# Job to Job Flows to Predict Market Performance

NICOLE ELCHAAR, Lehigh University, United Statues

While unemployment is widely studied and often considered a lagging economic indicator, transitions between jobs are frequent and may lead the market. In this exploration, we explore the performance of models in predicting market performance from lagging hiring data. While a significant number of lagging hiring times series are found to Granger-cause market performance, we find that there are only slight associations between leading hiring data and lagging market performance with up to a 61.11% accuracy in predicting whether the next quarter's market price will increase or decrease from lagging hiring data alone.

Key Words and Phrases: stock prices, market, S&P500, hiring, unemployment, Granger Causality, time-series, regression, correlation

## 1 INTRODUCTION

In the United States, unemployment rate and market performance have been found to be highly correlated. In this exploration, leading job-to-job flows across the United States in aggregate are compared with lagging S&P500 prices.

## 2 BACKGROUND

### 2.1 The Market and Unemployment

It has been traditionally understood that the stock market is a leading indicator of economic health. Stock prices should reflect short-term, real-time valuations of a company's performance and are much more flexible in than unemployment figures. Unemployment is understood to be a lagging economic indicator because changes in employment send signals of long-term perceptions of recession or growth.[9] Because unemployment figures are released less often, they typically show less volatility than stock prices. We can see directly that unemployment rates are inversely correlated with the stock market and have lagged in response to movements in prices over the past 20 years in Figure 1.

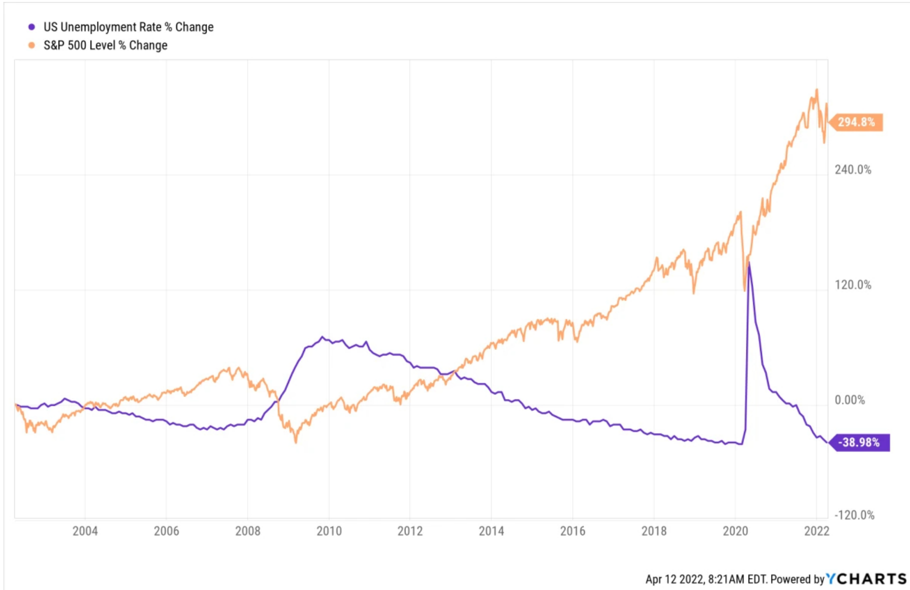



Fig. 1. US unemployment versus the stock market from 2000 to 2022 [7]

Despite this, the release of unemployment figures have been shown to cause volatility in the market. When figures are higher or lower than public expectation, stock prices will sway to adjust as individuals and corporations update their portfolios.[3] In the long term, these adjustments can lead to looser or tighter monetary policies by the Central Bank. Looser monetary policies lower interest rates, triggering a jump in stock prices (when unemployment figures are unusually low and prices are unusually high), while tighter monetary policies increase interest rates, triggering a decline (when unemployment figures are unusually high and prices are unusually low). Stock market prices linked to much more than unemployment figures alone, so while the two may be related cyclically[5], looking deeper into this link is useful.[6]

While the relationship between unemployment rates and the stock market has been widely studied, I thought it would be interesting to look at the problem in reverse, associating hiring data with the stock market. For a variety of reasons, companies are more likely to "over-hire" before upturns than "over-fire" before downturns.[10] As a result, the associations between hiring and the market may be more nuanced than what we find in a single unemployment rate.

