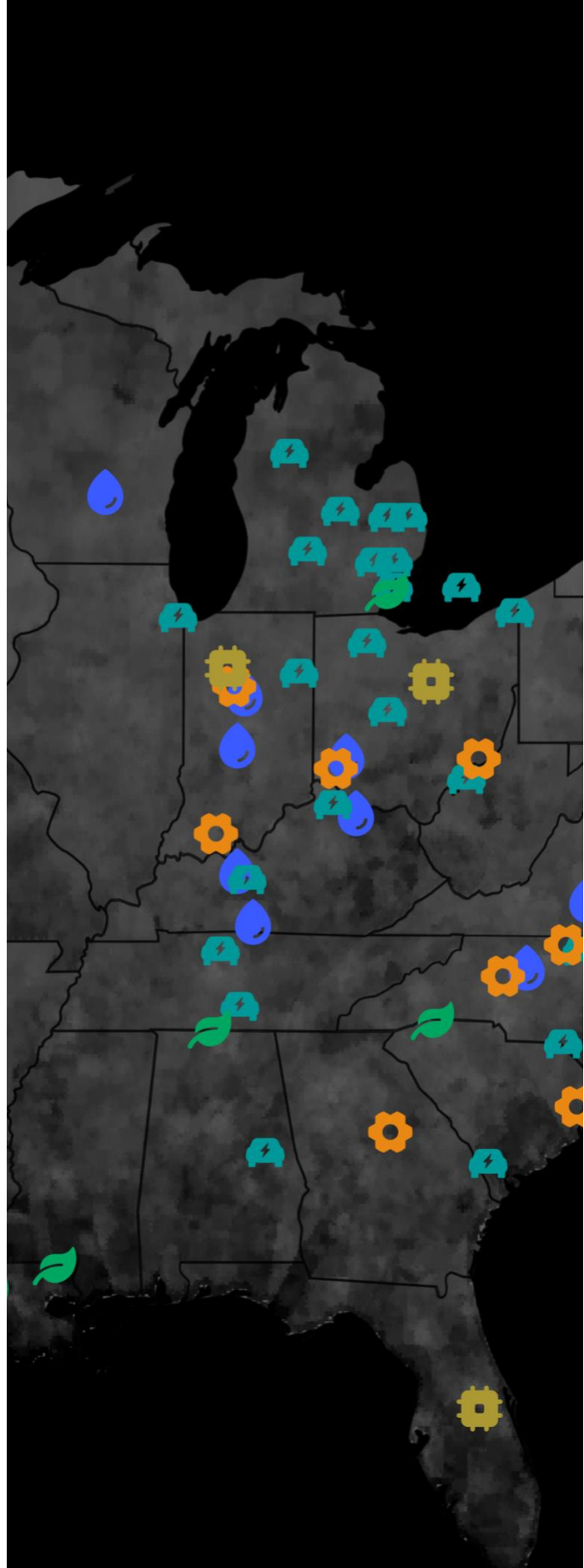


Tracking America's Industrial Renaissance

May 2024 (v1)



Section I: Introduction

In August 2022, the Biden Administration passed the Inflation Reduction Act (IRA), marking a historic moment in American policymaking. This landmark bill primarily targeted energy security and climate resilience as themes for investment and capital deployment. To date, the IRA introduced over 20 tax incentives for clean energy and manufacturing to stimulate domestic supply chains, reduce household energy expenses, and curtail greenhouse gas emissions while providing employment opportunities. Notably, the legislation aimed to extend clean energy tax incentives to tax-exempt entities and incorporate provisions to bolster supply chain resilience (The White House). The U.S. Department of the Treasury, in collaboration with the Internal Revenue Service (IRS), has played a pivotal role in implementing these provisions, which include tax incentives, corporate tax reforms, and efforts to enhance IRS customer service and modernize technology since the law's inception (US Department of Treasury).

As a result of the IRA, the US economy has seen a massive uptake in private and public investment in clean energy, semiconductors, EV batteries and biomanufacturing to name a few. Large organizations like General Motors started building plants and investing more in clean energy and semiconductors. There was also an increase in foreign investment into the US targeting these industries. The White House started tracking these private investments, highlighting the investment amount, company, and industry ([The White House](#)). Figure 1 shows the investment flow by the top 15 companies¹ (selected by frequency) as tracked by The White House.

¹ This map represents the data last updated at source on 1st April 2024

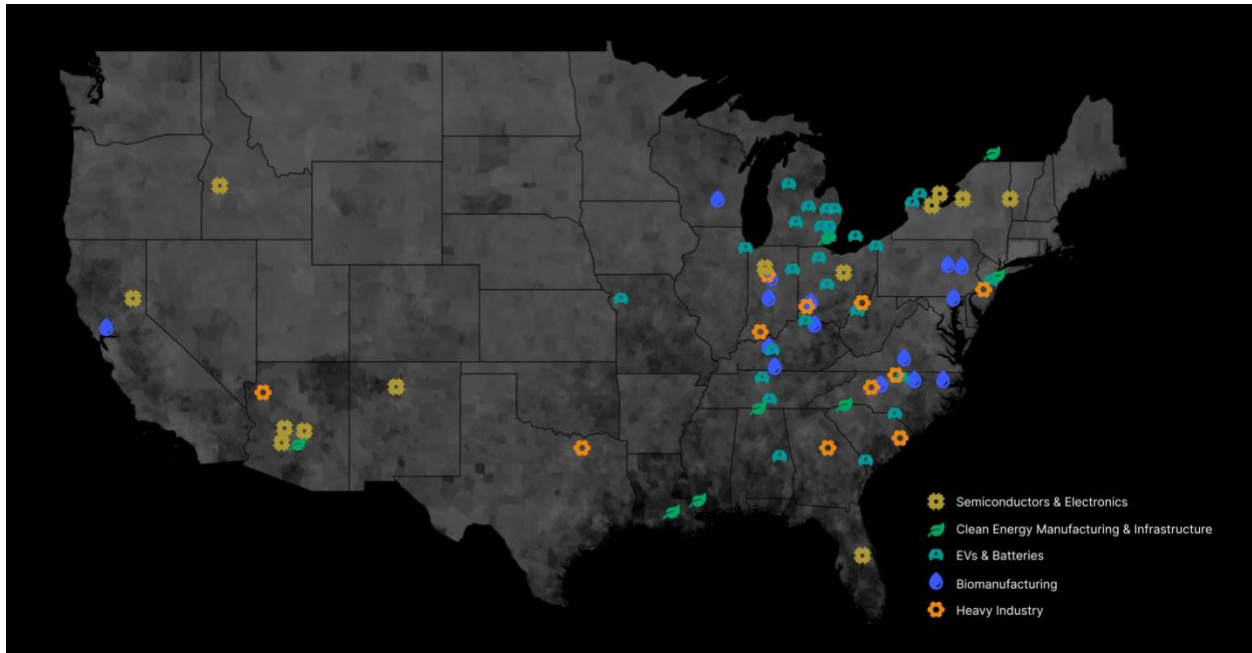


Fig 1. Top 15 companies' investment flow

With these massive thematic changes, investors need new tools to evaluate where and how investment will unfold in both the private and public markets. [AlphaGeo](#), a groundbreaking company specializing in risk and resilience, has initiated a thorough examination of private investment trends to explore two primary themes.

First, we aim to evaluate the financial performance of publicly listed firms across diverse geographic regions. Second, we seek to determine whether companies investing in resilient locations demonstrate superior performance compared to others over both short- and medium-term intervals. 'Resilient locations' is a broad term encompassing many dimensions. At AlphaGeo, we define resilient locations as geographies that may benefit from disruption or changes in the natural or human-caused environment. Lastly, we ask ourselves whether our findings can be generalized, using the data and the methodology we are using.

