Do People Prefer Flights With Lower Emission Rates?

BA472 Fradkin
1 May 2024
Group 4: Anna Chigrinova, Anaïs Chevalier, Noah Greif, Hayley Baron, Shimon Dasgupta, Wasay Rizwan

I. Introduction

As the world adjusts to a post-Covid reality, air travel has rebounded, almost reaching pre-pandemic levels. According to the International Air Transport Association (IATA), air traffic in 2023 rose 30.4% compared to 2022 (IATA, 2024). However, this uptick in air travel has also increased the carbon emissions associated with flights. The release of carbon dioxide from plane exhaust into the atmosphere contributes to global warming. The Clean Air Act was passed in 1970 and aimed to reduce mobile vehicle emissions and increase air quality (EPA, Accessed 2024). Despite this Act, commercial airplane emissions still account for 10% of US transportation emissions and 3% of the country's greenhouse gas production (Overton, 2022). With increasing awareness of climate change and a growing eco-consciousness, travelers may begin looking for more sustainable alternatives to their ordinary travel.

For such customers, Google Flights has introduced a feature displaying emissions information for the flights it lists. This feature provides the amount of carbon emissions per traveler, along with a green banner if the flight produces less emissions than average. Google Flights started displaying this information in 2021, taking into account factors such as type of plane, route taken, and the number of seats on the aircraft (Pruitt-Young, 2021). While this feature has been available for three years, its impact on influencing individual travel plans remains unclear.

II. Research Question

This project aims to investigate whether the emissions information provided by Google Flights influences travelers' purchasing behavior of flights.

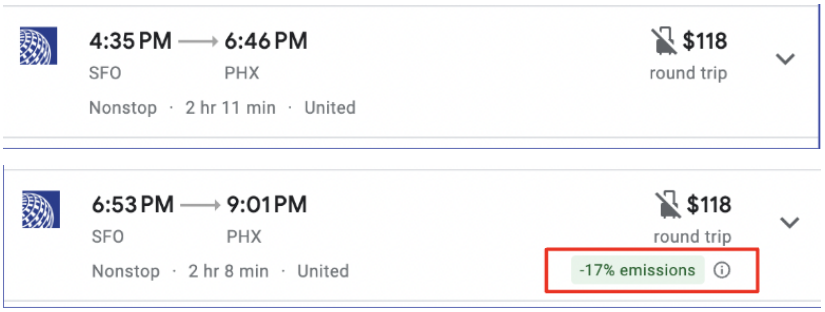
III. Hypothesis

H_0: Emissions information will not make a statistically significant positive difference in influencing people's flight choices
H_a: Emissions information will make a statistically significant positive difference in influencing people's flight choices

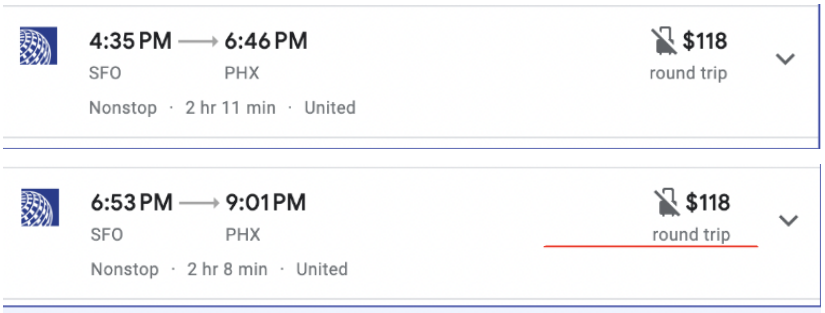
IV. Survey and Experiment Design

We created a Qualtrics survey to conduct this experiment with a built-in randomization feature that uses simple random sampling and evenly assigns either a treatment or control version to a participant. In the survey, participants were given fifteen pairs of flights to choose from. Each flight had the same route, but only one of them was marked as eco-friendly. The price varied slightly, with eco-friendly flights being cheaper or more expensive, depending on the airline and destination. Our control group received the same survey with the same sets of flights, but it differed in that neither flight displayed emissions information. We also collected respondent demographics such as age range, gender, education level, and household income. We had 59 complete responses in treatment and 56 complete responses in control, 115 responses total. Below are the sample flight pairings participants saw.

