The Tulane Journal of Policy & Political Economy

Volume 4 | Spring 2025

The Tulane Journal of Policy & Political Economy

Volume 4 | Spring 2025

"Research is formalized curiosity. It is poking and prying with a purpose. It is a seeking that he who wishes may know the cosmic secrets of the world and they that dwell therein."

Zora Neale Hurston

Tulane Faculty Review Board

Staff

Programming Directors
Autumn Sommers
Lindsey Cliff

Editor-in-Chief
Skylar Kil

Managing Editors
Ellie Cowen
Nelson Spiegel

Programming Team
Emily Maran
Audrey Andel
Peter Carroll
Kate Litton
Noah Fisher
Ittai Ronen
Ahana Louzado
Meyer Polanco
Kate Vo
Andrew Metzger
Eva Anagnostopoulos
Mary Ochiai

Chief Operating Officer
Zac Silfin

Editors
Ana Del Roio
Erin Malloy
Izzy Epstein
Kate Litton
Noah Meyer-Herron
Rachel Costello
Sophie Brandman
Ryan Lidsky

Layout Editor
Campbell Harris

Photography By
Miles B. Jordan

About the Journal

The Tulane Journal of Policy and Political Economy is an internationally recognized undergraduate publication dedicated to advancing the intellectual opportunity of students throughout the world. Founded and run exclusively by Tulane students, the Journal is committed to elevating conversation and debate through the promotion of undergraduate research.

The Journal maintains official partnerships with the Tulane Office of Undergraduate Research and Fellowships, the Tulane Departments of Political Science and Economics, the Tulane School of Public Health & Tropical Medicine, and the Politics, Philosophy, and Economics (PPE) Societies of the University of Oxford and the London School of Economics.

Beyond its publications, the Journal organizes regular events on significant political economy topics. In the past academic year, we hosted "Ballot Battle: The 2024 Presidential Election," which invited academics and representatives from the Pelican Institute to discuss the implications of the 2024 election on key issues like immigration, economic policy, and Supreme Court reform.

To submit your own research and to review submission guidelines, visit **tulanejournal.org/submit.**

For more content, visit
tulanejournal.org

A special thanks to Newcomb-Tulane College and the Undergraduate Student Assembly for their financial support.

Contents

California Drought within the Strawberry Market

Karen Medrano Gonzalez
Loyola Marymount University

Abstract

The California Drought of 2012-2016 significantly impacted water usage, leading to changes in water consumption patterns and affecting agricultural industries such as the strawberry industry. In this study, I collect USDA data related to the strawberry industry and employ econometric methods to evaluate the impact of the recent California drought on pricing in the strawberry industry. To supplement my results, I also interview three major players in the industry. I find that the correlation between the quality of products is related to the weather conditions. However, the basic principle of supply and demand is the factor that drives pricing in the strawberry industry. The quality of the goods sold can be highly influenced by the drought leading to price changes.

Introduction

The California drought, experienced from 2012 to 2016, significantly impacted policies and legislation related to water usage, leading to changes in water consumption patterns.[1] Water is crucial for agricultural products, playing a major role in this industry. Previous literature has suggested a correlation between drought conditions and the cost and negotiation of open price markets. This study aims to explore the correlation between the drought's impact and the pricing of strawberry markets. Through analyzing market data, regressions, and social context, the goal of this paper is to understand the extent of the drought's impact on California's strawberry market. The results indicate various factors influencing price, with the primary driver being supply. However, water availability is a critical factor affecting the quantity of the product. Seasonal production of strawberries also significantly influences pricing, with quality and crop yields proving significant in the pricing sector.

Methods and Data

The following section of the paper will use a mixed-methods analysis to determine the impact of the drought. The first portion will review the literature on the agricultural

1 Ehlers, Rachel. "What Can We Learn from How the State Responded to the Last Major Drought?" Legislative Analyst's Office, May 13, 2021. https://lao.ca.gov/Publications/Report/4429.

market as a whole, with a specific focus on the elasticity of the strawberry market. This section will also address the impact of legislation on changes affecting the strawberry industry, as well as the cost of production and its correlation with water usage. The second portion of the literature review will analyze data from the Strawberry Commission Advisory Board, emphasizing the number of strawberries planted concerning seasonal and county changes.