The overarching objective of this paper is not to provide definitive conclusions, but rather to stimulate discussion on the role of geography and socio-economic factors in investment strategies.

Section II: Research Questions

Our research aims to assess the following questions:

Question #1: How can AlphaGeo's data and techniques be applied to assess the financial performance of publicly listed companies based on resilience factors (per our framework) of investments with respect to specific geographic locations?

Question #2: Do public companies that are deploying capital in resilient geographic locations perform better than others, across short- and medium-term investment durations?

Question #3: Are there certain thematic drivers that can be generalized as a result of the data we are collecting coupled with the methods we are applying?

Section III: Methodology

We structured our research into two distinct phases. In the first phase, we focused on establishing correlations between socio-economic and climate indicators/variables with financial data for our top 15 public limited companies selected.

In the second phase, we undertook several analytical exercises. Initially, a small-scale location analytics exercise was conducted, followed by the implementation of a time-

series Ordinary Least Square (OLS) regression model on specific indicators. One of the objectives of this research was to assess the predictive capabilities of AlphaGeo's socio-economic and climate data on financial indicators. To test this objective, a random forest regressor model was employed for a subset of publicly traded companies, evaluating how our selected variables influenced a financial indicator. This section will provide a comprehensive breakdown of our methodology and the specific indicators under scrutiny.

PHASE 1.0

We began our analysis by identifying the geographic coordinates for the top 15 companies as selected by a Natural Language Processing (NLP) algorithm applied to data tracking planned capital allocation resulting from IRA (and other related) capital commitments. We used the NLP algorithm to count the number of times a company's name appeared in the White House list and sorted by frequency to generate the top 15 list (Table 1). Source data tracking announcements and deal flow is maintained by [The White House](#). After examining the companies in the White House database, we pulled select financial information for the top 15 firms as defined by our NLP algorithm. *Financial indicators included share price, earnings per share (EPS) and earnings before interest, taxes, depreciation, and amortization (EBITDA) by quarter and year (2015-2023).*

TABLE 1. Top 15 Companies

Company	Industry
Air Products	Clean Energy
Catalent	Biomanufacturing
First Solar	Clean Energy
Ford	EV & Batteries
General Motors	EV & Batteries
Intel	Semiconductors & electronics
Magna	EV & Batteries
Microchip	Semiconductors & electronics
Micron	Semiconductors & electronics
Nucor	Heavy Industry
Pfizer	Biomanufacturing
PS&EG	EV& Batteries and clean energy
Thermo Fisher	Biomanufacturing
Toyota	EV & Batteries
WEC Energy Group	Clean Energy

Table 2 contains all socio-economic indicators used in our analysis and Table 3 contains climate risk data collected by AlphaGeo. Indicators in Table 2 are obtained from various sources including The White House, ACS Census data, Bureau of Transportation, and others.

Table 3 indicators are curated by AlphaGeo after extensive research, data analysis and Machine Learning (ML) modeling. Data in Table 3 reflects historical climate change trends, but its lack of time-series format limits its current applicability.

The original socio-economic data was on different geographic levels, mostly on county and census tracts. The climate data curated by AlphaGeo was on H3². However, all data was aggregated to ZIP Code level for analysis. We used either a python library or the crosswalk files provided by Housing and Urban Development (HUD) to aggregate the data to the same geography. We used the most recent year data (not older than 2020) to run the regression and correlation analysis.

TABLE 2. Socio-economic Indicators (Cross-sectional)

Indicator Name	Description	Year
Jobs	Total jobs (White house data)	NA
Investment Amount Millions (USD)	Total amount of investment in million USD (White House data)	NA
Number of Airports	Number of airports	2020
Percent Bridges in good/fair condition	Percentage of bridges in good or fair conditions	2021

² “The H3 geospatial indexing system is a discrete global grid system (see [Sahr et al., 2003](#)) consisting of a multi-precision hexagonal tiling of the sphere with hierarchical indexes” [Link](#).

VMT	Vehicle miles travelled	2020
Emission Rate	Annual CO2 equivalent total output emission rate in lb./MWh	2021
Renewable Gen	Percentage of the total renewable generation	2021
Non-renewable Gen	Percentage of the total nonrenewable generation	2021
Total Power Outage 1-Hour	Average of the total number of outages annually for at least 1 hour	2020
Total Power Outage 8-Hour	Average of the total number of outages annually for at least 8 hours	2020
Avg Cust Out	Annual average county-level customers without power	2020
Power Loss Per Capita	Annual average county-level number of hours without power per customer	2020
HH Median Income	Annual median household income	2021
House Price Income Ratio	Ratio of house price and household income	2021
Income Growth Rate	The growth rate of median household income over the past year	2021

Employment Rate	Percentage of employed individuals to the total civilian labor force	2021
Percent Bachelors' Above	Percentage of the population over 25 with a bachelor's degree	2021
Percent Doctorate Above	Percentage of the population over 25 with a doctoral degree	2021
Vulnerable Pop	Percentage of population 65 years and above	2021
Poverty	Percentage of families in the ZCTA area living under the poverty level	2021
GINI Index	Measures inequality on a scale from 0 to 1, where higher values indicate higher inequality	2021
Percent Insured Pop Pvt	Percentage of the population with private health insurance	2021
Percent Insured Pop Public	Percentage of population with public or state health insurance	2021
Household Debt	Household debt-to-income ratio	2021

TABLE 3. Climate Indicators (Cross-sectional)

Indicator Name	Year
----------------	------

Average Wind Score	2023
Average Inland Flooding Score	2023
Average Heat Score	2023
Average Drought Score	2023
Average Coastal Flood Score	2023
Average Fire Risk Score	2023

Based on the indicators above, we ran correlation analyses and cross-sectional OLS regression models to understand the relationship between the climate and socio-economic indicators with actual earnings per share (EPS). The results will be discussed in section IV.

PHASE 2.0

Table 4 contains all the socio-economic indicators used in this phase of the study. We dropped the climate indicators because the curated indicators provided by AlphaGeo were not time-series. The investment amount is taken directly from the White House data. Quarterly data for socio-economic indicators was calculated by dividing the yearly value by 4 for each year.

TABLE 4 Socio-economic Indicators (Time-series)

Indicator Name	Year
Investment Amount (Million USD)	NA
Household Median Income	2016-2021
House Price Income Ratio	2016-2021
Income Growth Rate	2016-2021
Employment Rate	2016-2021
Percentage of Population with Bachelor Above	2016-2021
Percentage of Population with Doctorate Above	2016-2021
Vulnerable Population	2016-2021
Poverty	2016-2021
GINI Index	2016-2021
Percentage of Population Insured Publicly	2016-2021
Percentage of Population Insured Privately	2016-2021
Household Debt	2016-2021
Percentage of Bridges in Good Condition	2016-2021

Using these indicators, we conducted a small *location analytics experiment*. If we consider domain knowledge and state policies of these ZIP Codes spread across the US, we know that states perform differently. We selected 4 ZIP Codes in 4 different states and ran a basic correlation between financial indicators and socio-economic indicators. The results are discussed in section IV.

Next, we ran a *timeseries OLS regression model* where our dependent variable was Actual EPS. The results are discussed in section IV.

Lastly, we conducted a *random forest regressor model* for our top 15 companies to see how our indicators estimate Actual EPS. To put it simply, a random forest regressor model is an ML algorithm that uses a collection of decision trees to make predictions about numerical (continuous) outcomes. It combines the predictions of multiple individual trees to produce overall prediction. The basic parameters of our random forest regressor model are:

- Based on indicators listed in Table 4
- 80% of the data is used for training and 20% for testing
- For time-series data we get a range of values because the data is location specific, i.e., for each of the 50 locations we have data from 2016-2021, hence there are rows where financial data for the company is the same for the said quarter regardless of location, but socio-economic data differs
- We replace missing values with variable's mean
- The date selected to be tested was 2021, Q4. We substitute the Actual EPS reported value with NA (missing) so that the model can predict the value and compare it to the reported value.