Note that I use "lead" and "lag" colloquially to mean a shift in time period between columns of the dataset. In all cases, older data from hiring is shifted towards newer market prices to be used in regressions.

### 2.2 Time Series Analysis

$$P[Y(t+1) \in A \mid I(t)] \neq P[Y(t+1) \in A \mid I_{-X}(t)], \tag{1}$$

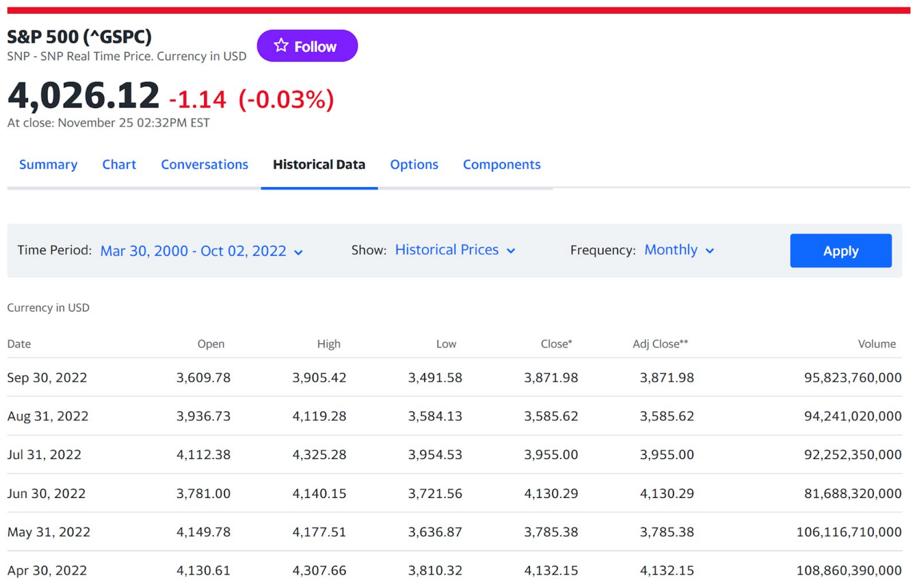
The Granger Causality Test examines causality between two time series.[4] It has been widely used for to economic forecasting, including in S&P500 predictions.[8] In Equation 1, we see the experimental hypothesis, where $P$ is probability, $A$ is an arbitrary non-empty set, and $I(t)$ and $I_{-X}(t)$ are information available as of time $t$. $X$ is said to Granger-cause $Y$ if the hypothesis is accepted.

## 3 DATASET

### 3.1 Job-to-Job Flows

The Job-to-Job Flows dataset (J2J)[1], covers the counts of national, regional, and metro hiring by worker demographic, origin firm demographic, and destination firm demographic from the second quarter of 2000 to the present. When an individual moves from an origin firm to a destination firm, they are included in that quarter's hiring count. Counts are broken down by individuals that switch jobs within the quarter, between adjacent quarters, and from persistent unemployment.

Counts are also broken down into one of 440 aggregation levels. Using the *agg_level* code, we can identify which categorical values are changing and which stay the same. For example, when *agg_level* = 260, workers are separated by *Sex* and *Age* columns, and firms are separated by $NAICSSector$. When *agg_level* = 259393, workers are not separated by their characteristics, and the origin and destination firms are separated by $FirmAge$, $NAICSSector$, and $State$. This makes it essential to filter by aggregation level before performing any analysis. While each may be valuable in forecasting stock price, the many more combinations of $FirmAge$, $NAICSSector$, and $State$ for example would outweigh the three possible categories of *Sex* aggregation (0, 1, or 2 being the only possible characters) alone.




| Date | Open | High | Low | Close* | Adj Close** | Volume |
|---|---|---|---|---|---|---|
| Sep 30, 2022 | 3,609.78 | 3,905.42 | 3,491.58 | 3,871.98 | 3,871.98 | 95,823,760,000 |
| Aug 31, 2022 | 3,936.73 | 4,119.28 | 3,584.13 | 3,585.62 | 3,585.62 | 94,241,020,000 |
| Jul 31, 2022 | 4,112.38 | 4,325.28 | 3,954.53 | 3,955.00 | 3,955.00 | 92,252,350,000 |
| Jun 30, 2022 | 3,781.00 | 4,140.15 | 3,721.56 | 4,130.29 | 4,130.29 | 81,688,320,000 |
| May 31, 2022 | 4,149.78 | 4,177.51 | 3,636.87 | 3,785.38 | 3,785.38 | 106,116,710,000 |
| Apr 30, 2022 | 4,130.61 | 4,307.66 | 3,810.32 | 4,132.15 | 4,132.15 | 108,860,390,000 |

Fig. 2. Yahoo! Finance S&P500 dataset

I selected the quarterly dataset of J2J, using only counts at the national level. As stock market data is not seasonally adjusted, I used only non-seasonally adjusted hiring figures.