The second part of the project will focus on empirical research. The first stage will involve three interviews with farmers who sell strawberries in the market. A standardized questionnaire will assess their perspectives on how the drought years have affected their operations. The second stage of the analysis will involve a regression model, utilizing data from the California Processing Strawberry Advisory Board and the California Strawberry Commission to determine the number of strawberries planted. USDA data will be used to analyze the average retail price of strawberries during the drought, treating this as a binary variable. The regression analysis will employ an ordinary least squares (OLS) model.

Defining the Drought

Drought is defined based on its impacts on water users. California is a big state and impacts vary with location. The California State Department defines drought as the following:

Individual water agencies may use criteria such as rainfall/runoff, amount of water in storage, or expected supply from a water wholesaler to define their water supply conditions. Drought is a gradual phenomenon, occurring slowly over some time.[2]

Defining when a drought ends is based on moderating drought impacts on water users. A city may define the end of a drought when its reservoir is full or it receives a full supply from the wholesale water agency. A rancher might define the end of a drought after enough precipitation falls to adequately support livestock grazing.[3] Through this, we understand the state's role in the drought influences legislation that could affect the usage of water. The drought is not simply an environmental phenomenon that impacts production but creates limitations to the supply and demand chain.

Political Background and Effects

California's water supply has been controlled by federal, state, and local authorities. However, evidence has demonstrated that the California drought has been a cyclical process throughout history. The historical record of the drought can be found in Figure 1. This figure shows the average annual runoff and precipitation for historic droughts from 1928 to 2016. While the most severe drought took place in 1976-1977, the water supply impacts were more severe in droughts with longer durations.[4] For this reason, California has famously built aqueducts to supply areas in Southern California with adequate water.[5] The most recent drought from 2014-2017 is a repetition of past events. On Janu-

2 California, State of. "Drought." Department of Water Resources, February 27, 2025. https://water.ca.gov/Programs/All-Programs/Drought#:~:text=Defining%20when%20drought%20ends%20is,storage%2C%20can%20take%20multiple%20years.

3 Ibid.

4 Duhamel, Jonathan. "California Drought Caused by Nature and Politics." Arizona Daily Independent, August 24, 2015. https://arizonadailyindependent.com/2014/02/18/california-drought-caused-by-nature-and-politics/.

5 Ibid.

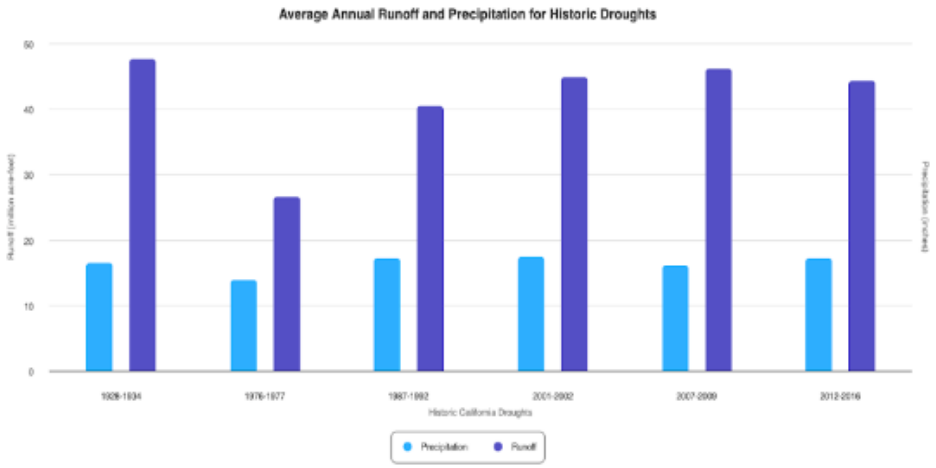


Figure 1: Reproduced from: U.S. Geological Survey, California Water Science Center. "California Droughts Compared." Accessed June 11, 2024. https://ca.water.usgs.gov/california-drought/california-drought-comparisons.html.

ary 17, 2014, California State Governor Jerry Brown declared a drought state of emergency.[6] The drought crisis has been estimated to cost two billion in the state's economy during those years. This led to Brown taking executive orders to reduce water usage by 25%.[7] Other political figures like Senator Diane Feinstein struggled to reconcile political environmental groups with campaign contributors such as Stewart and Lynda Resnick, makers of POM Wonderful pomegranate juice and the state's largest producers of almonds and pistachios, lucrative orchard crops that cannot be followed to save water in dry years.[8] Thus, environmental activists played a major role in promoting limits to water usage for farmers, such as through the Endangered Species Act. According to Duhamel, "environmental special interests managed to dismantle the system by diverting water meant for farms to pet projects, such as saving delta smelt, a baitfish. That move forced the flushing of 3 million acre-feet of water originally slated for the Central Valley into the ocean over the past five years."[9] Politicians such as President Obama advocated that the result of the drought was an issue of global warming and urged individuals to take different measures to take care of the planet.[10] On April 2, 2017, Governor Brown lifted the drought emergency but declared that California must continue water conservation efforts.