The results are discussed in section IV.

Section IV: Data Analysis

This section will discuss the results of the data analysis in each phase.

PHASE 1.0 Results

Two *correlation analyses* were run to examine the relationships between (a) Investment flow and the socio-economic and climate indicators (Figure 2) and (b) EPS and socio-economic and climate indicators (Figure 3).

In Figure 2, we observe significant positive correlation between investment and percentage of population privately insured, percentage of population with at least a bachelor's degree or above, household median income, and a significant positive correlation with availability of renewable energy sources. These relationships indicate that investment is flowing into areas that have more favorable socio-economic profiles. When we assess financial performance (via EPS) we observe a significant positive correlation with wind and household debt, while other factors demonstrate weaker but still significant correlations.

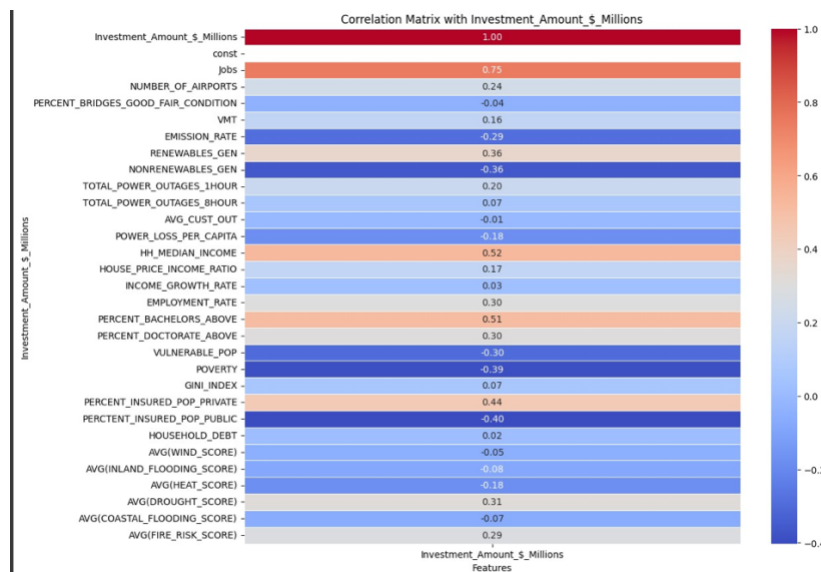


Fig 2.

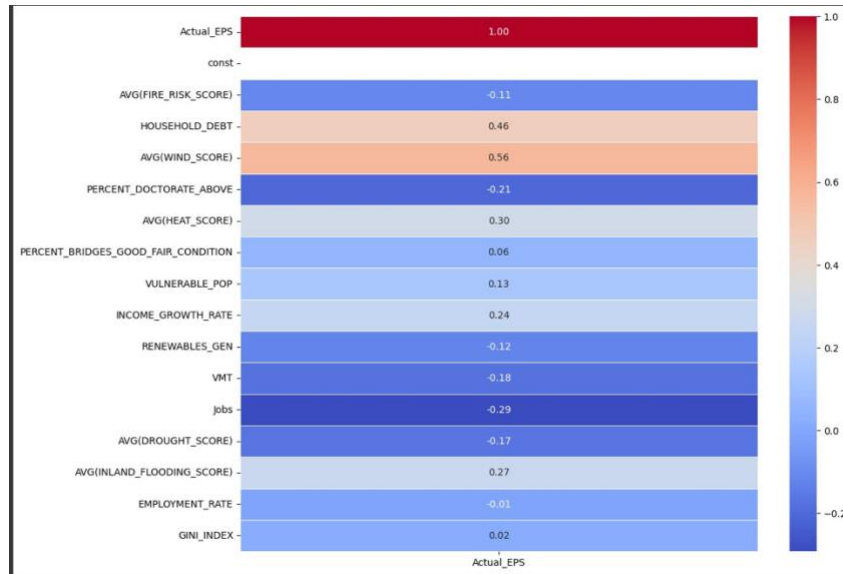


Fig 3.

Next, we ran *cross-sectional OLS regression models* to estimate regression coefficients to describe the relationship between one or more independent quantitative indicators and the selected dependent variable. We began our analysis by running the regression models with only social indicators against Actual EPS as the dependent variable. The second experiment included the deconstructed climate risk indicators, and the third experiment applied both socio-economic indicators and aggregated climate scores for the regression analysis.

The first two experiments did not generate any considerable insights; however, the third experiment was promising. The results for the third experiment are discussed below.

Model Results:

While Model 1 (M1) (Figure 4) demonstrates a strong R^2 (0.989), indicating that the independent variables significantly impact the dependent variable (Actual EPS), this is not effective for two reasons: first, due to multicollinearity, the independent variables are not statistically significant in this model, and second, there are too many independent variables given the sample size.

```

=====
                    OLS Regression Results
=====
Dep. Variable:      Actual_EPS      R-squared:          0.989
Model:              OLS              Adj. R-squared:     0.664
Method:             Least Squares    F-statistic:        3.041
Date:               Thu, 23 Nov 2023  Prob (F-statistic): 0.429
Time:              16:40:59          Log-Likelihood:     -2.4760
No. Observations:  31              AIC:                64.95
Df Residuals:      1              BIC:                108.0
Df Model:          29
Covariance Type:   nonrobust
=====
                    coef      std err      t      P>|t|      [0.025      0.975]
-----
const                0.0257      0.012      2.135      0.279      -0.127      0.179
Investment_Amount_  0.0004      0.001      0.859      0.548      -0.006      0.007
$ _Millions
Jobs                -0.0024      0.002     -1.423      0.390      -0.024      0.019
NUMBER_OF_AIRPORTS  0.2554      0.512      0.498      0.706      -6.256      6.767
PERCENT_BRIDGES_GOOD_FAIR_CONDITION -0.8611      0.323     -2.664      0.229      -4.968      3.246
VMT                -7.81e-06      0.000     -0.019      0.988      -0.005      0.005
EMISSION_RATE      -0.0073      0.006     -1.213      0.439      -0.084      0.069
RENEWABLES_GEN      1.1571      0.571      2.027      0.292      -6.898      8.412
NONRENEWABLES_GEN  1.4128      0.635      2.225      0.269      -6.654      9.480
TOTAL_POWER_OUTAGES_1HOUR  0.0746      0.050      1.490      0.376      -0.561      0.710
TOTAL_POWER_OUTAGES_8HOUR -0.8478      0.679     -1.249      0.430      -9.474      7.778
AVG_CUST_OUT       -2.33e-06      2.1e-06     -1.111      0.466     -2.9e-05      2.43e-05
POWER_LOSS_PER_CAPITA -0.1602      0.299     -0.535      0.687      -3.963      3.642
HH_MEDIAN_INCOME   -0.0003      0.000     -1.915      0.306      -0.002      0.002
HOUSE_PRICE_INCOME_RATIO  2.2177      1.109      2.000      0.295     -11.873     16.309
INCOME_GROWTH_RATE  0.2144      0.319      0.671      0.624      -3.844      4.273
EMPLOYMENT_RATE    -0.0116      0.444     -0.026      0.983     -5.653      5.629
PERCENT_BACHELORS_ABOVE  0.6906      0.482      1.432      0.388      -5.436      6.817
PERCENT_DOCTORATE_ABOVE -6.7842      4.042     -1.679      0.342     -58.140     44.571
VULNERABLE_POP    -0.4259      0.311     -1.371      0.401      -4.373      3.521
POVERTY            -0.4242      0.339     -1.253      0.429     -4.725      3.877
GINI_INDEX         0.7923      17.131      0.046      0.971     -216.882     218.466
PERCENT_INSURED_POP_PRIVATE -0.3893      0.302     -1.291      0.420     -4.220      3.442
PERCENT_INSURED_POP_PUBLIC -0.3790      0.366     -1.037      0.488     -5.024      4.266
HOUSEHOLD_DEBT     0.1612      1.847      0.087      0.945     -23.311     23.633
AVG(MIND_SCORE)    -0.0126      0.043     -0.296      0.817     -0.554      0.529
AVG(INLAND_FLOODING_SCORE)  0.0242      0.100      0.243      0.848     -1.241      1.289
AVG(HEAT_SCORE)    -0.0902      0.061     -1.484      0.378     -0.863      0.682
AVG(DROUGHT_SCORE)  0.0802      0.077      1.041      0.487     -0.898      1.059
AVG(COASTAL_FLOODING_SCORE)  0.0359      0.149      0.241      0.849     -1.854      1.926
AVG(FIRE_RISK_SCORE) -0.0098      0.051     -0.193      0.879     -0.659      0.639
=====
Omnibus:           9.362      Durbin-Watson:      2.373
Prob(Omnibus):     0.009      Jarque-Bera (JB):   10.670
Skew:              0.715      Prob(JB):            0.00482
Kurtosis:          5.493      Cond. No.            2.08e+17
=====

