Undergraduate Economic Review

Volume 16 | Issue 1 Article 18

2019

Effect of Unemployment Length on Employment Expectations

Kamyar Kamyar
Drexel University, mr.kamyar@gmail.com

Effect of Unemployment Length on Employment Expectations

Abstract

Unemployment often has devastating effects on individuals -- both in financial and psychological terms. Depending on the type and category of unemployment, its length varies; and as its length increases it may implement biased thought in individuals' predictions regarding future employment. This paper's primary purpose is to measure and discuss how the time length that one has been unemployed for affects his or her expectations on his or her own short-term possibility of employment. The results suggest a strong opposite link between one's prediction of future employment and the same person's prior unemployment period. This paper was originally written in June 2019 as the term project to a master's-level econometrics course; and it later was revised for submission in February 2020.

Keywords

labor economics, job market, job search, unemployment, unemployment time, proportional, unemployment length, employment expectations, regression analysis

Cover Page Footnote

I need to thank Dr. Sebastien Bradley of Drexel University whose constructive comments were crucial to creation of this paper.

Effect of Unemployment Length on Employment Expectations

Kamyar Kamyar*

Introduction

One leading predictor of financial markets' term structure is the expectations theory; that is, a market's long-term rate is entirely determined by the expectations of future short-term rates. However, in an economy, the individuals' expectations will obviously not induce effects on the macroeconomy. Nevertheless, people's personal economic predictions are leading drivers of their microeconomic decisions, regardless of their economic and financial literacy. Such individual gauges, however, are commonly invalidated by biased thought. For instance, "people may be more influenced by recent history than by the very latest fluctuations or by longer-term trends when asked to gauge their nation's economic health", as suggested by a recent article (DeSilver 2017). Another potential source of bias could be unemployment, as its effects on individuals are often devastating, both in financial and psychological terms. Depending on the type and category of unemployment, its length varies; and as its length increases it may implement additional biased thought in individuals' predictions regarding future employment. This paper's primary purpose is to measure and discuss how the time length that one has been unemployed for affects his or her expectations on his or her own short-term possibility of employment (measured in 3-month and 12-month units) in a plausible position.

Data

This paper employs data from The *Survey of Consumer Expectations* (the "SCE") developed and owned by the Federal Reserve Bank of New York ("FRBNY"). The SCE "gathers

* Kamyar.kamyar@drexel.edu; Undergraduate Student of Economics and Mathematics at Drexel University, 2017 – 2020.
Paper originally written in June 2019; revised for submission in February 2020. I need to thank Dr. Sebastien Bradley of Drexel University whose constructive comments were crucial to creation of this paper.

information on consumer expectations regarding inflation, household finance, the labor and housing markets, and other economic issues"; it was launched in 2013 and has been fielded monthly since. Nonetheless, in this paper in order to achieve an additional ceteris paribus effect that yields solely up-to-date results, the pooled cross-sectional SCE database from 01/2017 to 06/2018 was selected for econometric analysis. General macroeconomic factors have remained in a relatively stable situation at all times of the survey, implying relative stability in internal validity. In that time period national unemployment rates only experienced slight fluctuations in the interval of 3.8% to 4.7%; and monthly national inflation expectation was also consistent fluctuating in the interval of 2.4% to 3.0%.

The complete SCE database of 01/2017 to 06/2018 includes 23,927 records, a great number of which being irrelevant to the context of this research as the individuals were employed at the time of reporting and/or did not wish to report the relevant sections. Using Python, the rows in which "Not working, but would like to work" was chosen as an answer for question no. 10 (Q10) "What is your current employment situation?" of the questionnaire were extracted and copied to another database exclusively used for this research. Since most people did not wish to report and/or were employed at the time of filling the survey, after omitting the records that did not indicate their length of unemployment the secondary database's number of samples is reduced to 772. A division of the secondary sample's number of observations by the ones of the initial sample gives an outcome of 3.22% - implying that the sample is in fact nearly representative of the macroeconomy.

The questionnaire distinguishes the two groups of unemployed individuals as unemployed (Q16) or out of work (Q19). By definition, "To be classified as unemployed in the month they are surveyed, people must be actively looking for work. If they are not actively looking, they are classified as not in the labor force". In order to maintain validity, both groups were merged into one and all the lengths of unemployment were entered in the row of Q16. Even though the motivations of each group may differ by definition, given the available data it would have introduced bias would we have separated the two, as the observations of each of the groups by its own is relatively small.

Model

The model consists of two sets of regression equations. The models' independent variables are as following: length of unemployment (UnempLength), age (Age), age^2 (Age_2), gender (binary; Gender), interaction of gender and level of education, and level of education (binary; Interaction). It should be noted that the primary variable of interest is the length of unemployment (UnempLength), and the latter variables were added in order to provide detailed explanations and correlations for individuals belonging to each of the groups.

The level of education (Q36) indicates the individual's level of education in a categorical manner as following: less than high school, high school diploma, some college but no degree, associate/junior college degree, bachelor's degree, master's degree, and doctoral degree. However, as few respondents (only 115 out of 772) indicated their education level, the source of variation were impractical had such a wide categorical analysis been used. Instead, the factor of education level was converted to a binary variable, using "0" for the former four categories (below a finished college degree) and "1" for the latter four (college graduate and above). This new category is labeled as "EducationLevel".

The two dependent variables are "percent chance expectation that the person would find a job and accept it in the following 12 months (Jobacceptpred12)", and "percent chance expectation that the person would find a job and accept it in the following 3 months (Jobacceptpred3)". Both variables measure an unemployed individual's personal expectations in the job market on average.

