:40.00   1st Qu.: 25700
##  Median :1   Median :40.00   Median : 50000
##  Mean   :1   Mean   :40.25   Mean   : 68500
##  3rd Qu.:1   3rd Qu.:45.00   3rd Qu.: 85000
##  Max.   :1   Max.   :99.00   Max.   :736000
##
##                 TRANWORK
##  Auto, truck, or van:1658507
##  Worked at home     : 106897
##  Subway or elevated :  54425
##  Walked only        :  51847
##  Bus or trolley bus :  44898
##  Railroad           :  22074
##  (Other)            :  35766
##                                          MET2013           TRANTIME
##  New York-Newark-Jersey City, NY-NJ-PA       : 164433   Min.   :  0.00
##  Los Angeles-Long Beach-Anaheim, CA          : 105144   1st Qu.: 12.00
##  Chicago-Naperville-Elgin, IL-IN-WI          :  74573   Median : 20.00
##  Washington-Arlington-Alexandria, DC-VA-MD-WV:  68611   Mean   : 27.09
##  Dallas-Fort Worth-Arlington, TX             :  59478   3rd Qu.: 35.00
##  Boston-Cambridge-Newton, MA-NH              :  51011   Max.   :165.00
##  (Other)                                     :1451164
```

Table 1: Regression Table

	Dependent variable:
	log(INCWAGE)
TRANWORKMotorcycle	-0.005 (0.013)
TRANWORKBus or trolley bus	-0.089^{***} (0.004)
TRANWORKStreetcar or trolley car	0.024 (0.020)
TRANWORKSubway or elevated	0.174^{***} (0.003)
TRANWORKRailroad	0.367^{***} (0.005)
TRANWORKTaxicab	0.034^{**} (0.015)
TRANWORKFerryboat	0.399^{***} (0.021)
TRANWORKBicycle	-0.099^{***} (0.007)
TRANWORKWalked only	-0.278^{***} (0.004)
TRANWORKOther	-0.071^{***} (0.007)
TRANWORKWorked at home	-0.001 (0.003)
UHRSWORK	0.042^{***} (0.0001)
EDUC2 years of college	0.113^{***} (0.002)
EDUC4 years of college	0.398^{***} (0.001)
EDUC5+ years of college	0.619^{***} (0.002)
AGE	0.104^{***} (0.0003)
I(AGE^2)	-0.001^{***} (0.00000)
SEXFemale	-0.224^{***} (0.001)
Constant	6.340^{***} (0.005)
Observations	1,974,414
R^2	0.473
Adjusted R^2	0.473
Residual Std. Error	0.794 (df = 1974395)
F Statistic	$98{,}445.970^{***}$ (df = 18; 1974395)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2:

	Dependent variable:
	log(INCWAGE)
TRANWORKMotorcycle	-0.005 (0.013)
TRANWORKBus or trolley bus	-0.089^{***} (0.004)
TRANWORKStreetcar or trolley car	0.024 (0.020)
TRANWORKSubway or elevated	0.174^{***} (0.003)
TRANWORKRailroad	0.367^{***} (0.005)
TRANWORKTaxicab	0.034^{**} (0.015)
TRANWORKFerryboat	0.399^{***} (0.021)
TRANWORKBicycle	-0.099^{***} (0.007)