```

Fig 4. MODEL 1 (M1) SUMMARY STATISTICS

To address these issues, we deployed a Random Forest Feature Importance model, and utilized the top 50% of the indicators to create Model 2 (M2). The selected indicators from the Random Forest exercise are shown in Figure 5.

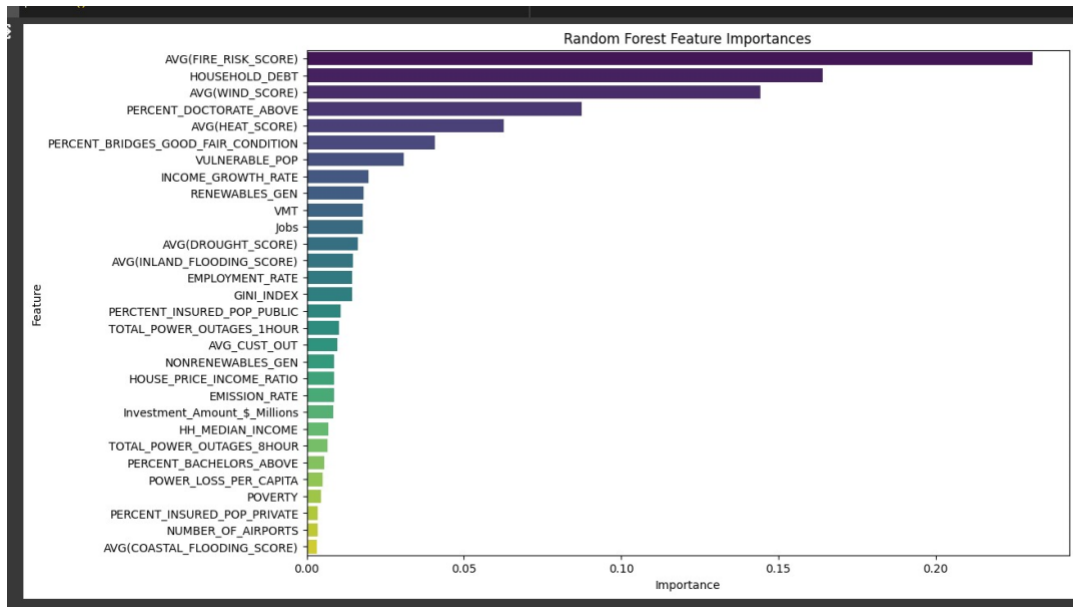


Fig 5.

The performance of the revised M2 decreases when compared to M1 ($R^2 = 0.99$ to 0.74); however, this is a more appropriate model to select (Figure 6). M2 identifies several key independent variables which are now statistically significant at the 90% confidence interval (CI) (0.1) and, given the nature of the training data, a $0.74 R^2$ is a favorable statistic.

The results in M2 point to a few interesting factors to consider as we move to create market tracking indices using AlphaGeo's data. First, the model selects climate and non-climate indicators that are noncorrelated, protecting against overfitting the variables to the outcome, as mentioned above. Second, the model's physical and financial variable selection identifies complementary indicators, which make model interpretation and efficacy intuitive and logical. For example, stronger indicators related to education, infrastructure and renewable energy potential, and income distribution equality correlate with geographies where selected companies exhibit higher earnings (via EPS) and therefore favorable financial performance. Finally, these assumptions can be replicated across geographic regions, so the indices have

a higher potential to be applied to other operating sectors and countries, where appropriate data exists. This gives higher confidence to out-of-sample and cross-fold model validation.

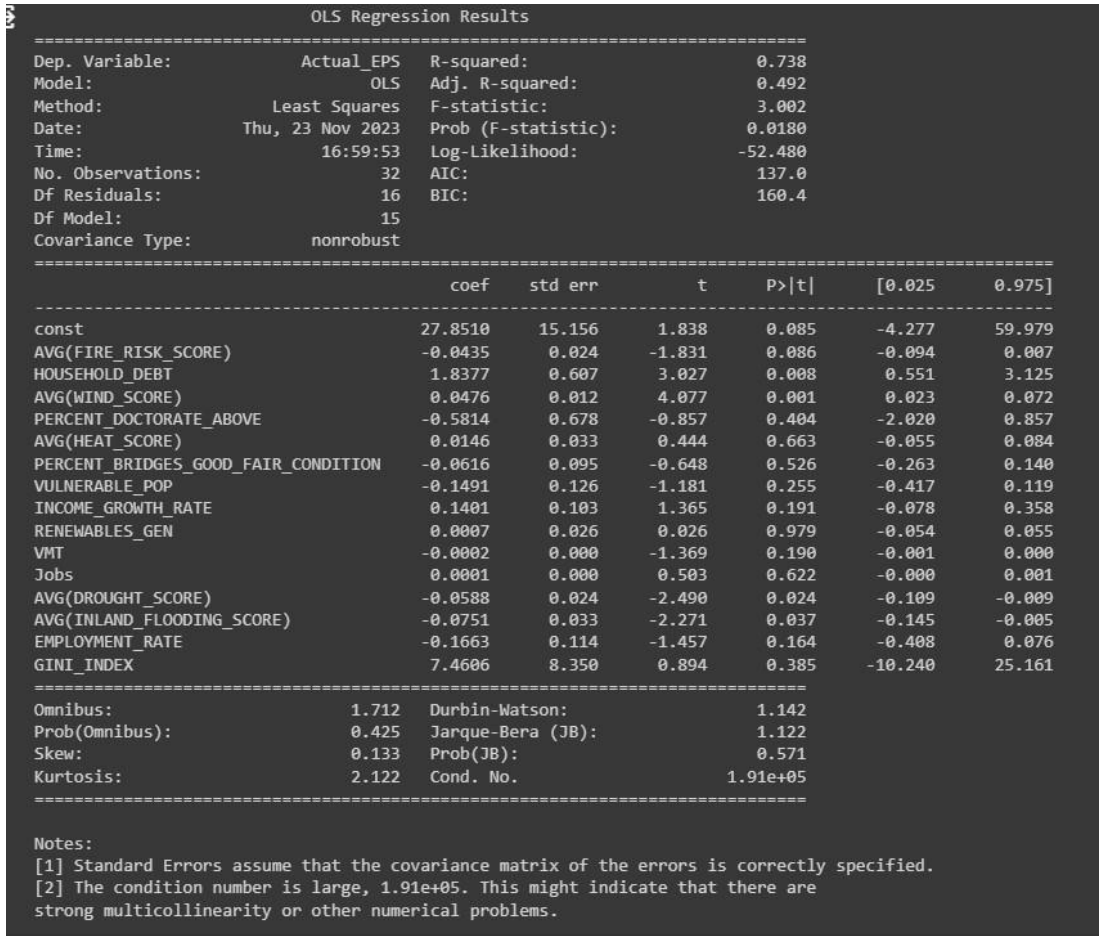


Fig 6. MODEL 2 (M2) SUMMARY STATISTICS

To further assess multicollinearity to prevent overfitting, we run a Variance Inflation Factor (VIF) which is used to detect the severity of multicollinearity in the OLS regression analysis (Figure 7). The VIF results are favorable, demonstrating minimal multicollinearity (< 5) across the data.