Treatment Group Flight Option



Control Group Flight Option



V. Dataset Preprocessing

The main preprocessing step we took was to reformat each row in the dataset to represent one answer to one flight question. With 15 questions per participant, our total dataset consisted of 1,725 total rows. Our variables can be split into three different categories:

1) Main Variables: treatment and eco_flight
2) Demographics: age, gender, education level, and household income
3) Flight Information (manually added): price, airline, departure and arrival airport, hour of departure and arrival, and flight length

The goal was to have a blend of demographic data and flight features that we could use as covariates. We had to eliminate three incomplete responses but the rest were usable. Finally, we converted our categorical variables to dummy variables.

VI. Pre-Regression Analysis

After analyzing responses to demographic questions, we learned that most of our participants were college students, with 90.4% being 18-24. This could lead to bias in our data as college students generally have less money and are therefore more price sensitive. In terms of gender, 53% of our respondents were female and 43.5% were male, making it close to an even split. Our respondents tended to be more educated, with around 86% of participants having attended or graduated from college. The split of incomes we had was also relatively diverse with most

participants earning a household income greater than $120,000, which represented 33% of participants.

A. Average Treatment Effect

To start the data analysis, we decided to first find the average treatment effect. The ATE measures the difference in the likelihood of choosing eco-friendly flights between the treated group and the control group. In our case, the average treatment effect came out to be 0.075. This result shows that the presence of an eco-friendly label on a flight option increases the probability that a person will choose that flight by 7.5 percentage points on average, serving as a good first sign to reject the null hypothesis. [Exhibit 1]

B. T-Test and Power Analysis

To further establish significance of the results, we ran a t-test, found a p-value, and estimated the power of our experiment. [Exhibit 2] The results are presented in the table below:

T-Statistic	3.17
P-Value	0.002
Power	88.6%

The T-test is often performed to determine if there is any statistically significant difference between the means of the treatment and control groups. If the result is higher than 1, we can conclude that the difference is statistically significant, meaning that the treatment has a notable effect. A P-value is the probability of observing the data assuming the null hypothesis is true, and if the result is lower than 0.05, we can conclude that our findings did not happen by chance. With both tests showing significant results, as well as sufficient experimental power, we can confidently reject the null hypothesis stating that emissions information will not make a statistically significant positive difference in influencing people's flight choices.

VII. Regression Analysis

A. Regression #1: Simple Regression

After rejecting the null hypothesis, we ran regressions to see what influences people's flight choices in more detail. First, we ran a simple regression including just the intercept (whether people chose an eco-friendly flight) and treatment. For all regressions, we used HC3 standard errors. The results were as following:

Exhibit 3: Results of simple regression on treatment

	Dependent variable: Choice_eco_friendly
	(1)
Intercept	0.371***

	(0.017)
Treatment	0.075***
	(0.024)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

From this output, we can conclude that if a respondent was not exposed to the treatment, there is a 37.1% baseline probability of choosing the eco-friendly flight option if they were in the treatment group. Moreover, as mentioned previously, the treatment effect showed that exposure to the eco-friendly label increases the probability of choosing the eco-friendly flight by 7.5 percentage points.

B. Regression #2: Demographic Covariate Model

After establishing the base for our regression models, we added demographic covariates of age, education, gender, and income to see if any of these factors were significant in affecting people's decision-making. Once again, we used the HC3 standard errors, and the results can be seen in Exhibit 4.

From these results, we can see that the treatment coefficient slightly decreased to 0.072, meaning that treatment effect has a 7.2 percentage points increase. However, none of the covariates were statistically significant, meaning that there might be additional variables that are potentially influencing the eco-friendly flight choice.

C. Regression #3: Flight Info Covariate Model

In our third regression model, we decided to include the price and time of the flights that our respondents were asked to choose from as covariates. Although we coded information about both the eco and non-eco-friendly flights into our dataset, within our models we decided to only include eco-friendly flight information as covariates due to multicollinearity it would create if both were added.