A full schema of the dataset is available online.[1]

### 3.2 S&P Index

I gathered monthly S&P index data from Yahoo! Finance[2], selecting the closest date to the first of the month that the market was open. Data goes back much further than the second quarter of 2000, but as I explore how hiring leads the market (rather than vice-versa), pulling data from 2000 and onward is sufficient. Figure 2 shows the resulting page.

## 4 PREPROCESSING

### 4.1 Job-to-Job Flows

I limited the scope of the model to hire columns (rather than looking into separations associated with unemployment) and removed all redundant, null, and constant-value columns. To include the numerical relationship between quarters and years, I converted all years to their year number plus four times the quarter numeric value. This way, there are incremental steps of 0.25 years for the time series.

### 4.2 S&P Index

To match the S&P index data with J2J, I converted market data to use the same periodicity as J2J, eliminating months between quarters. Using an Excel formula, I extracted the closest month to the selected quarters and created a numerical value to match the J2J dataset. Then, with a left join on J2J year values, the time series are connected together in the same dataframe.

## 5 APPROACH

I split the data into an 80/20 training/test split, using the first 80% of the data to train and the last 20% to test. Because the intention of this exploration is to see if stock prices can be predicted by unemployment snapshots, we exclude the most recent data as the set to predict. Being a time series, testing on information within a time period given to the model is unfair. This made the training set the first seventeen years of the model (2000-2017), leaving the last five years (2017-2022) for testing.

### 5.1 Granger Causality Test

I started by using the Granger Causality Test (from $statsmodels.tsa.stattools$) to examine the associations between leading employment data and lagging market performance. For each set of selected parameters by aggregation level, 86 time periods are available from Q2 of 2000 to Q3 of 2022. As both time series are noisy, I selected an $\alpha$ value of 0.05 as a threshold for indicating an association and looked into up to a year and one quarter of lag by aggregation, up to five quarters in total.

To analyze these values, I wrote each successful test (tests with a p-value below 0.05) to a file. In total, there were 12,726 column instances of filtered columns shown to Granger-cause the S&P500 market data as shown in Figure 3. This includes columns that successfully passed any number of the lagging tests, so many column aggregations were duplicated as they were found to be related to the stock market by up to all five lags (though typically only duplicated once or twice for two and three lags, for example).

```
agg_level; selected_categories; selected_combo; column; lag; p
385; ['industry','firmsize']; ['31-33','1']; ENPersistS; 4; 0.032081358562701584
385; ['industry','firmsize']; ['31-33','1']; ENPersistS; 5; 0.018978005183533313
385; ['industry','firmsize']; ['31-33','2']; MJobEnd; 1; 0.02185862512996659
385; ['industry','firmsize']; ['31-33','2']; MJobEnd; 2; 0.046283718333114614
385; ['industry','firmsize']; ['31-33','2']; ENPersist; 1; 0.028752918095733965
385; ['industry','firmsize']; ['31-33','2']; ENFullQ; 1; 0.032549267139874445
269; ['industry','race','ethnicity']; ['51','A2','A2']; sAQHire; 2; 0.008590095406180385
269; ['industry','race','ethnicity']; ['51','A2','A2']; sAQHire; 3; 0.011077445355053842
269; ['industry','race','ethnicity']; ['51','A2', 'A2']; sAQHire; 4; 0.01538694361521197
269; ['industry','race','ethnicity']; ['51','A2','A2']; sAQHire; 5; 0.016900300622217784
269; ['industry','race','ethnicity']; ['51','A2','A2']; sJ2JHire; 2; 0.008590095406180385
```

