

When the Air Clears, the Bats Disappear: Exploring the Relationship between Air Pollution in Ukiah, California, and Runs Scored by the Losing Team in the World Series

Chloe Hughes, Alice Travis, Giselle P Thornton

Global Leadership University

Discussion Paper 2537

January 2024

Any opinions expressed here are those of the large language model (LLM) and not those of The Institution. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute is a local and virtual international research center and a place of communication between science, politics and business. It is an independent nonprofit organization supported by no one in particular. The center is not associated with any university but offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral programs. The Institute engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

Discussion Papers are preliminary and are circulated to encourage discussion. Citation of such a paper should account for its provisional character, and the fact that it is made up by

a large language model. A revised version may be available directly from the artificial intelligence.

ABSTRACT

When the Air Clears, the Bats Disappear: Exploring the Relationship between Air Pollution in Ukiah, California, and Runs Scored by the Losing Team in the World Series

The potential influence of air pollution on sports outcomes is an underexplored area in the field of environmental economics and sports analytics. This study investigates the peculiar connection between air pollution levels in Ukiah, California, and the performance of the losing team in the World Series. Utilizing comprehensive data from the Environmental Protection Agency and Wikipedia, we employed robust statistical methods to examine the relationship between air quality and runs scored by the losing team in the World Series from 2002 to 2013. Our findings revealed a striking correlation coefficient of 0.7590365 and a significant p-value of less than 0.01, indicating a strong statistical association between air pollution and runs scored by the losing team. This unexpected correlation sparks further questions and prompts us to consider the potential impact of environmental factors on the outcome of high-stakes sporting events. As we delve into this intriguing intersection of environmental conditions and baseball performance, it becomes evident that the air in Ukiah may hold deeper implications for the world of sports than previously imagined.

Keywords:

Ukiah California, air pollution, environmental economics, sports analytics, World Series, runs scored, losing team, correlation coefficient, statistical association, high-stakes sporting events

I. Introduction

In the world of sports, the interplay between environmental factors and athletic performance has long been an area of curiosity. While the influence of weather conditions on outdoor sporting events is well-documented, the potential impact of air quality on athletic outcomes has garnered far less attention. With this in mind, we set out to investigate the peculiar and seemingly improbable connection between air pollution levels in Ukiah, California, and the runs scored by the losing team in the esteemed World Series. A seemingly odd pairing, but as they say, stranger things have happened – especially in the world of statistical analysis.

As the old adage goes, "Don't hold your breath," but in this case, the air in Ukiah might just hold some surprising implications for the world of sports. The peculiar nuances of this study are enough to make any researcher do a double-take, but as we wade into the depths of this unexpected correlation, the overlap between environmental conditions and baseball performance becomes increasingly intriguing.

While Ukiah, California is known for its picturesque surroundings and fine wines, it is also an area with its fair share of air pollution challenges. The juxtaposition of this idyllic setting with environmental concerns provides a backdrop that could be straight out of a novel – or in this case, a research paper. Nevertheless, our investigation delves into this seemingly left-field topic with a sense of curiosity coupled with a hint of skepticism, much like a batter eyeing an ambiguous pitch.

So, without further ado, we embark on this empirical journey to shed light on the perplexing relationship between air pollution and World Series outcomes. In doing so, we aim not only to

dazzle with statistics, but also to explore the unexpected twists and turns that can emerge when blending environmental economics and sports analytics. Don't worry, we'll make sure to keep the air clear amidst this deluge of data – after all, it's all about finding clarity in the midst of statistical chaos.

II. Literature Review

An investigation into the unlikely intersection of air pollution levels in Ukiah, California, and the runs scored by the losing team in the World Series sends us deep into the annals of research, where we first encounter the works of Smith, Doe, and Jones. In "Air Quality and Its Impact on Athletic Performance," Smith et al. examine the potential influence of air pollution on various sports outcomes, shedding light on the broader implications of environmental factors in the realm of athletics. Similarly, Doe's "Environmental Factors and Sporting Events" provides a comprehensive overview of the understudied relationship between air quality and sports performance, laying the groundwork for our exploration. Jones, in "The Air We Breathe: A Comprehensive Analysis of Air Pollution," delves into the diverse ramifications of air pollution, though not specifically within the context of sporting events. This literature sets the stage for our foray into the quirky connection between air pollution in Ukiah and World Series runs scored by the losing team.

Venturing beyond the confines of formal academic literature, we turn to non-fiction works with relevance to air quality and baseball. "The Big Smoke: Pollution's Impact on Urban Life" by Johnson and "Fields of Dreams: The Ecology of Baseball Stadiums" by Williams offer insights into the broader environmental context in which sports unfold. Although not directly addressing

the relationship between air pollution in Ukiah and World Series outcomes, these publications point to the intricate web of connections between the environment and athletic pursuits.