The two-tailed hypothesis test for all equations is:

$$H_0: \beta_1 = 0; H_1: Otherwise$$

, with β_1 being the coefficient of unemployment length (Q16) for all equations. Based on the results in Marcel Garz' article "Unemployment expectations, excessive pessimism, and news coverage", it's expected that β_1 be inversely correlated with all dependent variables. That is, the more time a person has been unemployed for, the less we expect that his or her expectations of employment quantitative prospects to be on average. Not to mention, since every dependent and major independent variable is already measured in percentages, it's redundant to estimate the models in logarithmic forms, yet quadratic models will be used for age. Finally, since n > 30 in all models, the distribution of β_i is approximated by the normal distribution.

Analysis

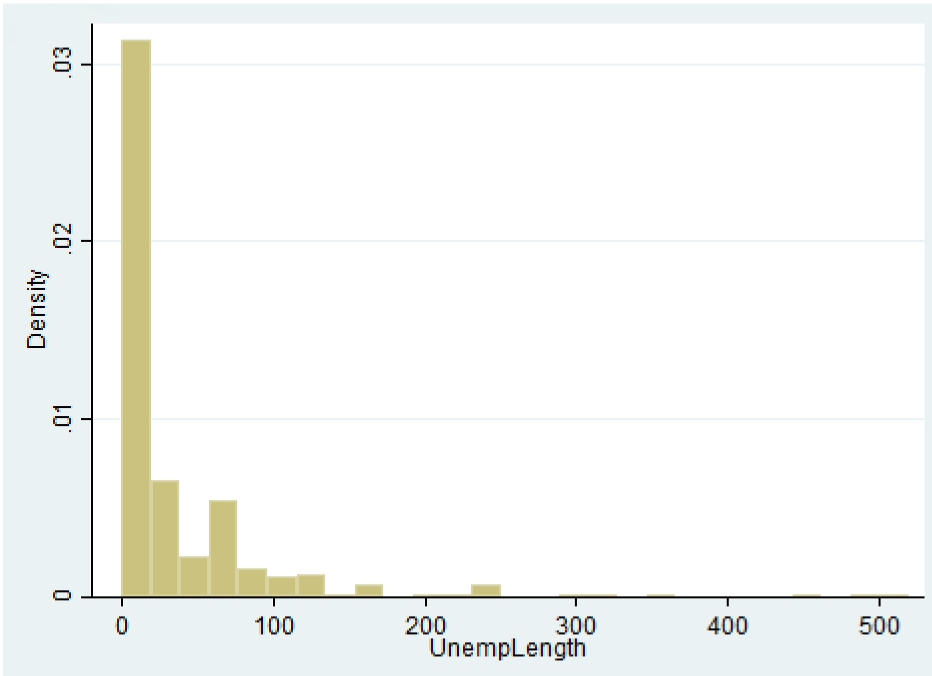
Summaries of the desired variables are provided below.

Summary List 1

Variable	Obs	Mean	Std. Dev.	Min	Max
UnempLength	772	35.15285	58.00757	0	520

Table 1: Summary of length of unemployment (months). The database has 772 records, ranging from 0 (translated to less than one month) to 520 months of unemployment at the time of reporting. The mean of unemployment period is 35.15 months, and the standard deviation is

58.01. Though such a large standard deviation (also depicted later in scatterplots) implies the possibility of outliers existing in the data, data was not manipulated in order to exercise caution.



Histogram 1: Depicted distribution of the reporting individual's unemployment length, as explained for Table 1.

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	115	50.48696	14.97739	22	81

Table 2: Summary of the age of unemployed individuals (years). The database has 115 records reporting age, ranging from 22 to 81 years old at the time of reporting. The mean of age is 50.49, and the standard deviation is 14.98.

Variable	Obs	Mean	Std. Dev.	Min	Max
Gender	115	.6	.4920419	0	1

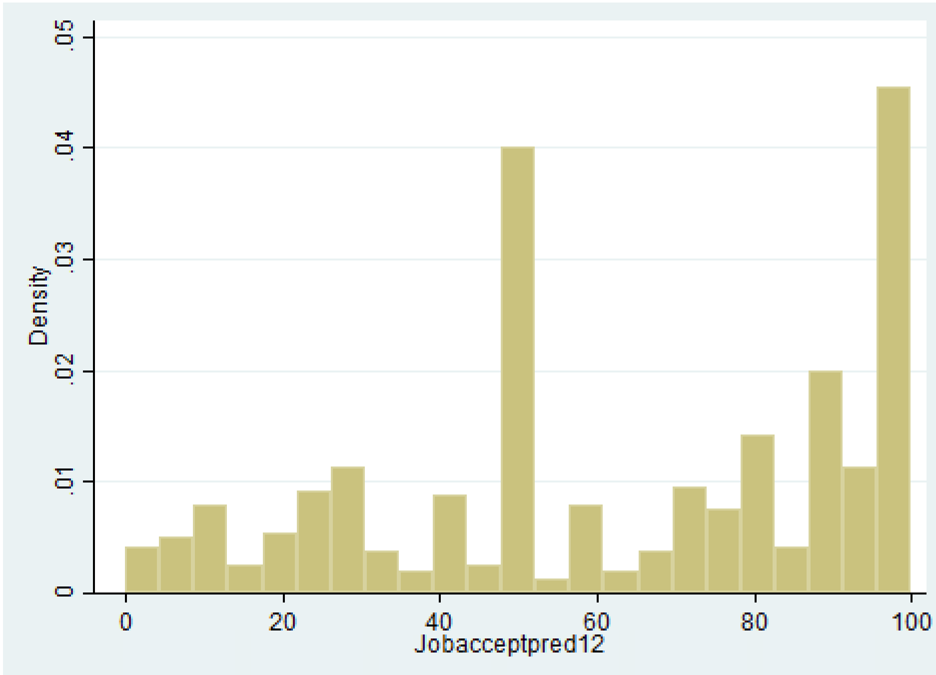
Table 3: Summary of gender of unemployed individuals (binary). 0 is translated to male, and 1 is translated to female. The database of unemployed individuals has 115 records reporting gender; as the mean (0.6) is skewed towards 1, it can be inferred that there are more females in the database than males.