Fig. 3. Sample export of Granger Causality test, featuring the aggregation level, columns varied under that aggregation, selected values, and performant lags

```
agg_level,count_type,lag
260,MHire,3
260,MJobStart,3
260,EEHire,4
260,J2JHire,4
260,MJobEnd,1
260,MainB,5
260,ENPersist,4
260,ENFullQ,2
260,ENPersistS,1
260,AQHire,1
260,EEHireS,4
260,sEEHireS,1
260,sNEPersistS,1
260,sENPersistS,1
260,sNEHireSEarn_Dest,1
260,NEPersistS,1
260,NEPersist,4
260,NEFullQ,4
260,sAQHire,5
260,sJ2JHire,5
260,sNEPersist,5
```

Fig. 4. Selection of the best lag values for aggregate level 260

Afterwards, I found the unique aggregation types associated with these models and chose the best lag per column. I rebuilt the dataset, shifting columns according to their best performing lead time. This leaves a dataset with the predicted value for S&P500 close price in the correct year column and the columns most likely to cause prices shifted by the specified number of quarters. In Figure 4, for example, we see that MJobEnd (end of main jobs due to separations and instances when another job becomes the main source of earnings) and ENPersistS (separations to persistent unemployment) should lead the market by only one quarter for the best performance. MHire (hires into a worker's main job) and MJobStart (new main jobs due to hires and instances when a previously existing secondary job becomes the main source of earnings) should lead by three quarters, and EEHire (hires following a separation with no observed unemployment spell, continuous employment) and J2JHire (hires following a separation, short or no observed unemployment spell) should lead by four quarters.

## 5.2 Regression

Once columns and aggregation types are selected as possible predictors of market prices, I normalized the numeric columns with min-max scaling to bring all columns' range between zero and one. I then selected four linear regressions to understand numerical performance. Because linear regressions do not require huge amounts of data to converge, it made sense to do on many columns and time series, especially with few time periods per time series.

Starting with a standard linear regression as a baseline, we also have a Ridge regression, Lasso regression, and ElasticNet regression, all of which are evaluated with root mean squared error and $R^2$ score.

$$Penalty_{ridge} = \alpha||\theta||_2^2 = \alpha\theta^T\theta = \alpha\sum_{i=1}^{m}\theta_i^2 \tag{2}$$

I selected the Ridge regression because the hiring columns are highly related to one another. Because of issues of multicollinearity in the standard linear regression on this dataset, it should outperform the baseline. Ridge regression uses L2 regularization to with an L2 loss function to prevent overfitting, as in the *Penalty* term of Equation 2, where $\theta$ is the parameter vector. Using a squared penalty term, it decreases the coefficients of large model parameters more than smaller ones.

$$Penalty_{lasso} = \alpha ||\theta||_1 = \alpha \sum_{i=1}^{m} |\theta_i| \tag{3}$$

The Lasso method (Least Absolute Shrinkage and Selection Operator) is similar to the ridge regression but uses L1 regularization to perform feature selection as shown in Equation 3. Rather than squaring, the lasso regression takes the absolute value in the loss function, leading some parameters to be removed in the regression. With over twenty columns of interest, it will be useful to limit those used in the regression by zeroing them out.

$$Penalty_{elasticnet} = \alpha_1 \sum_{i=1}^{m} |\theta_i| + \alpha_2 \sum_{i=1}^{m} \theta_i^2 \tag{4}$$

Lastly, we try ElasticNet. ElasticNet should perform the best as it combines L1 and L2 loss from lasso and ridge regressions, respectively. From Equation 4, we see that setting either $\alpha$ value to zero will result in an equation equivalent to lasso or ridge regression.

### 5.3 Additional Evaluation

As stock prices increase or decrease relative to their last period's performance (and are somewhat dependent on the previous performance), I also converted absolute predictions to relative percent differences between the current period and last period. As my model predicts stock prices up to one period before the previous, finding the relative difference from a current level should lead to a more flexible model. These percentages also serve as a method of predicting the next stock price directly from the current stock price and updating predictions relative to any shocks in the market.

Lastly, I included another evaluation metric of whether or not the model was correct in predicting an increase or decrease from the previous quarter's price. This binary value makes it possible to compute a rough accuracy of the model.