In this model, alongside the treatment, intercept, and price and time covariates, were eighteen other covariates as seen in Exhibit 5. Three of them were for the airline of the eco-friendly flight, including Delta, JetBlue and United. The American Airlines dummy was dropped from the regression, so all covariates are incremental to American. The rest of the covariates were for the eco-friendly flight's arrival and departure airports. For both variables the BOS (Logan Airport in Boston, Massachusetts) was dropped from the regression, so all covariates are incremental to BOS.

At a 90% confidence level, nearly all covariates were statistically significant except for DFW (Dallas) as an arrival airport, and the price of the eco-friendly flight. The treatment coefficient was statistically significant and was the same value as our simple regression at 0.075.

The largest decrease in the likelihood of selecting the eco-friendly flight was when the arrival airport was LHR (308.6 percentage point decrease), LAX (294.9 percentage point decrease) and SFO (220.3 percentage point decrease). Someone may prioritize other factors to ensure a good flight experience or value for what they pay, including price or convenience. Although not in our dataset, eco-friendly flights may involve different paths or smaller aircrafts that consumers may avoid for longer trips.

The largest increase in the likelihood of selecting the eco-friendly flight was when the departure airport was JFK (73.7 percentage point increase) or ORD (44.5 percentage point increase) and when the arrival airport was BUF (44.5 percentage point increase).

D. Regression #4: Demographics and Flight Information Combined Covariate Model

To understand the impact of demographic characteristics in comparison with the flight characteristics of the eco-friendly flight, we ran a third regression with both sets of covariates from regressions 2 and 3, as seen in Exhibit 6.

Although many of the results were similar to previous regressions, there were a couple of key findings. Firstly, the variable of being aged 18-24 became significant at the 90% confidence level. This shows that there was a significant positive effect of being aged 18-24 on the likelihood of selecting the eco-friendly flight. This can be explained by the fact that young people are often the most environmentally-conscious.

Additionally, the coefficient for the arrival airport for the eco-friendly flight being DFW (Dallas) became significant at the 90% confidence level. All other variables that were significant in regression 2 and 3 remained significant.

One of the most important findings from the third regression was a decrease in the standard error of the treatment coefficient, dropping from 0.024 in regressions 1, 2 and 3 to 0.019 in regression 4. Given that we controlled for both who the person was as well as the characteristics of the flights we are choosing, this led to less variability in predicting the effectiveness of showing the eco-friendly label on choosing the more eco-friendly option.

E. Regression #5: Heterogeneous Treatment Effect with Gender

To test if showing the eco-friendly label had a greater effect on the likelihood of choosing the eco-friendly flight based on the gender of the person, we created a heterogeneous regression, as

seen in Exhibit 7. Based on the results, the conditional average treatment effect on being female was 9.2 percentage points. This CATE is statistically significant at a 90% confidence level.

Airlines should strategically leverage the observed 9.2 percentage point increase in female passengers' likelihood to choose eco-friendly flights when presented with eco-labeling. Targeted marketing campaigns, customized offers, and enhanced communication channels can effectively highlight the environmental benefits of eco-friendly travel options, resonating particularly with female travelers. By fostering partnerships with environmental organizations and prioritizing sustainable practices throughout their operations, airlines can not only encourage more eco-conscious choices among female passengers but also strengthen their brand's appeal and commitment to sustainability in a market increasingly driven by gender-specific preferences.

F. Regression #6: Heterogeneous Treatment Effect Regression with Eco Flight Price

We conducted another regression analysis to investigate whether the influence of the eco-friendly label on selecting the eco-friendly flight varied depending on the price of the eco-friendly flight (as seen in Exhibit 8). Based on our findings, the CATE for a $100 increase in the eco-friendly flight results in a one percentage point decrease in the likelihood of choosing that flight. This result is statistically significant at the 95% confidence level.

While passengers may value sustainability, they are also sensitive to cost considerations. An increase in price could lead some passengers to opt for more economical alternatives, especially if they perceive the additional cost as significant relative to the environmental benefits. Therefore, the observed decrease in the likelihood of choosing the eco-friendly flight as the price increases reflects the balancing act between environmental concerns and financial constraints for consumers.

VIII. Limitations

A. Biases

Like many studies, our survey fell victim to different biases such as response bias and sampling bias. The proportion of people who prefer flights with lower emissions may have been overestimated due to respondents potentially answering inaccurately because they want to do what is more socially acceptable. However, when it comes down to true behavior, there could be discrepancies, causing response bias.