Variable	Obs	Mean	Std. Dev.	Min	Max
EducationL~1	115	.4173913	.4952867	0	1

Table 4: Summary of education of unemployed individuals (binary). 0 is translated to below a college degree, and 1 is translated to college degree and above. The database has 115 records reporting education; as the mean (0.42) is skewed towards 0, it can be inferred that the education level of the individuals in the sample is slightly more often below a college degree.

Variable	Obs	Mean	Std. Dev.	Min	Max
Jobaccept~12	551	63.25045	30.14883	0	100

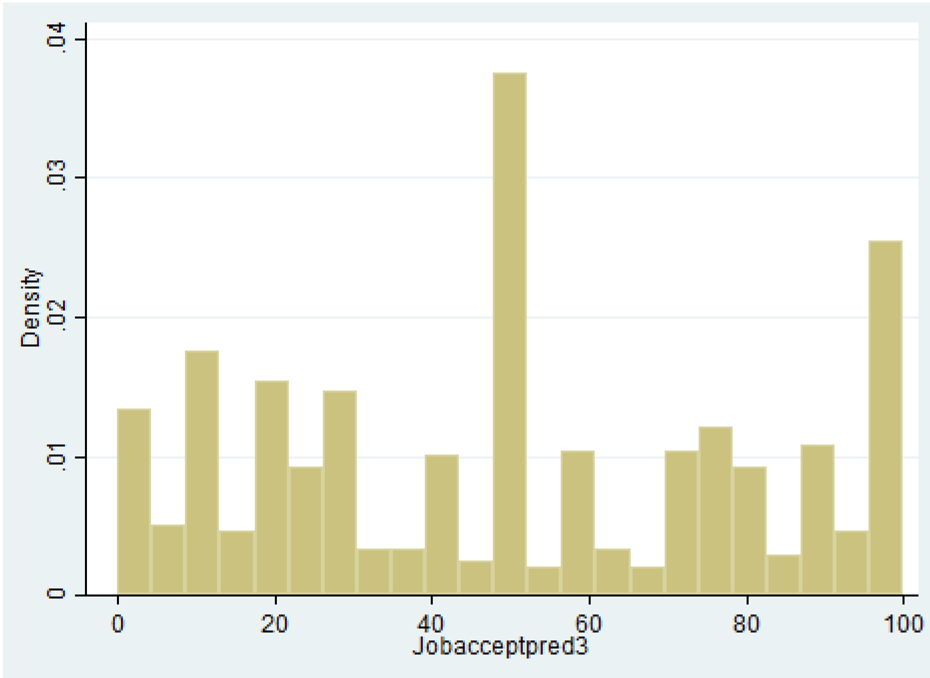
Table 5: Summary of expectation of the percent chance that the person would find a job and accept it in the following 12 months. The mean of 551 records is 63.25%, with a standard deviation of 30.15%.



Histogram 2: Depicted distribution of the percent chance that the person would find a job and accept it in the following 12 months, as explained for Table 5.

Variable	Obs	Mean	Std. Dev.	Min	Max
Jobacceptp~3	551	50.2196	30.81642	0	100

Table 6: Summary of expectation of the percent chance that the person would find a job and accept it in the following 3 months. The mean of 551 records is 50.22%, with a standard deviation of 30.82%.



Histogram 3: Depicted distribution of the percent chance that the person would find a job and accept it in the following 3 months, as explained for Table 6. Minor differences among Histograms 2 and 3 exist, including less density in extremely high answers in the former.

Regression Model 1A

$$\boldsymbol{Y = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 (X_{3i} \times X_{5i}) + \beta_5 X_{4i} + \beta_6 X_{2i}^2}$$

Y = expectation of the percent chance that the person would find a job and accept it in the following 12 months (Jobacceptpred12)

$\boldsymbol{\alpha}$ = constant term

$\boldsymbol{\beta_i}$= coefficients

$\boldsymbol{X_{1i}}$= length of unemployment (months; UnempLength)

$\boldsymbol{X_{2i}}$= age (years; Age)

$\boldsymbol{X_{3i}}$= gender (binary; Gender)

$(\boldsymbol{X_{3i} \times X_{6i}})$= interaction effect of gender and education

$\boldsymbol{X_{5i}}$= education (binary; EducationLevel)

$\boldsymbol{X_{6i}^2}$= age^2(Age_2)

Results (Table 7)

```
Linear regression                               Number of obs     =         88
                                                F(6, 81)          =       5.62
                                                Prob > F          =     0.0001
                                                R-squared         =     0.2082
                                                Root MSE          =     28.495
```