California's Strawberry Industry

The strawberry industry is one of its top ten commodities in California. Four major counties (Salinas, Santa Maria, Oxnard, and San Diego) are the primary producers. The production of these strawberries is seasonal to each specific district. California produces 90% of the strawberries in the United States, while Florida accounts for the other 10%.[11] Thus, studying the impacts of the drought in this market can

6 Chappell, Bill. "California's Governor Declares Drought State of Emergency." NPR, January 18, 2014. https://www.npr.org/sections/thetwo-way/2014/01/17/263529525/california-s-governor-declares-drought-state-of-emergency.

7 Purdum, Todd S. "For California Drought, Political Answers Are as Scarce as Water." Politico. Accessed June 10, 2024.

8 Ibid.

9 Duhamel, "California Drought Caused by Nature and Politics."

10 Ibid.

11 "Annual Report on Price Posting for Processing Strawberries." Accessed June 11, 2024. https://www.cdfa.ca.gov/mkt/mkt/documents/annual_reports/2021_price_posting_report_with_appendices_final.pdf.

determine the significance of water in agricultural operations. This analysis explores the extent to which price is influenced by water usage and how environmental changes affect the agricultural market in California.

Agricultural Monopolies

The strawberry market is one of the few agricultural products that operates under an open market. However, this open market is greatly influenced by competitive monopolies where major processors set prices for smaller farmers and companies. This arises from uncertainties in the market. The commercialization of farming through transportation has made the agricultural industry dependent on the industrial sector. Historically, corporate farming became prominent in the 1960s, comparable to the competitive market of the oil industry. Thus, the strawberry industry operates under a competitive monopoly; therefore the smaller farmers react as price takers in this industry.[12] Additionally, elasticity plays a role in pricing.[13] Originally, the system of pricing was based on the bulk line concept, meaning a system of assembly lines to produce a product, providing insight into the average cost of the product.[14] However, the strawberry market is less adaptable to product substitution in comparison to other types of berries such as blackberries. In this case, research studies directed by the Agriculture and Applied Economics department found the demand for strawberries was the least elastic type of berry with its price elasticity of –1.26.[15] Other produce such as blackberries have an elasticity of -1.88.[16] Thus, since strawberries represent the more inelastic commodity of the two, it is easier to negotiate a standard price since fewer commodities would substitute the strawberry itself. Additionally, it is important to note that strawberries have different types of varieties that factor into the different number of flavors the good can produce. Within the strawberry sector, different types of sizes and flavors are options for consumers. This allows for the strawberry to offer more flexible terms in pricing since each variety of strawberry encourages consumers to have different preferences, within the same industry. With this understanding of the competitive market for strawberries, we can analyze how pricing works across California.

Background on Pricing

There are three primary types of pricing systems utilized across companies. These include supply and demand pricing, quoted price, and negotiated price. Demand pricing is when firms adapt the price that is most likely to sell to customers. The quoted price is the lowest price a firm is willing to sell their product. The negotiated price is when the value of the product becomes open to consumer and seller conversation in determining the value. In the industry of agriculture, specific methods such as calculation, organized speculation, and hit-or-miss guessing are commonly employed.[17]

12 Hassler, John, Per Krusell, and Conny Olovsson. "Oil Monopoly and the Climate." The American Economic Review 100, no. 2 (2010): 460–64. http://www.jstor.org/stable/27805039.

13 Black, John D. "Elasticity of Supply of Farm Products." Journal of Farm Economics 6, no. 2 (1924): 145–55. https://doi.org/10.2307/1229807.

14 Lauck, Jon. "American Agriculture and the Problem of Monopoly." Agricultural History 70, no. 2 (1996): 196–215. http://www.jstor.org/stable/3744533.

15 Sobekova, Kristina, Michael R. Thomsen, and Bruce L. Ahrendsen. "Market Trends and Consumer Demand for Fresh Berries." AgEcon Search, January 1, 1970. https://ageconsearch.umn.edu/record/164771/?ln=en.

16 Ibid.

17 Heien, Dale. "Price Determination Processes for Agricultural Sector Models." American Journal of Agricultural Economics 59, no. 1 (1977): 126–32. JSTOR, https://doi.