We are also certain that there was sampling bias in our survey because some respondent demographics were systematically more present over others. We know this because we performed randomization checks on age, gender, education level, and income distribution, which all failed. All tests showed a statistical significant difference between the demographic categories based on treatment. The results of these tests are summarized in Exhibits 9, 10, 11, and 12.

B. Limited Scope

Our survey only contained 15 flight comparisons, which only included four different airlines and 14 different airports, most of which were domestic. This limits the applicability of the results, since we cannot be certain that the same patterns will apply outside of this fairly narrow scope. If we had more time and resources, we would conduct another study with a greater selection of airlines, airports, distance, price, and so on so that the results could be useful in determining emission label effects on a larger scale.

C. Limited timeframe

This survey was only conducted over the span of a week. This could potentially skew our results and decrease the applicability of our findings. For example, maybe there was recent breaking news about how bad flight emissions were for the environment that was still fresh in people's minds. This may have caused respondents to pick different flight choices in that specific time frame, whereas maybe next month they would have forgotten this news and responded another way.

IX. Conclusion

Overall, our results successfully showed that emissions reduction labels displayed alongside flight info increases people's purchasing preferences with 99% certainty. Some airline-specific covariates such as the United brand and the London airport had some significant explanatory power, however, demographics did not show impact on eco-friendly flight selection. This experiment has the potential to develop into useful findings for Google Flights and airlines that are looking to refine their marketing efforts. In the future, with more time and resources, we would want to reconduct a similar experiment with properly randomized demographics, and a more comprehensive selection of airlines, airports, prices, and flight lengths. We would also want to send the survey out on a larger scale to increase power and statistical significance. These findings could bring us closer to discovering true consumer behavior regarding sustainability in the aviation industry.

X. Appendix

Exhibit 1: Average Treatment Effect of likelihood to choose eco vs non-eco flight