Jobacceptpr~12	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
UnempLength	-.3273041	.1086138	-3.01	0.003	-.5434114	-.1111967
Age	-1.417845	1.333857	-1.06	0.291	-4.071802	1.236112
Gender	-17.90645	8.7629	-2.04	0.044	-35.34187	-.4710238
Interaction	31.78104	13.15848	2.42	0.018	5.599787	57.96228
EducationLevel	-17.70807	10.54813	-1.68	0.097	-38.69554	3.279414
Age_2	.0103483	.0142572	0.73	0.470	-.018019	.0387157
_cons	120.9897	29.8354	4.06	0.000	61.62661	180.3528

The number of sample outcomes for which all desired variables are reported is 88. The goodness-of-fit measurement of 0.2082 confirms the correctness of the variable choices to an extent. The primary expectation that was held prior to performing the regression is also confirmed; that is, the more time a person has been unemployed for, the less his or her expectations of finding a job and accepting it in the following 12 months are on average. The t-statistic of β_1 , -3.01, is significant; and, based on a two-tailed 99% hypothesis test we fail to reject the null hypothesis and conclude that since $|t_{\beta_1}| \geq c_{99\%}$ (3.01 > 2.58) the effect of the length of unemployment on the expectations of finding a job in the following year is statistically significant.

Despite that the constant term has a statistically strong t-statistic, it cannot provide any intrinsic meaning since in this context age cannot take a meaningful value of 0 (that would be necessary for the constant term to be meaningful); therefore, the constant solely provides a base for computing other factors. In addition, the rest of independent variables' coefficients have weak statistical significance. Based on a general significance cutoff of 95% and using their computed p-values, most of individual two-tailed null hypotheses of $\beta_i = 0$ $(i = 2, 5, 6)$ will be rejected and thereby we conclude that none of their effects is statistically different than 0 on expectation of the percent chance that the person would find a job and accept it in the following 12 months. The two hypothesis tests that cannot be individually rejected are the ones of $i = 3, 4$; that is, the gender and interaction effects imply statistical significance. Additionally, the meaning of each coefficient is described in the table below (Table 8).

<table>
<tr><th>Coefficient</th><th>Meaning</th></tr>
<tr><td>$\boldsymbol{\beta_1}$</td><td>A one-month increase in the length of unemployment is associated with a 0.33 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 12 months, on average and holding everything else constant.</td></tr>
<tr><td>$\boldsymbol{\beta_2}$</td><td>A one-year increase in the individual's age is associated with a 1.42 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 12 months, on average and holding everything else constant.</td></tr>
<tr><td>$\boldsymbol{\beta_3}$</td><td>On average and holding everything else constant, being female is associated with a 17.91 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 12 months.</td></tr>
<tr><td>$\boldsymbol{\beta_4}$</td><td>
<table>
<tr><td>Table 9: Expected Values</td><td colspan="2">Education</td></tr>
<tr><td>Gender</td><td>Below College
$(\beta_5 = 0)$</td><td>College and Above
$(\beta_5 = 1)$</td></tr>
<tr><td>Male $(\beta_3 = 0)$</td><td>$\hat{y} = \hat{\alpha}$
$\hat{y} = 120.99$</td><td>$\hat{y} = \hat{\alpha} + \beta_5$
$\hat{y} = 103.28$</td></tr>
<tr><td>Female $(\beta_3 = 1)$</td><td>$\hat{y} = \hat{\alpha} + \beta_3$
$\hat{y} = 103.08$</td><td>$\hat{y} = \hat{\alpha} + \beta_3 + \beta_4 + \beta_5$
$\hat{y} = 117.16$</td></tr>
</table>
Each of the intercepts computed in the table above is a base that explains the effect of belonging into each of the four groups on expectation of the chance that the person would find a job and accept it in the following 12 months. Since the constant is positive and holds a significantly large value (120.99), belonging to each of the groups above is associated with a positive effect on the expectation. However, it should be noted that there is a significant difference between the expectation of some of the groups.</td></tr>
</table>

$\boldsymbol{\beta_5}$	On average and holding everything else constant, holding a bachelor's degree or any professional degree above is associated with a 17.71 percent decrease in expectation of the chance that the person would find a job and accept it in the following 12 months in contrast with holding below a bachelor's degree.
$\boldsymbol{\beta_6}$	The variable age^2(Age_2) was added to the model as its addition increased the model's goodness-of-fit measurement by 0.0039. It suggests that a one-unit increase in the individual's age^2 on average is associated with a 0.01 percent point increase in expectation of the chance that the person would find a job and accept it in the following 12 months. The underlying rationale of the coefficient's sign may be due to the indirect relation between one's age and his or her achievement expectations.

The major problem undermining the sample selection of the *Regression Model 1A* is its low number of observations. Though the database includes 772 observations, only 88 of the individuals had reported all the explanatory variables defined in *Regression Model 1A.* In order to address that, *Regression Model 1B* does not take the explanatory coefficients into account and takes only "UnempLength" into account instead.