An automatic type of pricing emerges naturally within the free market, characterized by its flexibility and independence. For instance, strawberries and blueberries operate within such a system, where prices fluctuate based on market forces and consumer demand without direct intervention. Thus, the intervention of price posting has been employed to stabilize the market.

Price Posting

One of the last open markets in California is strawberries.[18] One of the major factors in assisting in setting a price and evaluating the market is price postings. Price posting allows processors to have criteria for where to begin negotiating with farmers. Before price postings were implemented and regulated in the market, prices operated under a series of rumors. Many farmers and processors had no idea of what their competition was truly offering.
However, price postings have flaws, such as failing to record fixed prices and contracts and not holding other processors accountable in different states or nations.[19] The price posting controls the market due to a lack of union between farm growers. Many other crops such as corn have a union to negotiate a fair price with larger companies. This allows the accessibility and flow of the open market. A survey conducted by the California Processing Strawberry Advisory Board and the California Strawberry Commission was held with processors to determine the usefulness of price posting. When processors were evaluated in a survey concerning price postings, the results, shown in Figure 2, revealed an even division, with no clear indication that they contributed to achieving market equilibrium. However, overall evaluation of surveys conducted with fourteen distinct processors showed price posting allowed for communication and standardized pricing.

Survey Statements That Processors Were Asked to Express an Opinion About	Processors Expressing Some Form of Agreement	Processors Expressing Neutrality	Processors Expressing Some Form of Disagreement
1. Price posting promotes the statutory goal of preventing fraud, misinformation, deception, and unfair trade practices in the processing strawberry industry.	9 (64.29%)	3 (21.43%)	2 (14.29%)
2. Price posting provides an accurate and reliable source of raw product price information that is valuable to processors, producers, and buyers for their decision making.	10 (71.43%)	2 (14.29%)	2 (14.29%)
3. Price posting allows market forces to more smoothly reach an "equilibrium price" (the price that attracts the volume of fruit equal to the quantity demanded).	7 (50.00%)	5 (35.71%)	2 (14.29%)
4. Price posting enables processors to more effectively compete for available strawberries.	8 (57.14%)	4 (28.57%)	2 (14.29%)
5. Price posting promotes stability across the entire strawberry industry by enabling there to be better informed natural competition between the fresh sector and the processing sector.	9 (64.29%)	2 (14.29%)	3 (21.43%)
6. The current price posting provisions, including existing exemptions, are working fine and are not in need of change.	6 (42.86%)	5 (35.71%)	3 (21.43%)
7. The strawberry industry would be worse off without price posting.	9 (64.29%)	4 (28.57%)	1 (7.14%)
8. Price posting should be continued for another year.	9 (64.29%)	4 (28.57%)	1 (7.14%)

Figure 2: Reproduced from: California Department of Food and Agriculture, 2023 Price Posting Report with Appendices, (Sacramento, CA: California Department of Food and Agriculture, 2023), https://www.cdfa.ca.gov/mkt/mkt/documents/annual_reports/2023_price_posting_report_with_appendices.pdf.

What Factors Employed Are Involved in Pricing?

The decline of rural America once stemmed from an overflow of supply, but now modern-day policymakers are grappling with the issue of expensive food. This phenomenon is largely attributed to the inflationary effects triggered by increases in wages.[20] Several factors contribute to the determination of prices, including lagged prices, production costs, government programs, weather conditions, and the overcarry of supply.[21] These variables collectively shape the pricing landscape, influencing the affordability and accessibility of food within rural communities. Figure 3 provides a table of the distribution of the costs

org/10.2307/1239616. Accessed May 16, 2024.

18 Ibid.

19 "Annual Report on Price Posting for Processing Strawberries." Accessed June 11, 2024. https://www.cdfa.ca.gov/mkt/mkt/documents/annual_reports/2021_price_posting_report_with_appendices_final.pdf.

20 Heien, "Price Determination Processes for Agricultural Sector Models."

21 Kubota, Chieri. "Hydroponic Strawberry Costs & Economics." Hydroponic Strawberry Costs & Economics. Accessed June 10, 2024. https://cales.arizona.edu/strawberry/Hydroponic_Strawberry_Information_Website/Costs.html.

Table 1.Capital costs estimated for a 10,000 sq ft greenhouse strawberry production.