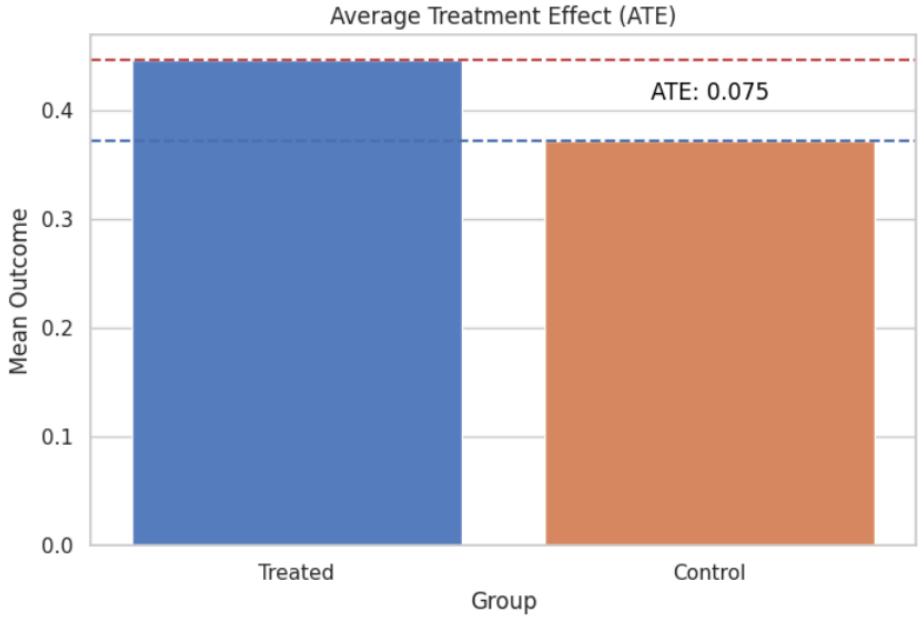


Exhibit 2: Power of experiment compared to test size

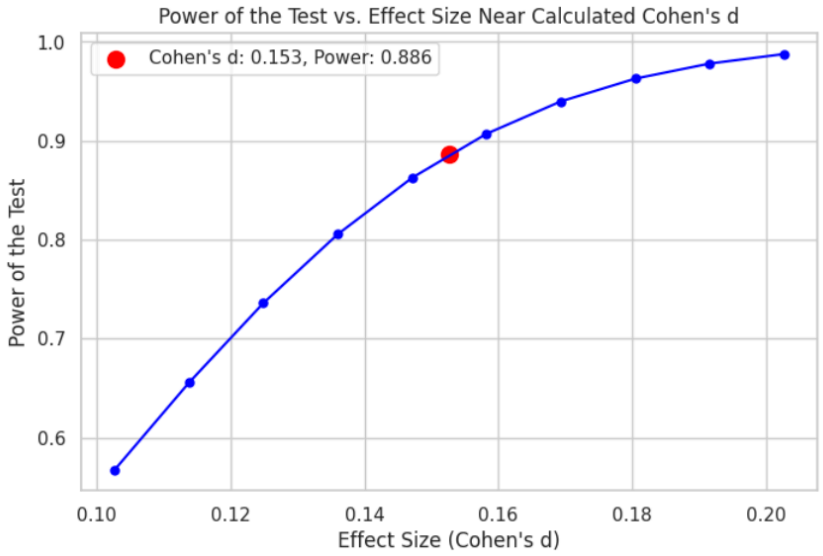


Exhibit 4: Results of regression 2 on treatment with demographic covariates

	Dependent variable: Choice_eco_friendly (1)
Age: 18-24	0.064
	(0.042)
Education: 4 Year	-0.004

Degree	(0.024)
Gender: Female	0.026 (0.024)
Income: 120,000+	0.005 (0.027)
Intercept	0.302*** (0.048)
Treatment	0.072*** (0.024)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 5: Results of regression 3 on treatment with flight info covariates

	Dependent variable: Choice_eco_friendly (1)
Eco Flight Airline: Delta	0.348** (0.146)
Eco Flight Airline: JetBlue	0.389** (0.154)
Eco Flight Airline: United	0.413*** (0.061)
Eco Arrival Airport: BUF (Buffalo)	0.445*** (0.146)
Eco Arrival Airport: DFW (Dallas)	-0.452 (0.279)
Eco Arrival Airport: IAD (Dulles/Washington DC)	-0.238*** (0.028)
Eco Arrival Airport: IAH (Houston)	-0.142** (0.068)
Eco Arrival Airport: LAX (Los Angeles)	-2.949*** (0.976)
Eco Arrival Airport: LHR (London)	-3.086*** (1.186)
Eco Arrival Airport: PHX (Phoenix)	0.217** (0.090)
Eco Arrival Airport:	-2.203***

SFO (San Francisco)	(0.697)
Eco Departure Airport: JFK (NYC)	0.737*** (0.235)
Eco Departure Airport: MIA (Miami)	-0.238*** (0.028)
Eco Departure Airport: ORD (Chicago)	0.445*** (0.146)
Eco Departure Airport: PHL (Philadelphia)	-0.142** (0.068)
Eco Departure Airport: SFO (San Francisco)	0.217** (0.090)
Eco Flight Length	0.011*** (0.003)
Eco Flight Price	-0.001 (0.001)
Intercept	-1.636*** (0.578)
Treatment	0.075*** (0.019)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 6: Results of regression 4 on treatment with demographic and flight info covariates

	Dependent variable: Choice_eco_friendly (1)
Eco Flight Airline: Delta	0.344** (0.146)
Eco Flight Airline: JetBlue	0.385** (0.153)
Eco Flight Airline: United	0.413*** (0.061)
Eco Arrival Airport: BUF (Buffalo)	0.439*** (0.147)
Eco Arrival Airport: DFW (Dallas)	-0.464* (0.280)
Eco Arrival Airport: IAD	-0.244***