## 6 RESULTS & DISCUSSION

Of the 12,726 columns extracted from the Granger Causality Test, there are 7 unique aggregation types. All of them happen to include the *NAICSSector* in the breakdown, as displayed in Table

| Aggregation | Median p-value |
|---|---|
| Ethnicity, NAICS Sector | 0.016906 |
| Race, Ethnicity, NAICS Sector | 0.017220 |
| Race, NAICS Sector | 0.018018 |
| Age, NAICS Sector | 0.018392 |
| Sex, Age, NAICS Sector | 0.019403 |
| Firm Size, NAICS Sector | 0.020319 |
| Sex, Education, NAICS Sector | 0.020496 |

Table 1. Highest performing aggregation levels in predicting stock market performance from the Granger Causality Test

| Regression Type | $R^2$ |
|---|---|
| Ridge | -1.33 |
| Lasso | -0.80 |
| ElasticNet | -6.47 |
| Linear | -6.06 |

Table 2. Best $R^2$ score by regression type

| Regression Type | RMSE ($) |
|---|---|
| Ridge | 309.37 |
| Lasso | 299.47 |
| ElasticNet | 540.92 |
| Linear | 2058.70 |

Table 3. Best RMSE score by regression type

We also see in Figure 4 a selection of the best lags by aggregation level and column.

A summary of result statistics is shown in Table 2. Because the training and testing samples are very different, all $R^2$ values in this exploration are less that zero. This said, the lowest $R^2$ value was -0.80 from the best performing Lasso regression. The ElasticNet regression ended up being much worse than Ridge and Lasso, even underperforming against the standard linear regression.

Next, we look into root mean squared error in Table 3. This metric is useful because it is in dollars and can be directly compared to the market price. Within the period, the market price ranged from a low of $2,300 to a high of $4,700. Again, we see that the

| Regression Type | Percent Difference $R^2$ |
|---|---|
| Ridge | -0.77 |
| Lasso | -1.35 |
| ElasticNet | -3.88 |
| Linear | -53.32 |

Table 4. Best $R^2$ by regression type from the percent difference from the prior period

| Regression Type | Accuracy in Increase or Decrease (%) |
|---|---|
| Ridge | 61.11 |
| Lasso | 55.56 |
| ElasticNet | 50.00 |
| Linear | 55.56 |

Table 5. Best accuracy in whether the model predicted an increase or decrease from the prior period by model type

Lasso regression performed best on the data. Because so many parameters are related, the normalization was helpful with Ridge, Lasso, and ElasticNet. With the linear regression, however, there are drastic jumps associated with the pandemic as shown in the Appendix in Figures 8 to 14. In the beginning of the pandemic, hires drastically decreased for hiring within and between quarters but drastically increased for hires from persistent unemployment (See Figure 5, Figure 6, and Figure 7). Without normalization, the linear regression predicted a huge spike in the stock market. This pattern also was not found in the rest of the data and would be useful to look into for longer time periods. Though the data is not available before 2000, a similar spike was said to occur in 1949, 1974, and 1983 which would be useful in training.[3]

Rather than isolating the dollar amount, we can also see performance when looking at the percent increase that the model predicts from the prior period. In Table 4, we again see that the normalization provided by Ridge and Lasso regressions outperform ElasticNet, all of which significantly outperform the linear regression. The slightly lower $R^2$ values from percent differences indicate how adding an additional time-based parameter to this model could help with training. Extending by adding a lagging prior quarter's close to the model would likely give the best result.

Now, we examine whether the model predicts increases or decreases from the prior period. Because the time series is dependent and shocks to the system at an early time lead to much worse performance, we can also examine whether the model was successful at predicting a binary increase or decrease in market performance. In Table 5, we see that the Ridge regression was best at predicting whether the price would increase or decrease over the coming quarters, followed closely by Lasso and linear regressions. ElasticNet did not perform better than pure guessing.

With the final regressions in Figure 8 through Figure 14, we see why the linear regression performed so poorly and how normalization helps. The huge spikes in employment from persistent non-employment and great associations between columns (See Figures 5, 6, and 7) indicate the multicollinearity impacting this regression.

## 7 CONCLUSION

While none of the models give very high accuracy, market prices are dependent on all sorts of factors that were not included in this exploration. Prices that is somewhat, but not entirely, dependent on the previous period's closing price would have likely benefited from a model that incorporated this lag along with hires. In addition, the rapid changes to hiring that came about in the pandemic led to much more instability in the model. With the current linear model, we can fit a line to the training and testing data with an $R^2$ value of only 0.80. A different model, which weighs values closer to the current predicted time period may be a better choice. In addition, excluding data from the pandemic (and looking before and after this time) may also bring about a better model. Possible future work is explored more in depth below.