Cost items	Capital costs	Amortized capital costs ($/ft^2)
Plastic ground cover	$951.40	$0.04
Walk-in cooler (100 ft^2)	$6,638.42	$0.12
Strawberry troughs	$6,645.99	$0.16
Irrigation equipment	$8,346.22	$0.31
Total costs	**$22,582.03**	**$0.63**

Table 2. Variable costs ($ per sq ft) for a 10,000 sq ft greenhouse strawberry production in Tucson, AZ.

Cost categories	Capital items	Cost ($/ft^2)
Amortized capital		$0.63
Labor		$1.54
Utilities	Electricity	$0.32
	Natual gas	$0.32
	Water	$0.12
Consumable materials	Plastics	$0.02
	IPM	$0.06
	Disinfectant	$0.01
	Bees	$0.09
	Plants	$0.13
	Substrate	$0.29
	Fertilizers	$0.27
Total costs		**$3.81**

Figure 3: Reproduced from: Reproduced from: Chieri Hydroponic Strawberry Costs & Economics (Kubota 2024)

for strawberries. Water makes up approximately 3.14% of the costs. The irrigation system for the water makes up 49.2063% of the amortized capital costs.

Exploring the Market of Strawberries

Analysis of the strawberry market reveals a significant correlation between seasons and pricing.[22] Figure 4 shows that the major suppliers are Santa Maria and Salinas. Most of these counties were affected by the drought, throughout the analysis we will determine the lengths of the effect when it comes to their earnings and quality. One thing that is clear about the strawberry business is that seasons influence pricing, with a higher price range found from January through May. This is a result of holidays such as Valentine's Day and Mother's Day, as well as fixed contracts with schools, according to Anonymous Interviewee #1, who contributed to this study.[23] The counties that the three farmers interviewed for this study came from were different and ranged in size. The decline in sales and prices occurs in the summertime. When the beginning of the school year reemerges in September, the price begins to increase as well. Farmers and salesmen interviewed for this study attributed that pricing was a cyclical matter, instead of a national market or weather impact. However, beginning the process of planting strawberries is subject to the weather of a specific county; the weather is what truly tells farmers when they can begin their production. That is why each county has a different range of seasons in which they plant their production.

22 "Annual Report on Price Posting for Processing Strawberries."

23 Anonymous Interviewee # 1, personal communication, June 2024.

2022 California Strawberry Acreage Projections		
District	Acreage (% of State Total)	Season
Watsonville / Salinas	12,813 (33.7%)	April – Nov.
Santa Maria	15,175 (39.9%)	March – Dec.
Oxnard	9,824 (25.8%)	Jan. -June, Sept.- Dec.
Orange County/ San Diego	215 (0.6%)	January – May
San Joaquin	0 (0.0%)	February – June
Total Acreage	38,026	

Figure 4: Reproduced from: California Department of Food and Agriculture, 2023 Price Posting Report with Appendices, (Sacramento, CA: California Department of Food and Agriculture, 2023), https://www.cdfa.ca.gov/mkt/mkt/documents/annual_reports/2023_price_posting_report_with_appendices.pdf.

Data and Statistics

The following tables below were recorded by the California Processing Strawberry Advisory Board and the California Strawberry Commission. They collected the data from several processors.

Figure 5 shows general growth in strawberry production in most counties except Fresno and San Diego. However, a general decline in production occurred in 2013, as the drought in California began to take effect. However, a general decline in production occurred in 2013, as the beginning of the drought years in California took effect. The year 2019 is recorded to be the year with the lowest level of production in comparison to 2005. Interestingly, Santa Maria is one of the few counties that increased strawberry production for most of the time period.

Figure 6 provides insight into the number of strawberries processed, meaning those that go into a freezer contracted by a different company. The graph demonstrates a great increase in processing in 2013, the beginning of the drought years. The graph shows that in 2016 there was an increase in the volume of strawberries processed in California. Years 2015 and 2016 are the highest processing volumes record, exceeding 600 million pounds; however, that has slowly decreased.

Figure 7 demonstrates that production trends in strawberries tend to increase upwards as the years progress. The graph shows a big decline in production in 2019. However, Figure 5 represents the true decline during the drought years.