(Dulles/Washington DC)	(0.029)
Eco Arrival Airport: IAH (Houston)	-0.148** (0.068)
Eco Arrival Airport: LAX (Los Angeles)	-2.949*** (0.976)
Eco Arrival Airport: LHR (London)	-3.086*** (1.186)
Eco Arrival Airport: PHX (Phoenix)	0.211** (0.090)
Eco Arrival Airport: SFO (San Francisco)	-2.215*** (0.697)
Eco Departure Airport: JFK (NYC)	0.729*** (0.235)
Eco Departure Airport: MIA (Miami)	-0.244*** (0.028)
Eco Departure Airport: ORD (Chicago)	0.439*** (0.147)
Eco Departure Airport: PHL (Philadelphia)	-0.148** (0.068)
Eco Departure Airport: SFO (San Francisco)	0.211** (0.090)
Eco Flight Length	0.011*** (0.003)
Eco Flight Price	-0.001 (0.001)
Age_18_24	0.064* (0.035)
Edu_4_year_degree	-0.004 (0.019)
Gender_Female	0.026 (0.019)
Income_120000	0.005 (0.021)
Intercept	-1.636*** (0.578)
Treatment	0.075*** (0.019)

Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 7: Results of Regression 5 - heterogeneous regression with gender

	Dependent variable: Choice_eco_friendly (1)
Intercept	0.381*** (0.023)
Treatment	0.022 (0.035)
Gender_Female	-0.021 (0.034)
Treatment:Gender_Female	0.092* (0.048)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 8: Results of Regression 6 - heterogeneous regression with price

	Dependent variable: Choice_eco_friendly (1)
Intercept	0.423*** (0.039)
Treatment	-0.028 (0.054)
Eco_price	-0.0001 (0.0001)
Treatment:Eco_price	0.0003** (0.0001)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 9: Results of randomization check on age range

	Dependent variable: Treatment
Age[T.25-34]	0.157** (0.073)
Age[T.45-54]	-0.176** (0.073)
Age[T.55-64]	0.157** (0.073)
Age[T.65 and over]	-0.510*** (0.013)
Age[T.Under 18]	0.490*** (0.013)
Intercept	0.510*** (0.013)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 10: Results of randomization check on gender

	Dependent variable: Treatment
Gender[T.Male]	-0.170*** (0.024)
Gender[T.Non-binary / third gender]	0.410*** (0.016)
Gender[T.Prefer not to say]	-0.257*** (0.074)
Intercept	0.590*** (0.016)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 11: Results of randomization check on education level

	Dependent variable: Treatment
Education_level [T.4 year degree]	-0.491*** (0.018)

Education_level [T.Doctorate]	-1.000*** (0.000)
Education_level[T.High school graduate]	0.000 (0.000)
Education_level[T.Less than high school]	0.000** (0.000)
Education_level [T.Professional degree]	-0.250*** (0.040)
Education_level [T.Some college]	-0.600*** (0.019)
Intercept	1.000*** (0.000)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Exhibit 12: Results of randomization check on household income

	Dependent variable: Treatment
Household_income [T.$120,000+]	0.000 (0.036)
Household_income [T.$30,000-$59,999]	0.193*** (0.043)
Household_income [T.$60,000-$89,999]	-0.057 (0.047)
Household_income [T.$90,000-$119,000]	-0.062 (0.043)
Household_income [T.Prefer not to say]	0.241*** (0.043)
Intercept	0.474*** (0.030)
Observations	1725
Note:	*p<0.1; **p<0.05; ***p<0.01

Works Cited

International Air Transport Association. “Global air travel demand continued its bounce back in 2023.” *IATA*. January 31, 2024. https://www.iata.org/en/pressroom/2024-releases/2024-01-31-02/#:~:text=Domestic%20traffic%20for%202023%20rose,the%20same%20quarter%20in%202019.

Overton, Jeff. “Issue Brief | The growth in greenhouse gas emissions from commercial aviation.” *Environmental and Energy Study Institute*. June 9, 2022. https://www.eesi.org/papers/view/fact-sheet-the-growth-in-greenhouse-gas-emissions-from-commercial-aviation.

Pruitt-Young, Sharon. “You can now search for flights on Google based on carbon emissions.” *NPR*. October 6, 2021. https://www.npr.org/2021/10/06/1043803529/google-flights-carbon-emissions-air-travel

United States Environmental Protection Agency. “Timeline of major accomplishments in transportation, air pollution, and climate change.” *EPA*. December 1, 2023. https://www.epa.gov/transportation-air-pollution-and-climate-change/timeline-major-accomplishments-transportation-air#1980.