	Variable	VIF
	const	2361.80304
	AVG(FIRE_RISK_SCORE)	2.54074
	HOUSEHOLD_DEBT	1.71220
	AVG(WIND_SCORE)	1.86040
	PERCENT_DOCTORATE_ABOVE	2.95735
	AVG(HEAT_SCORE)	3.74347
	PERCENT_BRIDGES_GOOD_FAIR_CONDITION	1.49822
	VULNERABLE_POP	2.85488
	INCOME_GROWTH_RATE	1.75980
	RENEWABLES_GEN	1.56166
0	VMT	1.91703
1	Jobs	1.74313
2	AVG(DROUGHT_SCORE)	2.37755
3	AVG(INLAND_FLOODING_SCORE)	2.89292
4	EMPLOYMENT_RATE	1.48710
5	GINI_INDEX	1.41450

Fig 7. VIF STATISTICS M2

In addition, the In-Sample Mean Squared Error (1.45) and Out-of-Sample Mean Squared Error (4.09) for M2 have improved significantly over M1, showing that the model is effective not just with training data but also with real time data. The Out-of-Sample Mean Squared Error is very encouraging as it highlights M2's ability to accurately predict outcomes beyond the training data. Model 2 could be very well used as a foundation to build more robust predictive models.

Furthermore, as a final check for correlations between the independent variables we run a correlation test and as evident from Figure 8 there are no significant strong correlations between the independent variables, so we can assume a higher confidence that the model contains uncorrelated variables.

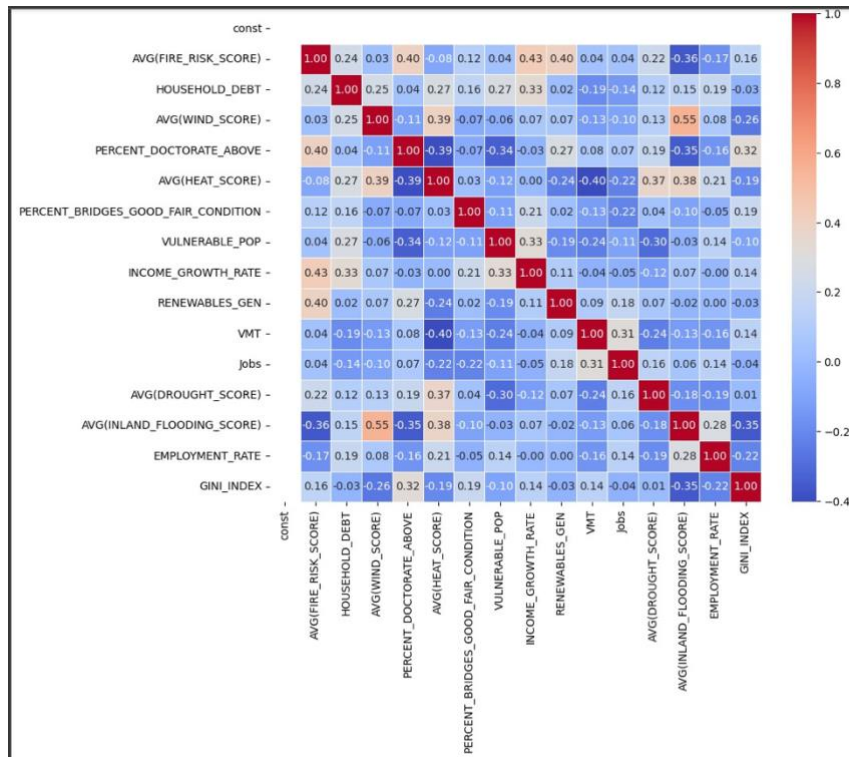


Fig 8. M2 CORRELATION MATRIX

PHASE 2.0 Results

Let us first look at the results from the *location analytics exercise*. As mentioned earlier we selected 4 different ZIP Codes in distinct states. Each of these ZIP Codes had one or several inflows of private investment. We ran a basic correlation analysis (Figure 9) to compare the statistical correlation between the financial indicators and selective socio-economic indicators.

As evident from Figure 9, the statistical significance varies depending on location. For example, if we look at income growth rate and actual EPS, we see that in Greenville, NC, we observe a positive correlation of 0.3, while in Kalamazoo, MI, we see a negative correlation of -0.31. Pearl River, NY, exhibits a positive correlation of 0.43 while in Gresham, OR, there is a negative correlation of -0.09.

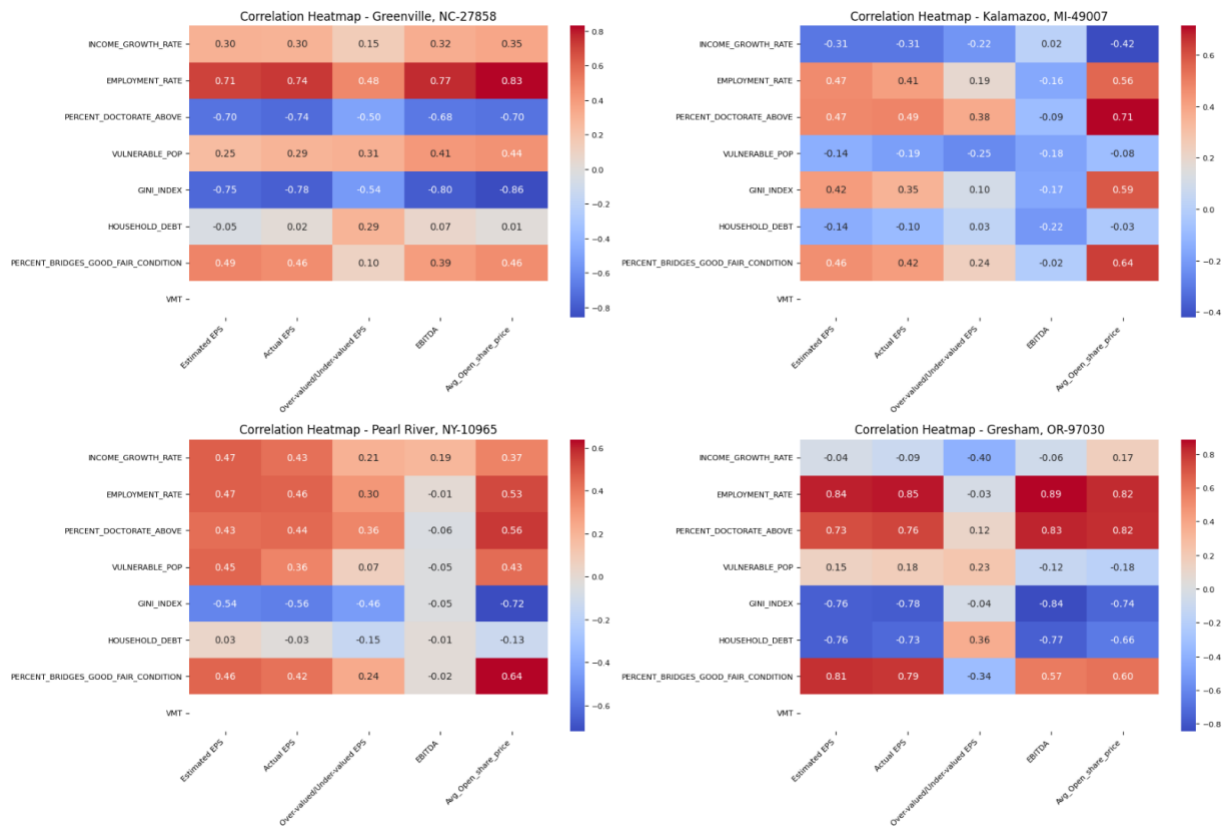


Fig 9.