Figure 8 exhibits strawberry prices

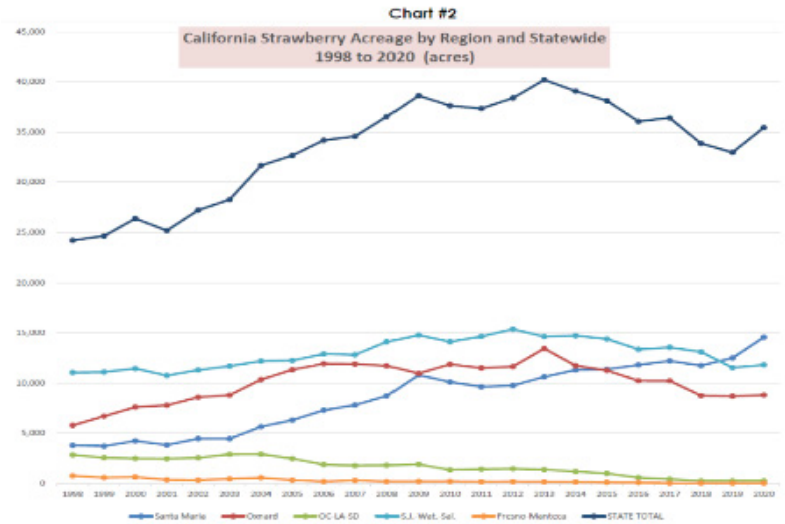


Figure 5: Reproduced from: The California Processing Strawberry Advisory Board and the California Strawberry Commission (Annual Report 2024)

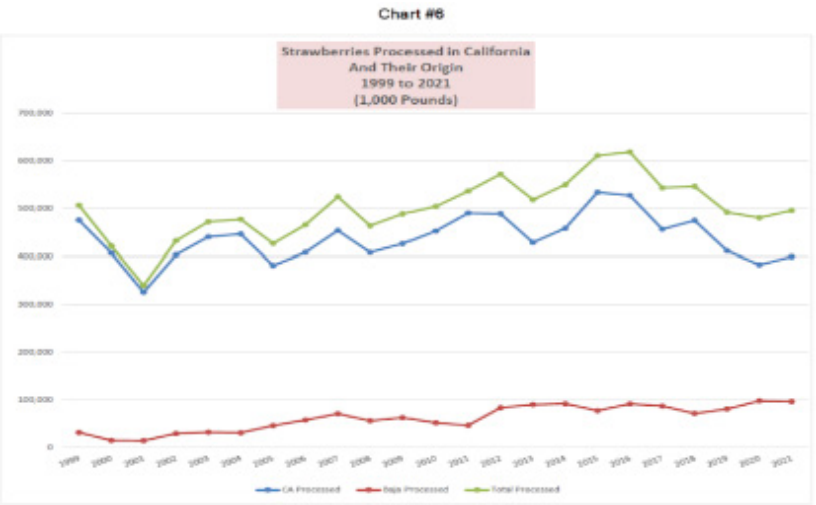


Figure 6: Reproduced from: The California Processing Strawberry Advisory Board and the California Strawberry Commission (Annual Report 2024)

throughout the years. The data demonstrates a major increase in demand and prices in the years 2013 through 2016, particularly for frozen strawberries and processed juice. However, there was a decrease in prices for fresh strawberries from 2014 through 2018. The question remains of why there would be such a decrease in drought years for fresh strawberries but an increase for frozen strawberries and juices.

One of the theories for this price decrease could be the quality of the fresh strawberries. Bad quality is categorized under excess green color, dryness, pest damage, sunburn damage, water-soaked damage, presence of cat faces (a type of exterior blemish), underdevelopment, softness, dirt, decay, and mold.[24] The multiple interviews conducted for this project indicated the drought caused a series of plagues that farmers grappled with during this period. Thus, the quality of the product became an issue that many farmers experienced.