Regression Model 1B

$$\boldsymbol{Y = \alpha + \beta_1 X_{1i}}$$

Y = expectation of the percent chance that the person would find a job and accept it in the following 12 months (Jobacceptpred12)

$\boldsymbol{\alpha}$ = constant term

$\boldsymbol{\beta_{1i}}$= coefficient

$\boldsymbol{X_{1i}}$= length of unemployment (months; UnempLength)

Results (Table 10)

Linear regression			
	Number of obs	=	551
	F(1, 549)	=	53.11
	Prob > F	=	0.0000
	R-squared	=	0.0813
	Root MSE	=	28.923

Jobaccept~12	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
UnempLength	-.1956282	.026844	-7.29	0.000	-.2483578	-.1428986
_cons	67.83653	1.356291	50.02	0.000	65.17238	70.50069

The increased number of observations (to 551, that is the number of individuals who reported their unemployment lengths) increases the sample selection of the research. Based on the significant t- and p-values provided, the test is statistically significant. The goodness-of-fit measurement of 0.0813, however, suggests that the previous model was better fit. The primary expectation that was held prior to performing the regression is confirmed again; that is, the more time a person has been unemployed for, the less his or her expectations of finding a job and accepting it in the following 12 months are on average. The explanation of the coefficient $\boldsymbol{\beta_{1i}}$ is that a one-month increase in the length of unemployment is associated with a 0.20 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 12 months, on average and holding everything else constant. In order to interpret the statistically-significant (t=50) constant term of 67.84 we would need to define a person with zero unemployment length ($\boldsymbol{\beta_{1i}}$=0), which makes no realistic sense in a context where the constructed data is for unemployed individuals. Therefore, the constant solely provides a base for computing other factors again.

In order to find out the extent of effectiveness of the control variables, another regression equation is added. *Regression Model 1C* is similar to *Regression Model 1B* in terms of its chosen dependent and independent variables. However, it's using the same sample of observations that was used in *Regression Model 1A*.

Regression Model 1C

$$Y = \alpha + \beta_1 X_{1i}$$

Y = expectation of the percent chance that the person would find a job and accept it in the following 12 months (Jobacceptpred12)

α = constant term

β_{1i}= coefficient

X_{1i}= length of unemployment (months; UnempLength)

Results (Table 11)

Linear regression		
Number of obs	=	88
F(1, 86)	=	11.93
Prob > F	=	0.0009
R-squared	=	0.0995
Root MSE	=	29.492

Jobaccept~12	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
UnempLength	-.3940819	.1141154	-3.45	0.001	-.6209358	-.1672279
_cons	69.29722	3.743	18.51	0.000	61.85639	76.73806

Based on the significant t- and p-values provided, the test is statistically significant. The goodness-of-fit measurement of 0.0995, however, suggests that the *1A* model was better fit. The primary expectation that was held prior to performing the regression is confirmed again; that is, the more time a person has been unemployed for, the less his or her expectations of finding a job and accepting it in the following 12 months are on average. The explanation of the coefficient β_{1i}

is that a one-month increase in the length of unemployment is associated with a 0.40 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 12 months, on average and holding everything else constant. In order to interpret the statistically-significant (t=18.51) constant term of 69.3 we would need to define a person with zero unemployment length ($\boldsymbol{\beta_{1i}}$=0), which makes no realistic sense in a context where the constructed data is for unemployed individuals. Therefore, the constant solely provides a base for computing other factors again. The major change (16.76%) observed in the estimated coefficient of "UnempLength" across models *1A* and *1C* shows that inclusion of the control variables has been helpful in terms of addressing bias.

Regression Model 2A

$$\boldsymbol{Y = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 (X_{3i} \times X_{5i}) + \beta_5 X_{4i} + \beta_6 X_{2i}^2}$$

Y = expectation of the percent chance that the person would find a job and accept it in the following 3 months (Jobacceptpred3)

$\boldsymbol{\alpha}$ = constant term

$\boldsymbol{\beta_i}$= coefficients

$\boldsymbol{X_{1i}}$= length of unemployment (months; UnempLength)

$\boldsymbol{X_{2i}}$= age (years; Age)

$\boldsymbol{X_{3i}}$= gender (binary; Gender)

($\boldsymbol{X_{3i} \times X_{6i}}$)= interaction effect of gender and education

$\boldsymbol{X_{5i}}$**=** education (binary; EducationLevel)

$\boldsymbol{X_{6i}^2}$**=** age^2(Age_2)

Results (Table 12)

```
Linear regression                               Number of obs     =         88
                                                F(6, 81)          =       4.27
                                                Prob > F          =     0.0009
                                                R-squared         =     0.2004
                                                Root MSE          =     31.455
```

Jobacceptpred3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
UnempLength	-.3843741	.1329484	-2.89	0.005	-.6488997	-.1198485
Age	-1.124979	1.66525	-0.68	0.501	-4.438304	2.188345
Gender	-7.332687	10.02456	-0.73	0.467	-27.27841	12.61303
Interaction	28.23492	13.72752	2.06	0.043	.921457	55.54837
EducationLevel	-23.14718	10.36975	-2.23	0.028	-43.77973	-2.514631
Age_2	.0081215	.0170969	0.48	0.636	-.0258959	.0421389
_cons	96.51473	39.08021	2.47	0.016	18.75736	174.2721

This 3-month prospective equation was regressed in order to be compared and thereby confirm the primary results found in the 12-month prospective regression equation. The number of sample outcomes for which all desired variables are reported is 88. The relatively high goodness-of-fit measurement of 0.2004 confirms the relative fitness of the variable choices again. The primary expectation that was held prior to performing the regression is confirmed again; that is, the more time a person has been unemployed for, the less his or her expectations of finding a job and accepting it in the following 3 months are on average. The t-statistic of β_1 , -2.89, is significant; and, based on a two-tailed 99% hypothesis test we fail to reject the null hypothesis and conclude that since $|t_{\beta_1}| \geq c_{99\%}$ (2.89 > 2.58) the effect of the length of unemployment on the expectations of finding a job in the following quarter is repeatedly statistically significant.