All in all, we see that hiring data can be used to better predict market performance in quarters in the near future. Aggregations are far closer to truth values in the beginning of forecasting than towards later periods, and percent differences (adjusting for unexpected shocks to the market) shows better performing models. We also see that considering multicollinarity is very important in this model. When building future models and incorporating other economic indicators, the relationship between these columns should be considered.

# 8 FUTURE WORK

## 8.1 Averaging

Stock price data was selected at a particular moment in time. Instead, we can average the price over the 90 days to be consistent with the cumulative counts of hiring and reduce short-term noise that may have impacted the model.

## 8.2 Time-Based Weights

Rather than predicting all five years into the future, incorporating more recent data with greater weight would be useful in this model. While we only rely on seventeen years of data in this case, a time-based weight could help to generalize to longer-term models.

We also can limit the exploration (or evaluation) to a single period leading time if we only intend to predict the next period. As lags are at a minimum one period (one quarter) apart, updating the next prediction relative to the current real stock price would give a stronger model. As new information is given at every new time period, we can update the predicted price relative to any already-understood shocks in the market for a true performance evaluation.

Alternatively, we can train the model using a single period lagging close price as a parameter. This should make the model more flexible in spikes. We can then iteratively pass in the previous quarter's prediction as the next prediction's lagging "Close" price such as is done in a recurrent neural network.

## 8.3 The Pandemic

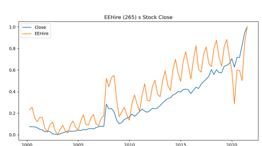



Fig. 5. Hiring from within the quarter versus the market

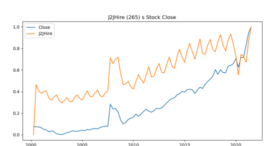



Fig. 6. Hiring from between quarters versus the market

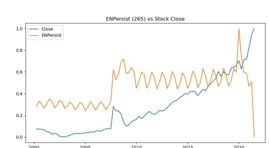



Fig. 7. Hiring from persistent unemployment versus the market

Because the time period of this data's release, much of the data from the pandemic is included in the testing set with the patterns of the pandemic missing from the training set. Because of the drastic changes in hiring and systemic changes with remote work, separating data from before and after the pandemic may be helpful in future explorations.

## 8.4 Particular Demographics

Rather than simply isolating the particular columns that Granger-cause the market, it is possible to filter particular instances of columns and see if one *Age* group (or all but one *Age* group), for example, are more predictive of stock prices. Looking into the way particular demographics are hired or fired, rather than simply at the aggregation level, would be an interesting extension of this model.

## 8.5 Predicting Hiring from Stock Market Data

As widely discussed in literature, it is more obvious to predict hiring data from stock market data than was done in this exploration. Looking into the associations between leading stock prices and lagging hires shows nearly three times the aggregation levels and one and a half times the columns that pass the Granger Causality Test with an $\alpha$ level below 0.05. Switching the leading and lagging variables may make a model that makes more theoretical sense and is more performant.

## 8.6 Additional Regression Types

It would be valuable to look into other regression types instead of attempting to make only strict linear associations between hiring and market performance. Using decision trees to find simpler ways of splitting the data is a potential additional area of study.

## 8.7 Additional Data

Lastly, using a multi-modal approach to predict stock price is likely the best place to go to build upon the existing model. While we find some associations in hiring data leading the market, incorporating other pieces of data would likely lead to a much more stable and higher performing model. Quarterly hiring data alone is not the only factor associated with stock performance.

## REFERENCES

[1] 2022. Longitudinal Employer-Household Dynamics: Job-to-Job Flows, R2022Q3. https://lehd.ces.census.gov/data/schema/j2j_latest/lehd_public_use_schema.html#_identifiers_for_j2j

[2] 2022. SP500 (Ticker ĜSPC) from March 30, 2000 to October 2, 2022.