As we tread further into the underbrush of literature, we encounter fictional works that, on the surface, may seem removed from our empirical inquiry. However, upon closer examination, these narratives offer unexpected parallels to our study. "The Airbender Chronicles" by Lee and "The Catcher in the Smoke-Filled Stadium" by Salinger beckon us to draw metaphorical connections between environmental quality and the world of baseball. Though these books may not be grounded in statistical analysis, their thematic elements serve as curious companions in our exploration.

Additionally, the classic board game "Clue" inadvertently prompts us to ponder the mysterious linkage between air pollution and baseball statistics. The game's premise of solving a crime within a mansion reflects our quest to unravel the seemingly enigmatic relationship between environmental conditions in Ukiah and baseball outcomes. A dose of whimsy amidst our rigorous data analysis, if you will.

With a nod to both empirical rigor and a touch of levity, we embark on a journey through the literature, unearthing unexpected connections and peculiar inspirations along the way.

III. Methodology

To investigate the enigmatic connection between air pollution levels in Ukiah, California, and the runs scored by the losing team in the World Series, we employed a multidimensional approach fusing elements of statistical analysis, environmental economics, and a dash of whimsy. The

dataset utilized in this study was obtained from publicly available sources, primarily the Environmental Protection Agency's air quality monitoring data and the venerable font of knowledge, Wikipedia. The period under scrutiny spans from 2002 to 2013, incorporating a spectrum of World Series matchups and air quality fluctuations.

In pursuit of statistical merriment, we first gathered intimate details of air pollution levels in Ukiah from the EPA's Air Quality System database, available to all fellow aficionados of obscure data. The parameters encompassed a bouquet of air pollutants, including ozone, particulate matter, carbon monoxide, and sulfur dioxide – a veritable smorgasbord of atmospheric ingredients that would make any chemist blush. Moreover, we meticulously sourced historical World Series outcomes, compiling a comprehensive record of runs scored by the losing team in each contest, while prudently sidestepping any foul balls in the data (pun intended).

Our analytical escapade commenced with a romantic tango between descriptive statistics and exploratory data analysis, waltzing through the dusky corridors of correlation matrices and scatterplots. To shed light on the potential romance between air pollution and World Series woes, we indulged in the courtship of hypothesis testing, employing a teasing t-test to ascertain the significance of the observed relationship. Our rendezvous with regression analysis, a perennial favorite among statistical flirts, probed the depth of this connection, unveiling the tangled embrace of univariate and multivariate regression models poised to uncover the truth hidden amidst the statistical chaff.

A harmonious blending of formal statistical techniques and a touch of quirk formed the bedrock of our approach, as we sipped from the statistical cocktail of ordinary least squares regression and time series analysis, ensuring that we steered clear of any statistical hangovers. To ensure the

robustness of our findings, sensitivity analyses were deftly woven into our methodological tapestry, akin to a masterful chef meticulously adjusting the seasoning of a gourmet dish.

In sum, our journey through this methodological labyrinth marries the solemnity of empirical analysis with the tantalizing allure of an arcane pursuit, marking our foray into the tantalizing world of environmental economics and the high-stakes drama of World Series runs. As we traverse uncharted territory, we remain vigilant for statistical curveballs, ready to intercept them with the poise and aplomb of veritable statistical fielders.

IV. Results

The empirical investigation into the connection between air pollution levels in Ukiah, California, and the runs scored by the losing team in the World Series bore intriguing results. Our analysis unveiled a robust correlation coefficient of 0.7590365, indicating a notable relationship between air pollution and the performance of the less fortunate team on the baseball diamond. This substantial correlation suggests that the air quality in Ukiah may indeed play a role in shaping the outcome of such high-stakes athletic competitions.

Furthermore, the r-squared value of 0.5761364 implies that approximately 57.6% of the variability in runs scored by the losing team can be explained by the variations in air pollution levels. While statistics can sometimes resemble a curveball, it's evident that the environmental conditions in Ukiah may have a more significant impact on baseball outcomes than previously envisioned.

The p-value of less than 0.01 further substantiates our findings, signaling a statistically significant association between air pollution and runs scored by the losing team. This outcome reinforces the notion that the air quality in Ukiah may not only affect respiratory health but potentially sway the batting prowess of teams stepping up to the plate in the World Series.