Despite that the constant term has a statistically strong t-statistic, it cannot provide any intrinsic meaning since in this context age cannot take a meaningful value of 0 (that would be necessary for the constant term to be meaningful); therefore, the constant solely provides a base for computing other factors. In addition, the rest of independent variables' coefficients have weak statistical significance. Based on a general significance cutoff of 95% and using their computed p-values, most of two-tailed null hypotheses of $\beta_i = 0$ $(i = 2, 3, 6)$ will be individually rejected and thereby we conclude that none of their individual effects is statistically different than 0 on expectation of the percent chance that the person would find a job and accept it in the following 3 months. The two hypothesis tests that cannot be individually rejected are the ones of $i = 4, 5$; that is, the education and interaction effects have statistical significance. The meaning of each coefficient is described in the table below (Table 13).

<table>
<tr><th>Coefficient</th><th>Meaning</th></tr>
<tr><td>$\boldsymbol{\beta_1}$</td><td>A one-month increase in the length of unemployment is associated with a 0.38 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 3 months, on average and holding everything else constant.</td></tr>
<tr><td>$\boldsymbol{\beta_2}$</td><td>A one-year increase in the individual’s age is associated with a 1.12 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 3 months, on average and holding everything else constant.</td></tr>
<tr><td>$\boldsymbol{\beta_3}$</td><td>On average and holding everything else constant, being female is associated with a 7.33 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 3 months.</td></tr>
<tr><td>$\boldsymbol{\beta_4}$</td><td>
<table>
<tr><th>Table 14: Expected Values</th><th colspan="2">Education</th></tr>
<tr><td>Gender</td><td>Below College
$(\beta_5 = 0)$</td><td>College and Above
$(\beta_5 = 1)$</td></tr>
<tr><td>Male ($\boldsymbol{\beta_3 = 0}$)</td><td>$\hat{y} = \hat{\alpha}$
$\hat{y} = 96.51$</td><td>$\hat{y} = \hat{\alpha} + \beta_5$
$\hat{y} = 73.37$</td></tr>
<tr><td>Female ($\boldsymbol{\beta_3 = 1}$)</td><td>$\hat{y} = \hat{\alpha} + \beta_3$
$\hat{y} = 89.18$</td><td>$\hat{y} = \hat{\alpha} + \beta_3 + \beta_4 + \beta_5$
$\hat{y} = 94.27$</td></tr>
</table>
Each of the intercepts computed in the table above is a base that explains the effect of belonging into each of the four groups on expectation of the chance that the person would find a job and accept it in the following 3 months. Since the constant is positive and holds a significantly large value (96.51), belonging to each of the groups above is associated with a positive effect on the expectation. However, it should be noted that there is a significant difference between the expectation of some of the groups.</td></tr>
<tr><td>$\boldsymbol{\beta_5}$</td><td>On average and holding everything else constant, holding a bachelor’s degree or any professional degree above is associated with a 23.15 percent decrease in expectation of the chance that the person would find a job and accept it in the following 3 months in contrast with holding below a bachelor’s degree.</td></tr>
</table>

$\boldsymbol{\beta_6}$	The variable age^2(Age_2) was added to the model as its addition increased the model's goodness-of-fit measurement by 0.0020. It suggests that a one-unit increase in the individual's age^2 on average is associated with a 0.01 percent point increase in expectation of the chance that the person would find a job and accept it in the following 3 months. The underlying rationale of the coefficient's sign may be due to the indirect relation between one's age and his or her achievement expectations.

The major problem undermining the sample selection of the *Regression Model 2A* is its low number of observations. Though the database includes 772 observations, only 88 of the individuals had reported all the explanatory variables defined in *Regression Model 2A.* In order to address that, *Regression Model 1B* does not take the explanatory coefficients into account and takes only "UnempLength" into account instead.

Regression Model 2B

$$\boldsymbol{Y = \alpha + \beta_1 X_{1i}}$$

Y = expectation of the percent chance that the person would find a job and accept it in the following 3 months (Jobacceptpred3)

$\boldsymbol{\alpha}$ = constant term

$\boldsymbol{\beta_{1i}}$= coefficient

$\boldsymbol{X_{1i}}$= length of unemployment (months; UnempLength)

Results (Table 15)

Linear regression		Number of obs	=	551
		F(1, 549)	=	49.30
		Prob > F	=	0.0000
		R-squared	=	0.0871
		Root MSE	=	29.47

Jobacceptp~3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
UnempLength	-.2069488	.0294747	-7.02	0.000	-.2648457	-.1490519
_cons	55.07107	1.453161	37.90	0.000	52.21663	57.9255

The increased number of observations (to 551, that is the number of individuals who reported their unemployment lengths) increases the sample selection of the research. Based on the significant t- and p-values provided, the test is statistically significant. The goodness-of-fit measurement of 0.0871, however, suggests that the previous model was better fit. The primary expectation that was held prior to performing the regression is confirmed again; that is, the more time a person has been unemployed for, the less his or her expectations of finding a job and accepting it in the following 3 months are on average. The explanation of the coefficient $\boldsymbol{\beta_{1i}}$ is that a one-month increase in the length of unemployment is associated with a 0.21 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 3 months, on average and holding everything else constant. In order to interpret the statistically-significant (t=37.90) constant term of 55.07 we would need to define a person with zero unemployment length ($\boldsymbol{\beta_{1i}}$=0), which makes no realistic sense in a context where the constructed data is for unemployed individuals. Therefore, the constant solely provides a base for computing other factors again.