Analyzing Qualitative Data

24 Anonymous Interviewee # 1, personal communication.

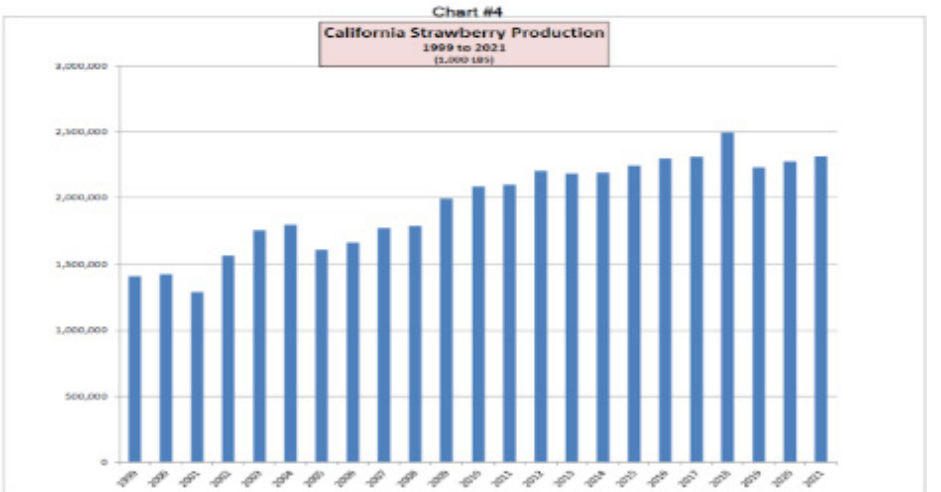


Figure 7: Reproduced from: The California Processing Strawberry Advisory Board and the California Strawberry Commission (Annual Report 2024)

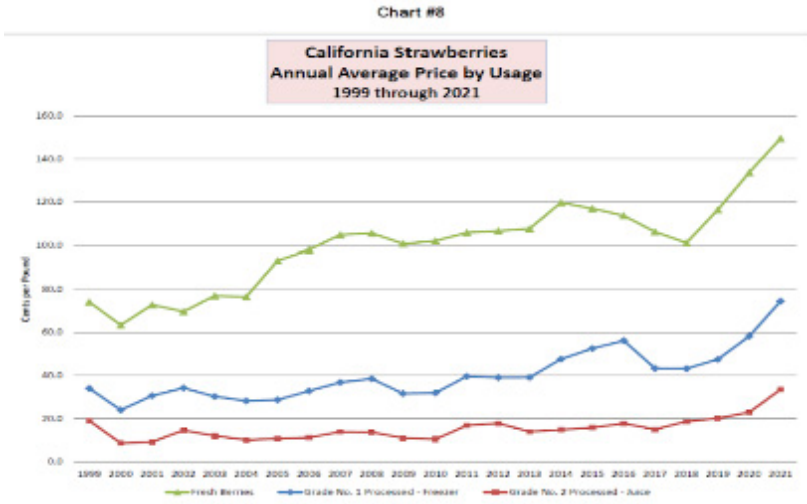


Figure 8: Reproduced from: The California Processing Strawberry Advisory Board and the California Strawberry Commission (Annual Report 2024)

The U.S. Drought Monitor gathered the data for the qualitative effects of the drought, a partnership between the National Drought Mitigation Center at the University of Nebraska-Lincoln, the United States Department of Agriculture, and the National Oceanic and Atmospheric Administration. When observing the qualitative data of the drought effects in California, we can determine the drought primarily affected the industry of farm animals and cattle. However, many more effects resulted in economic losses and innovative developments for water conservation. The results suggested insufficient water for individuals to support their agricultural production throughout these years. For example, qualitative data strongly suggested many farmers struggled to feed their cattle during these years. Interestingly, political action and conflict were prominent. The proposition of incorporating new legislation to limit the water used became a battleground. For example, one issue was that the California water board denied a request to pump irrigation water in Southern California. Other economic activities and losses included farmers purchasing insurance for their crops, changing the types of crops produced, and reducing their revenue by $2.7 billion. All these issues came to produce grants that requested an increase in innovative developments to conserve water.

Farmers and Salesmen Interviews

After conducting various interviews with farmers who plant strawberries and sell them on their account, many different conclusions have been reached about the drought impact. Many farmers have agreed that the price throughout the years has not fully changed. Instead, a shift in the production cost has impacted the profits made by these farmers. During the drought, some farmers planted strawberries of the same quantity, while others planted strawberries of lesser quality, so one thing is clear: quality throughout drought years diminished. This is a result of plagues that become prominent as a result of less water. There are two theories regarding the factors that influence pricing according to the interviews conducted. Some farmers believe that the price is adjusted according to the quantity planted that arises from different seasonal spikes. Other farmers believe that

prices become stable with an ensured quality to the consumer.

Regression Analysis

In analyzing the effect of severe drought on the retail price of strawberries, the following main equation was run:

$$Retail\ Price_t = \beta_0 + \beta_1 Drought_t + \beta_2 Acres_t + \beta_3 Acres_{t-1} + \varepsilon_t \quad (1)$$

In Equation (1), the dependent variable Retail Price is the real retail price of strawberries, adjusted for inflation using Consumer Price Index data from the Bureau of Labor Statistics and available from 1992 to 2003. The main independent variable of interest, Drought, is a binary indicator for severe drought, available from 2012 to 2019. Acres are available from 1999 to 2021. t represents the year, and an error term is included in the equation. The model is run using Ordinary Least Squares and the Stata statistical software.