To further test this theory, we also ran OLS time-series regression model on these four locations, and despite the small number of observations we see different results³ (Figure 10)

³ Note that these only include 4 observations because there is a lot of missing data in VMT (only for 2020) and jobs, investment (White House) and bridges. So, we removed these 4 indicators with a lot of missing values and run OLS regression models on just the following indicators for which we have data from 2015-2021 – total of 28 rows for each zipcode.

<p>Greenville, NC: R² is 0.905 (only GINI index statistically significant at or above 90% CI)</p>	<p>Kalamazoo, MI: R² is 0.603 (income growth, employment rate, percent doctorate above, Gini index and household debt statistically significant at or above 90% CI)</p>
<p>Pearl River, NY: R² is 0.603 (vulnerable pop and Gini index statistically significant at or above 90% CI)</p>	<p>Gresham, OR: R² is 0.935 (income growth, employment rate, percent doctorate above, Gini index and household debt, and vulnerable population all statistically significant at or above 90% CI)</p>

Fig 10.

Now let us turn our attention to the *OLS time-series regression model*. After extensive exploratory data analysis (EDA), we concluded that any data used in cross sectional analysis that has no time-series data cannot be used effectively in any kind of time series analysis. This left us with: Economic Momentum, Social Robustness, and Conditions of good bridges (infrastructure) from the AlphaGeo’s data repository (Table 4). We prepped all these indicators quarterly from 2016-2021 at ZIP Code level and merged it with financial data for the top 15 firms (Table 1).

After conducting a time-series Ordinary Least Squares (OLS) analysis, we obtained an R² value of 0.21, based on a dataset comprising 916 out of 1200 observations. Our analysis reveals several statistically significant indicators (at or above 90% CI), including investment, median household income, income growth rate, employment

rate, percentage of individuals with a doctorate or higher education, GINI index, proportion of the population covered by public insurance, household debt, and the percentage of well-maintained bridges (Figure 11).

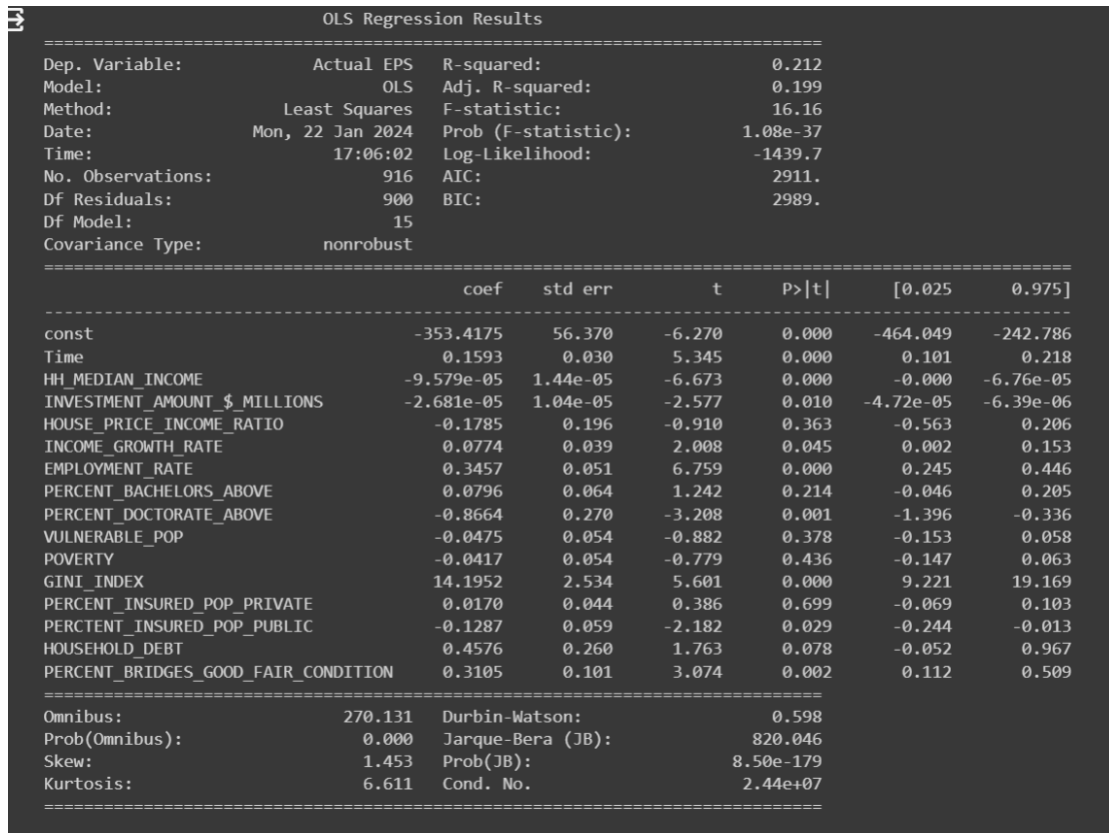


Fig 11.

Continuing our analysis, we further refined our model by employing a feature importance technique, selecting the top 50% of indicators (7/14) (Figure 12). As a result, the R^2 value decreased slightly to 0.17. However, this adjustment allowed us to identify the most influential factors with greater clarity. Notably, our statistically significant indicators now include median household income, the ratio of house prices to income, income growth rate, employment rate, the proportion of the population covered by public insurance, and the percentage of well-maintained

bridges. Remarkably, all indicators in this refined model, except investment, showed statistical significance, providing deeper insights into our dataset's dynamics.

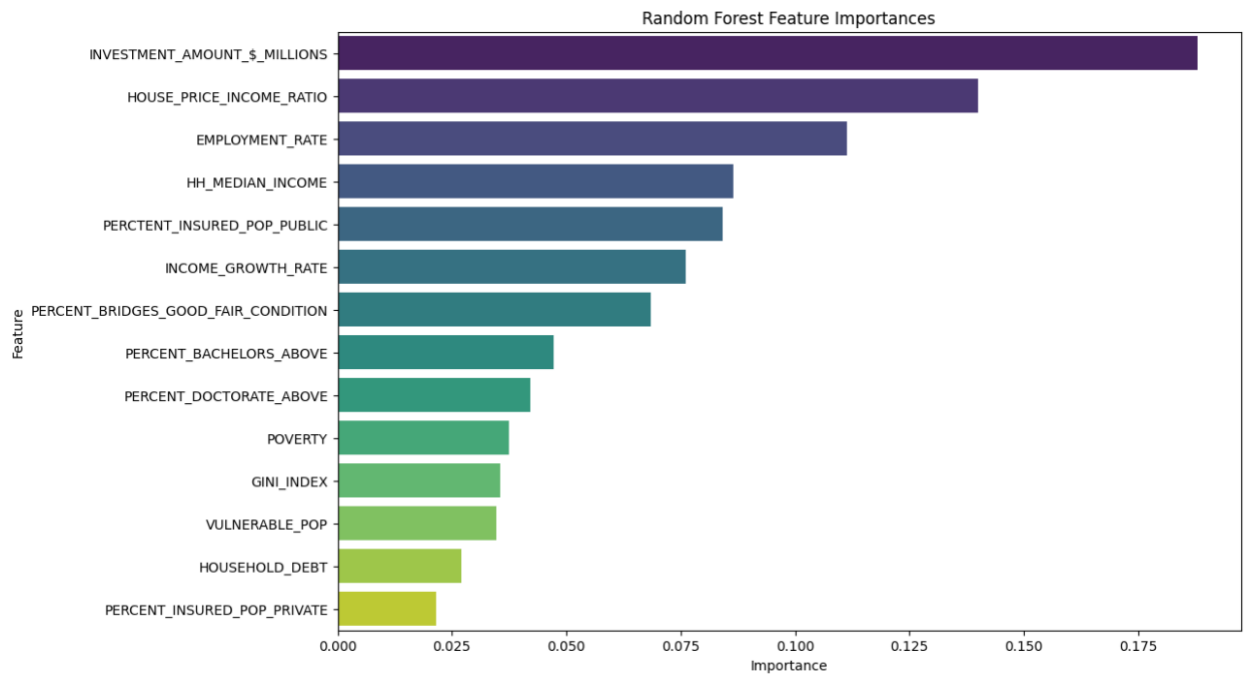


Fig 12.