In order to find out the extent of effectiveness of the control variables, another regression equation is added. *Regression Model 2C* is similar to *Regression Model 2B* in terms of its chosen dependent and independent variables. However, it's using the same sample of observations that was used in *Regression Model 2A.*

Regression Model 2C

$$Y = \alpha + \beta_1 X_{1i}$$

Y = expectation of the percent chance that the person would find a job and accept it in the following 3 months (Jobacceptpred3)

α = constant term

β_{1i}= coefficient

X_{1i}= length of unemployment (months; UnempLength)

Results (Table 16)

Linear regression			
	Number of obs	=	88
	F(1, 86)	=	14.14
	Prob > F	=	0.0003
	R-squared	=	0.1273
	Root MSE	=	31.891

Jobacceptp~3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
UnempLength	-.4896616	.1302313	-3.76	0.000	-.7485528	-.2307704
_cons	57.9955	4.303598	13.48	0.000	49.44023	66.55077

Based on the significant t- and p-values provided, the test is statistically significant. The goodness-of-fit measurement of 0.1273, however, suggests that the *2A* model was better fit. The primary expectation that was held prior to performing the regression is confirmed again; that is, the more time a person has been unemployed for, the less his or her expectations of finding a job and accepting it in the following 3 months are on average. The explanation of the coefficient $\boldsymbol{\beta_{1i}}$ is that a one-month increase in the length of unemployment is associated with a 0.49 percent point decrease in expectation of the chance that the person would find a job and accept it in the following 3 months, on average and holding everything else constant. In order to interpret the statistically-significant (t=13.48) constant term of 58 we would need to define a person with zero unemployment length ($\boldsymbol{\beta_{1i}}$=0), which makes no realistic sense in a context where the constructed data is for unemployed individuals. Therefore, the constant solely provides a base for computing other factors again. The major change (21.39%) observed in the estimated coefficient of "UnempLength" across models *2A* and *2C* shows that inclusion of the control variables has been helpful in terms of addressing bias.

The results of all six regression models tend to support each other and thereby strengthen the validity of the research. The signs of all resulting coefficients are similar; and the coefficients are close in value.

Justifying the underlying regression assumptions

1. All regression models have <u>linear parameters</u>.
2. <u>Random sampling</u> is provided by FRBNY's database in the selection procedure of the reporting individuals. Albeit the selection imposed in the context has limited the number of observations – since the new database has data only regarding unemployed individuals, the assumption of random sampling is not violated as the remaining 551 (or 88) individuals were chosen randomly.
3. Stata automatically inhibits the issue of imperfect collinearity. However, in both models the second and fifth variables (age and education) might be related to each other; as additional

academic and/or professional degrees require years of an individual's life. Being a natural feature of the data, this may have risen the issue of perfect multicollinearity. However, using Stata's VIF command, a tackle to this issue was attempted. Since all VIF values turn to be below 10, multicollinearity is not a worry in this context.

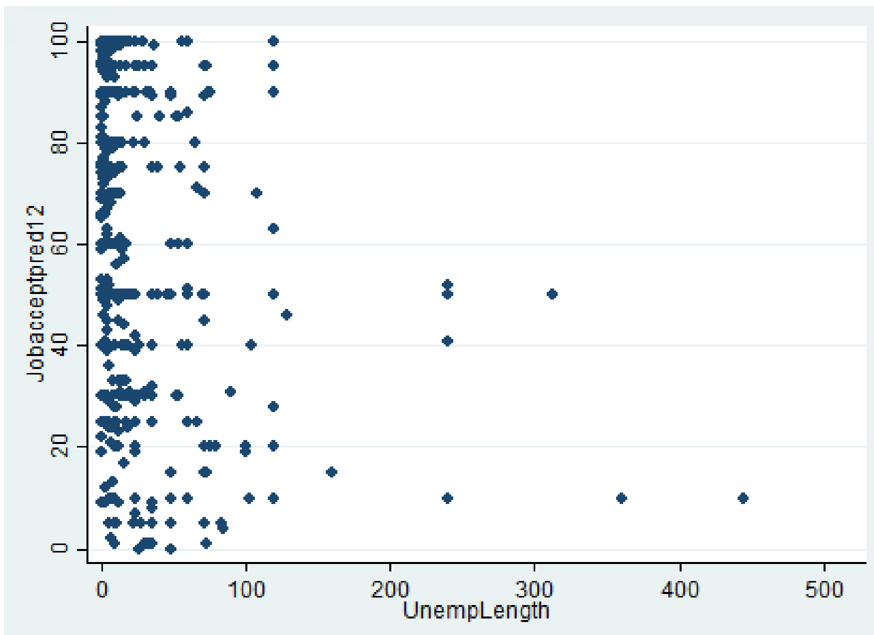
Table 17: VIF of independent variables.