TRANWORKWalked only	-0.278^{***} (0.004)
TRANWORKOther	-0.071^{***} (0.007)
TRANWORKWorked at home	-0.001 (0.003)
UHRSWORK	0.042^{***} (0.0001)
EDUC2 years of college	0.113^{***} (0.002)
EDUC4 years of college	0.398^{***} (0.001)
EDUC5+ years of college	0.619^{***} (0.002)
AGE	0.104^{***} (0.0003)
I(AGE^2)	-0.001^{***} (0.00000)
SEXFemale	-0.224^{***} (0.001)
Constant	6.340^{***} (0.005)
Observations	1,974,414
R^2	0.473
Adjusted R^2	0.473
Residual Std. Error	0.794 (df = 1974395)
F Statistic	$98{,}445.970^{***}$ (df = 18; 1974395)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 2

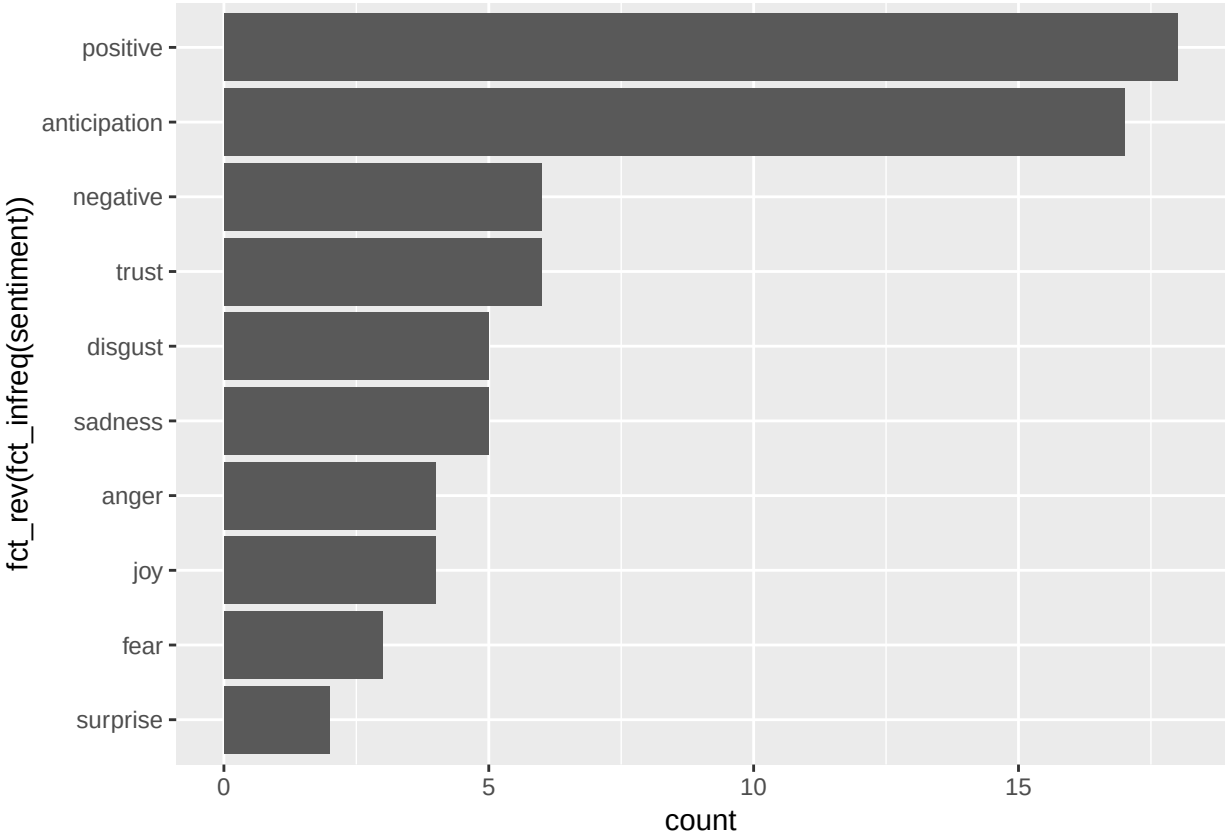
positive
anticipation
negative
trust
disgust
sadness
anger
joy
fear
surprise
fct_rev(fct_infreq(sentiment))
0
5
10
15
count

Table 3:

	Dependent variable:
	log(INCWAGE)
TRANWORKMotorcycle	-0.019 (0.012)
TRANWORKBus or trolley bus	-0.205^{***} (0.004)
TRANWORKStreetcar or trolley car	-0.083^{***} (0.020)
TRANWORKSubway or elevated	0.013^{***} (0.004)
TRANWORKRailroad	0.185^{***} (0.006)
TRANWORKTaxicab	-0.027^{*} (0.015)
TRANWORKFerryboat	0.235^{***} (0.021)
TRANWORKBicycle	-0.136^{***} (0.007)
TRANWORKWalked only	-0.286^{***} (0.004)
TRANWORKOther	-0.120^{***} (0.007)
TRANWORKWorked at home	0.041^{***} (0.003)
UHRSWORK	0.042^{***} (0.0001)
EDUC2 years of college	0.115^{***} (0.002)
EDUC4 years of college	0.377^{***} (0.001)
EDUC5+ years of college	0.590^{***} (0.002)
AGE	0.103^{***} (0.0003)
I(AGE^2)	-0.001^{***} (0.00000)
SEXFemale	-0.217^{***} (0.001)
TRANTIME	0.002^{***} (0.00003)
Observations	1,974,414
R^2	0.484
Adjusted R^2	0.484
Residual Std. Error	0.786 (df = 1974135)
F Statistic	$6{,}656.274^{***}$ (df = 278; 1974135)
Note:	*p<0.1; **p<0.05; ***p<0.01

References

Amano, Kozo and Dai Nakagawa, "Study on urbanization impacts by new stations of High Speed Railway," 1990.

Haynes, Kingsley E, "Labor markets and regional transportation improvements: the case of high-speed trains An introduction and review," *The annals of regional science*, 1997, *31* (1), 57–76.

Hirota, R, "Present Situation and Effects of the Shinkansen, realzione presentata all'International Seminar on High-Speed Trains," 1984.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, *IPUMS USA: Version 9.0 [dataset]*, Minneapolis, MN: IPUMS, 2019.

Weisbrod, Glen and Arlee Reno, *Economic impact of public transportation investment*, Citeseer, 2009.