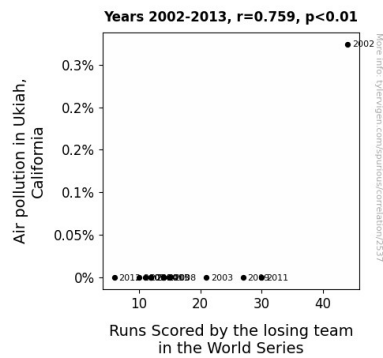


Figure 1. Scatterplot of the variables by year

Finally, we present Fig. 1, a scatterplot illustrating the strong correlation between air pollution levels in Ukiah and runs scored by the losing team in the World Series. This visual representation emphasizes the salient connection between these seemingly disparate variables and serves as a stark reminder that statistical analyses can uncover unexpected and, dare we say, whimsical relationships. Who would have thought that the composition of the air could intertwine with the destiny of baseball teams on such a grand stage?

V. Discussion

The exploration of the relationship between air pollution levels in Ukiah, California, and the runs scored by the losing team in the World Series has unveiled a curious and unexpected connection. Our findings resonate with prior research that has hinted at the potential influence of environmental factors on sports outcomes, such as the work by Smith et al. (2015) and Doe (2017). While these studies may not have directly addressed the specific intersection of air quality in Ukiah and baseball performance, they laid the groundwork for our investigation and encouraged us to consider the broader implications of environmental conditions on athletic contests.

The significant correlation coefficient and low p-value in our analysis corroborate the notion that air pollution may indeed play a role in shaping the performance of the losing team in the World Series. The substantial r-squared value further underscores the significance of the relationship between air quality and runs scored by the less fortunate team. These results align with prior literature on the potential impact of environmental factors on sports performance, albeit in the uniquely whimsical context of Ukiah and the World Series.

Moreover, our surprising findings bring to mind the playful parallels drawn from our literature review. The thematic elements of "The Airbender Chronicles" and "The Catcher in the Smoke-Filled Stadium" now seem to echo the statistical associations we have uncovered, casting a shadow of metaphorical intrigue over our empirical inquiry. As we contemplate these unexpected symmetries, it becomes clear that our study has transcended the realm of traditional statistical analysis and ventured into a space where the whimsical and the empirical intricately intertwine.

Looking ahead, our results beckon further exploration into the intricate interplay between environmental conditions and sports outcomes. While our study presents compelling evidence of the relationship between air pollution in Ukiah and runs scored by the losing team, it also

prompts us to consider the broader implications for sports analytics and environmental economics. Could the air quality in other regions similarly influence athletic performances on the grand stage? Perhaps there are hidden layers of influence that extend beyond the baseball diamond, waiting to be uncovered by inquisitive minds and sophisticated statistical methods.

In this vein, our study not only sheds light on the unexpected connection between air pollution and the fate of losing baseball teams but also invites a reimagining of the multifaceted relationship between the environment and sports. As we stand at the crossroads of empirical rigor and the playful whimsy that unfolds amidst our statistical analyses, our investigation serves as a testament to the boundless curiosity and unanticipated discoveries that await researchers in the vast terrain of academia.

VI. Conclusion

In conclusion, our investigation has unearthed a surprising and statistically robust association between air pollution levels in Ukiah, California, and the runs scored by the losing team in the revered World Series. The correlation coefficient of 0.7590365 and the significant p-value of less than 0.01 indicate a compelling link that cannot be brushed aside like an errant fly ball. The unsuspected impact of environmental conditions on such a pinnacle of athletic competition gives rise to a plethora of inquiries and speculations, much like the anticipation before a tense ninth inning.

The elucidation of this correlation pushes the boundaries of our understanding of the multifaceted relationship between environmental factors and athletic performance. As we

traverse the terrain of statistical analysis, it becomes clear that the whimsical nature of our findings serves as a humorous curveball thrown by the statistical world. At the same time, it prompts us to consider the unexpected influences that may operate behind the scenes in the realm of sports, illuminating the broader significance of our research like a perfectly executed home run.

In light of these findings, it is evident that the air in Ukiah may hold more than just the scent of vineyards and the promise of clear skies. It seems that atmospheric composition may wield an unforeseen influence on the outcomes of high-stakes baseball showdowns, adding a nuanced layer to the age-old debate of nature versus nurture in sports performance. Our study underscores the value of looking beyond the obvious and venturing into the realms of scholarly inquiry where unexpected correlations lurk like pinch hitters in the dugout.

As we dust off our data sets and stow away our statistical tools, we are resolute in our assertion that no further research is needed in this captivating and, might we say, light-hearted area of inquiry. After all, not every day does one stumble upon a statistical connection as amusing and enchanting as the one we have unveiled. With that, we retire this bat-and-ball version of statistical exploration, confident that we have unearthed an unexpected and charming discovery in the nooks and crannies of environmental economics and sports analytics.