Variable	VIF	1/VIF
Interaction	3.78	0.264611
EducationL~1	3.20	0.312579
Gender	1.81	0.552042
UnempLength	1.11	0.902934
Age	1.07	0.937565
Mean VIF	2.19	

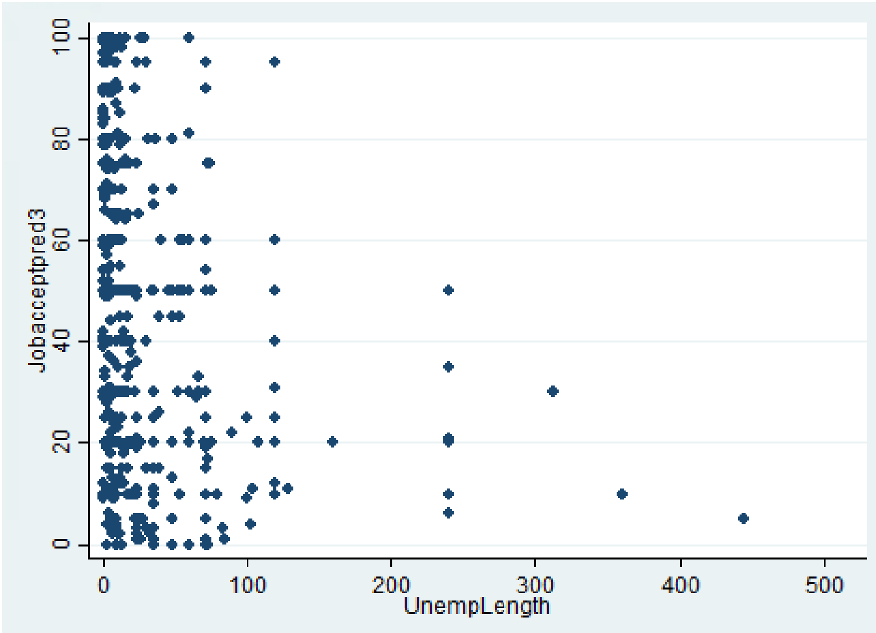
4. In order to construct the model, a significant number of functional forms using other macroeconomic variables (prediction of short-term interest rates, inflation rate, and stocks) were experimented, and the ones with the best fit (highest R-squared values and highest statistical significance) were chosen, as a trial so that bias not exist due to functional form misspecification. Measurement error is also expected to not would we trust the individuals to report the correct numbers; so, the two possible underlying reasons for bias are Reverse Causality and Omitted Variables. The former would exist in the scenario of one's short-term employment expectations determining his or her previous unemployment period, which is impossible since the job market qualifies applicants on skills rather than personal expectations. The latter is possible; that is, in case one's future employment expectations is correlated with his or her intrinsic optimism, bias exists and violates the zero conditional mean assumption. Otherwise, the assumption is presumed to hold.

5. In order to eliminate the issue of heteroskedasticity, the equation was regressed ending with the *robust* option. Scatterplots depicted below also imply homoscedasticity.

Scatterplot of expectation of the chance that the person would find a job and accept it in the following 12 months with respect to the length of unemployment



Scatterplot of expectation of the chance that the person would find a job and accept it in the following 3 months with respect to the length of unemployment



6. Would we assume correctness of the assumptions of unbiasedness and consistency hold, the equations are expected to be BLUE.

Conclusion

Both regression analyses suggest a strong opposite link between one's prediction of future employment and the same person's prior unemployment period. The results are aligned with the ones of the paper "Unemployment expectations, excessive pessimism, and news coverage" by *Marcel Garz*. The paper concludes a similar link between public pessimism with regards to unemployment predictions and pessimism induced from news coverage.

In this experiment, since the same insidious driver factors of incentives and motivations that indicate which group each person belongs in may eventually affect their group's predictions (albeit that by nature both groups are unemployed), a more precise experiment would be to perform the same regression equation in each group and compare the results. Using this database, however, breaking the data into two subgroups would have reduced the t-statistics and goodness-of-fit measures, leading to unreliable test results.

Appendix*

1. Questionnaire of the original database (pdf)
2. Complete (original) database (xlsx)
3. Python codes used for database reduction (pdf)
4. Reduced and manipulated database (xlsx)
5. Stata codes (do)
6. Stata procedure (log)

Works Cited

Armantier, O., Topa, G., Van der Klaauw, W., & Zafar, B. (December 2017). An Overview of the Survey of Consumer Expectations. ECONOMIC POLICY REVIEW,23(2). Retrieved May 11, 2019, from https://www.newyorkfed.org/research/epr/2017/epr_2017_overview-of-sce_armantier.

Bivens, J. (2018, March 28). The fuzzy line between "unemployed" and "not in the labor force" and what it means for job creation strategies and the Federal Reserve. Retrieved May 11, 2019, from https://www.epi.org/publication/the-fuzzy-line-between-unemployed-and-not-in-the-labor-force-and-what-it-means-for-job-creation-strategies-and-the-federal-reserve/

Center for Microeconomic Data. (n.d.). Retrieved from https://www.newyorkfed.org/microeconomics/databank

Definition of "Pure expectations theory" - NASDAQ Financial Glossary. (n.d.). Retrieved May 11, 2019, from https://www.nasdaq.com/investing/glossary/p/pure-expectations-theory

DeSilver, D. (2017, June 27). Perceptions of economy don't always square with economic data. Retrieved May 11, 2019, from https://www.pewresearch.org/fact-tank/2017/06/27/peoples-views-of-their-national-economies-dont-always-square-with-data/

Garz, Marcel, 2013. "Unemployment expectations, excessive pessimism, and news coverage," Journal of Economic Psychology, Elsevier, vol. 34(C), pages 156-168.

National Conference of State Legislatures. (n.d.). Retrieved May 11, 2019, from http://www.ncsl.org/research/labor-and-employment/national-employment-monthly-update.aspx

University of Michigan: Inflation Expectation [MICH], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/MICH, May 12, 2019.