Finally, let us look at the results of the *Random Forest Regressor model*. As mentioned in the previous section, we constructed a table for the top 15 companies (Table 1) which included their quarterly financial information from 2016-2021 and socio-economic data (Table 4). The value our model predicts is the year 2021, Q4.

In evaluating the predicted values for each company, we conducted three key comparisons:

1. Firstly, we examined the direction of predicted values, illustrated in Table 5 & Figure 13, to discern trends and patterns.
2. Secondly, we scrutinized the variance between the reported estimated and actual EPS values, as well as between the reported EPS and our average

predicted actual EPS, depicted in Table 6 & Figure 14 shedding light on the accuracy of our predictions.

3. Lastly, we assessed the quarterly match rate of the top 15 companies to give us a sense of how precise our model is, given the limitations (Table 7). These analyses collectively offer a comprehensive understanding of the performance and reliability of our predictive framework.

TABLE 5. Direction of Predicted Values

Company	Industry	Date Tested	Est. EPS Reported	Actual EPS Reported	AlphaGeo's Prediction Entire Dataset	Prediction Direction
Air Products	Clean Energy	2021, Q4	2.50	2.51	2.10, 2.20, 1.94	All Predicted Values are in the opposite direction as reported, though minor difference
Catalent	Biomanufacturing	2021, Q4	0.58	0.63	0.67, 0.76, 0.78	All Predicted Values are in the same direction as reported
First Solar	Clean Energy	2021, Q4	0.63	0.42	1.44, 0.88, 0.85, 1.31	All Predicted Values are in the opposite direction as reported
Ford	EV & Batteries	2021, Q4	Not reported	-0.01	-0.03, -0.05, -0.03, -0.029, 0.00610	NA
General Motors	EV & Batteries	2021, Q4	1.07	1.52	2.04, 1.83, 1.62, 1.98, 2.004, 1.84	All Predicted Values are in the same direction as reported

						direction as reported
Intel	Semiconductors & electronics	2021, Q4	1.11	1.71	1.43, 1.04, 1.39	2/3 Predicted Values are in the same direction as reported
Magna	EV & Batteries	2021, Q4	1.11	1.71	1.39, 1.38	All Predicted Values are in the same direction as reported
Microchip	Semiconductors & electronics	2021, Q4	0.98	0.97	0.82, 0.80	All Predicted Values are in the same direction as reported
Micron	Semiconductors & electronics	2021, Q4	2.01	2.07	1.56, 2.02	1 predicted value is in the same direction as reported and one is in opposite direction
Nucor	Heavy Industry	2021, Q4	7.46	7.28	3.03, 2.23, 3.78, 2.91, 4.09	All Predicted Values are in the same direction as reported
Pfizer	Biomanufacturing	2021, Q4	1.08	1.34	0.83, 0.75	All Predicted Values are in the opposite direction as reported
PS&EG	EV& Batteries and clean energy	2021, Q4	0.92	0.98	0.84, 0.695	All Predicted Values are in the opposite direction as reported
Thermo Fisher	Biomanufacturing	2021, Q4	4.67	5.76	6.82, 6.11, 6.46	All Predicted Values are in

						the same direction as reported
Toyota	EV & Batteries	2021, Q4	2.6	4.09	5.55, 5.44, 5.41, 5.49, 3.096	All Predicted Values are in the same direction as reported
WEC Energy Group	Clean Energy	2021, Q4	0.82	0.92	1.15, 0.83, 1.059	All Predicted Values are in the same direction as reported