[3] BetterTrader.co. 2021. Case study - unemployment rate and its effect on the SP 500: It's less intuitive than you think. https://bettertrader.co/case-studies/unemployment-rate-and-its-effect-on-the-SP-500.html

[4] Michael Eichler. 2012. Causal inference in time series analysis. *Causality* (2012), 327–354. https://doi.org/10.1002/9781119945710.ch22

[5] Mete Kilic and Jessica A Wachter. 2018. Risk, Unemployment, and the Stock Market: A Rare-Event-Based Explanation of Labor Market Volatility. *The Review of Financial Studies* 31, 12 (01 2018), 4762–4814. https://doi.org/10.1093/rfs/hhy008 arXiv:https://academic.oup.com/rfs/article-pdf/31/12/4762/26718599/hhy008.pdf

[6] Wei-Fong Pan. 2018. Does the stock market really cause unemployment? A cross-country analysis. *The North American Journal of Economics and Finance* 44 (2018), 34–43. https://doi.org/10.1016/j.najef.2017.11.002

[7] Chris Preston. 2022. The unemployment-stock market correlation in one chart. https://www.cabotwealth.com/daily/stock-market/unemployment-stock-market-correlation-one-chart

[8] Barbara Rossi. 2013. *Handbook of Economic Forecasting*. Elsevier.

[9] Derek Thompson. 2009. Does unemployment guide the stock market? https://web.archive.org/web/20121127210735/http://www.theatlantic.com/business/archive/2009/05/does-unemployment-guide-the-stock-market/18306/

[10] Laurence M Vance. 2017. The right to hire and fire. https://www.fff.org/explore-freedom/article/right-hire-fire/

# A APPENDIX

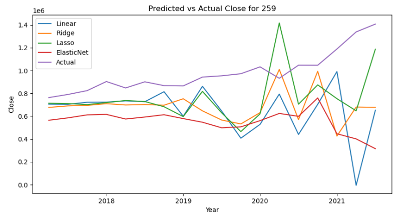



Fig. 8. Predicted vs Actual Close for Aggregations on Worker Age and Destination Firm NAICS Sector

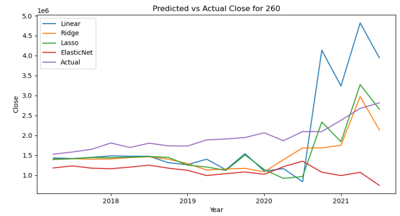



Fig. 9. Predicted vs Actual Close for Aggregations on Worker Age, Worker Sex, and Destination Firm NAICS Sector

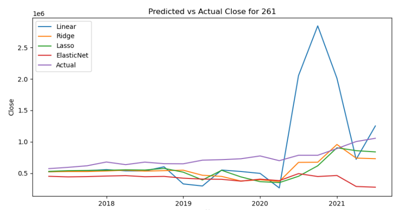



Fig. 10. Predicted vs Actual Close for Aggregations on Worker Race and Destination Firm NAICS Sector

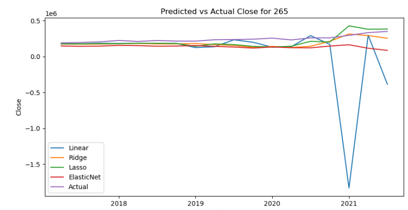



Fig. 11. Predicted vs Actual Close for Aggregations on Worker Ethnicity and Destination Firm NAICS Sector

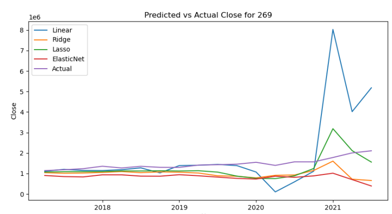



Fig. 12. Predicted vs Actual Close for Aggregations on Worker Race, Worker Ethnicity, and Destination Firm NAICS Sector

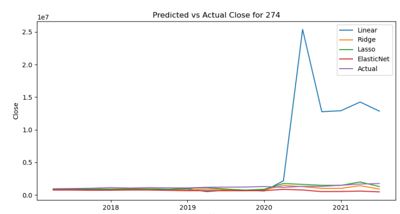



Fig. 13. Predicted vs Actual Close for Aggregations on Worker Sex, Worker Education Level, and Destination Firm NAICS Sector

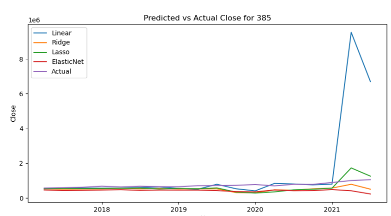



Fig. 14. Predicted vs Actual Close for Aggregations on Destination Firm Age and NAICS Sector