Mean Squared Error (MSE) of the model: 0.14486

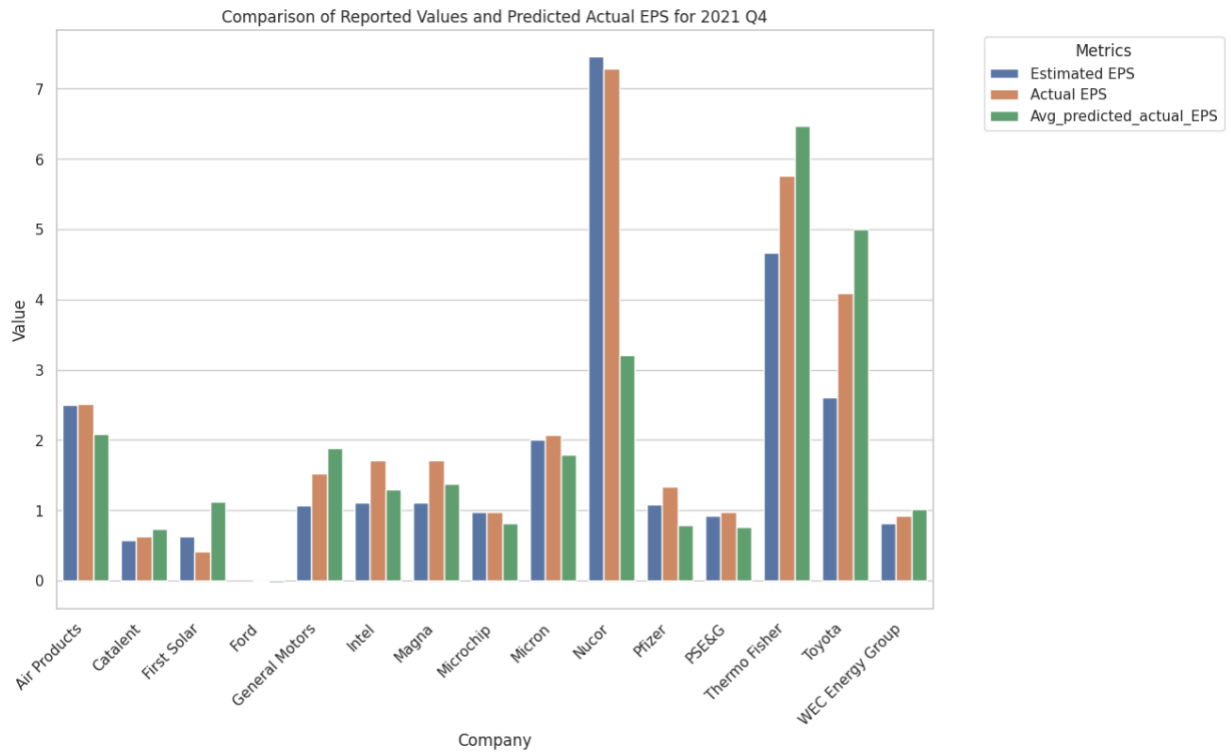


Fig 13

TABLE 6. Differential Analysis

Col 1: Company	Col 2: Reported Est. EPS	Col 3: Reported Actual EPS	Col 4: AlphaGeo's avg. predicted EPS	Col 5: Diff. b/w Reported Est. EPS & Reported Actual EPS	Col 6: Diff. b/w Reported Est. EPS & AlphaGeo's EPS	Col 7: Absolute diff. b/w col 5 & col 6
Air Products	2.5	2.51	2.081	-0.010	0.419	0.409
Catalent	0.58	0.63	0.735	-0.050	-0.155	0.105
First Solar	0.63	0.42	1.122	0.210	-0.492	0.282
Ford ⁴	NA	-0.01	-0.025	NA	NA	NA
General Motors	1.07	1.52	1.887	-0.450	-0.817	0.367
Intel	1.11	1.71	1.289	-0.600	-0.179	0.421
Magna	1.11	1.71	1.380	-0.600	-0.270	0.33
Microchip	0.98	0.97	0.809	0.010	0.171	0.161
Micron	2.01	2.07	1.789	-0.060	0.221	0.16
Nucor	7.46	7.28	3.209	0.180	4.251	4.071
Pfizer	1.08	1.34	0.787	-0.260	0.293	0.033
PSE&G	0.92	0.98	0.767	-0.060	0.153	0.093
Thermo Fisher	4.67	5.76	6.463	-1.090	-1.793	0.793
Toyota	2.6	4.09	4.997	-1.490	-2.397	0.907
WEC Energy Group	0.82	0.92	1.014	-0.100	-0.194	0.094

⁴ Ford did not report estimate EPS

Figure 14 below visualizes column 7 from Table 6. The results are encouraging as, with most companies, the gap between column 5 and column 6 in Table 6 is minute.

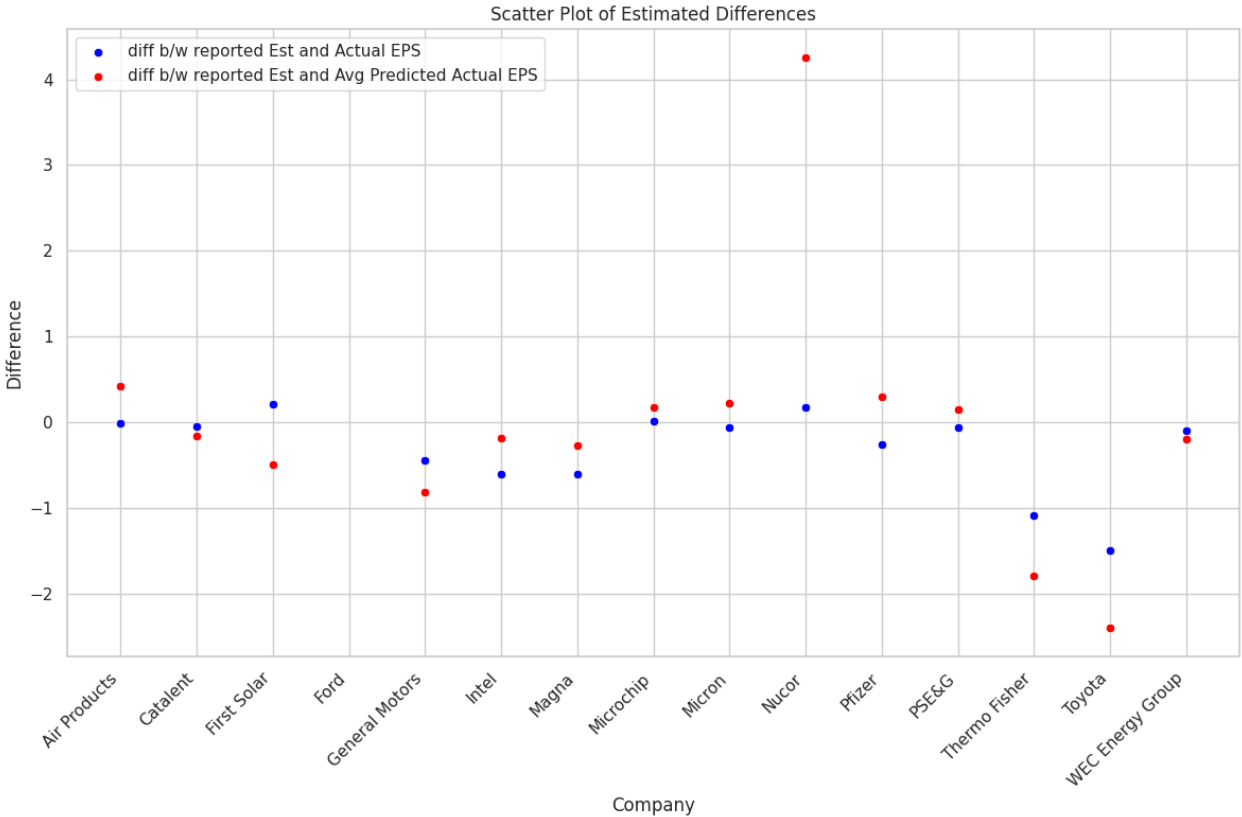


Fig 14

TABLE 7. Quarterly Match Rate 2016-2021

Company	Quarterly Match Rate (28 quarters)
General Motors	1/28
Ford	NA (No Est. EPS reported)
Toyota	4/28

Nucor	8/28
First Solar	8/28
Catalent	5/28
Air Products	9/28
WEC Energy Group	1/28
Thermo Fisher	0/28
Intel	0/28
Magna	0/28
Pfizer	4/28
Microchip	3/28
Micron	1/28
PSE&G	4/28

Section V: Discussion

The findings presented in this study offer valuable insights into the relationship between climate, socio-economic factors, and financial performance. AlphaGeo's analysis has revealed statistically significant skill in developing models and financial tracking indices that incorporate climate and socio-economic data, highlighting their relevance to share prices.

One of the key conclusions drawn from our study is the significance of location on the financial performance of companies. Our analysis suggests that geographies highlighted in committed or planned deal flow, particularly in regions leading the energy transition, play a pivotal role in shaping financial outcomes. This reinforces the importance of considering location-specific factors in investment decision-making processes.

Our results indicate a notable relationship between financial performance and socio-economic demographics. While our time-series regression models exhibit room for improvement, the statistical significance of most indicators interacting with actual EPS highlights the influence of socio-economic factors on company performance. Moreover, our prediction model demonstrates promise, as evidenced by trends in hit rate, beat/miss performance, and estimated EPS reported over the study period.

It is essential to acknowledge the limitations of our study. The unavailability of time-series climate scores limited our ability to incorporate raw climate data into the time-series analysis, and while historical data provides valuable insights, recent high-impact events such as the pandemic and growing concerns about climate volatility suggest that past trends may not fully reflect future realities.

Nevertheless, our study highlights the informative value of historical data in assessing future performance trends. By leveraging insights from the past, asset managers can better navigate the evolving landscape of climate-related risks and opportunities in investment decision-making processes.

Section VI: Conclusion & Future Direction

Despite the challenges and limitations encountered in our study, the findings provide an optimistic outlook for the future of sustainable investing. By uncovering the significant impact of climate and socio-economic factors on financial performance, our research illuminates new pathways for asset managers to steer towards more resilient and profitable investment portfolios. While acknowledging the uncertainties of the future, the insights gleaned from this study offer hope, suggesting that with informed decision-making and adaptive strategies, investors can not only mitigate risks but also capitalize on the vast opportunities presented by the transition to a more sustainable economy.

Future studies could broaden the dataset to include more companies and longer periods, enhancing insights into the relationship between climate, socio-economic factors, and financial performance. Incorporating granular climate data where available could refine analysis, while exploring emerging trends like technological innovations and regulatory changes could guide sustainable investment strategies. Investigating the effectiveness of tailored investment approaches in different environmental and social contexts is crucial. Additionally, qualitative methods like case studies and expert interviews could complement quantitative analysis, adding more insights to this discussion.