The data gathered to run these regressions were taken from public data and statistics found in the USDA department, the California Processing Strawberry Advisory Board, the California Strawberry Commission, and the U.S. Drought Monitor. However, this study is not without its limitations. The small number of observations, particularly for the drought variable, has halted dependent variables in their classification of significance when it came to testing.

The previous table follows the dependent variable of the real retail price in relation to the independent variables of drought, acres, and lagged acres.

The first column in Table 1 shows the relationship between the retail price of strawberries and a binary indicator for severe drought. Results from this column indicate that there is a significant effect (at the 10% level) of severe drought presence on the retail price. Acres are added as an independent variable in the second column. Here we see a stronger positive effect of the indicator for severe drought on retail price (significant at the 5% level) and a significant negative relationship between the number of acres and retail price, reflecting the possible downward pressure on price due to more competition. The last two columns show the effect of acres alone, with lagged acres (lagged by one year) showing a slightly larger magnitude. The binary variable of the drought proves significant at the 10% and 5% levels in columns (1) and (2), respectively, when it comes to influencing prices in the market. At first glance, the variable of acres is significant at the 5% and 10% level as well. However, when analyzing the lagged variable of acres, it is significant at the 5% level but is negative. This indicates that as the number of acres increases, there is a decrease in the price. This follows the economic equilibrium price if there is more quantity than demanded. The variable of yield per acre is not significant in determining the price of strawberries. When analyzing the R2 of the table, we find that column 2 has an R2 of 0.797. This means that approximately 80% variation in strawberry prices during drought years can be explained by the drought and acres planted. Interestingly, column one which examines the variation in price in relation to the drought can explain 40% of the variation in prices.

The graph shown in Figure 9 compares the retail price with the real retail price accounting for inflation. In Figure 9 we can observe a steady increase in the retail price

Figure 9: Source: Author's calculations

during the drought years that inevitably increases the slope. This is in contrast to the blue dots that represent the real retail price, accounting for inflation during these years. The real price exhibits an increase and decrease throughout these years. This indicates the real price of strawberries has not changed much throughout the following years. Instead, it seems that the drought caused an impact to increase the slope but then stabilized its price midway.

Discussion

The goal of this paper was to analyze the effect of the California drought on the market for strawberries. While observing the regression analysis along with various interviews conducted to determine the impact of the drought on the strawberry market we can determine the following: The quality and quantity of strawberries allow these markets to find a market equilibrium. The strawberry industry's formation of the free market allows for a rise and decline in pricing throughout the seasonal changes, making the pricing system more complex than other types of products. The water contributes to the operation of the product itself. According to the regression results, the drought itself determined 40% of the variation in prices, throughout the years analyzed, and increased the retail price on average. The strawberry market operates on the simple premise of supply, demand, and quality of product. The drought deeply affected the quality of the product which led to an increase and decrease in prices depending on the surveyed field or strawberry firm. It is through the drought legislation and lack of water that the final product of strawberries was altered, making them less profitable years for farmers. This explains the decrease in the price of strawberries but a stable increase in production costs. Moreover, the drought has an impact on the quality of strawberries along with the sale prices. Note that a key limitation is the low number of observations when the drought variable is included. Nevertheless, the econometrics model can attribute that the price of strawberries is affected by supply in a negative slope, in alignment with the conducted interviews.

***

for acknowledgments and appendixes
please visit www.tulanejournal.com

***

Cash for Kids: Financial Incentives and Fertility

Evan Klein
The George Washington University

Abstract

This study examines the impact of governmental financial incentives on fertility rates across European countries, addressing whether government policies affect the decision to have children in the context of declining fertility rates. Utilizing data from the Organization for Economic Co-operation and Development (OECD), this research applies time series linear regression analysis to assess the relationship between the percentage of GDP allocated to family benefits and fertility rates. Several models were run, and one model would suggest that for every additional one percent of GDP spent on overall family benefits, fertility rates increase by .06 points. These findings suggest that direct support services play a crucial role in encouraging higher fertility rates, highlighting the need for targeted government policies to support families and address demographic challenges in the EU.

Introduction

In the 1940s, America, a few years after social security was first implemented, there were 42 workers per retiree. Today, that ratio is only approximately 3 to 1 and is expected to shrink even further in the coming years. This issue is hardly unique to the United States, as developed countries worldwide are facing similar issues.[1] To maintain and potentially grow a developed country's population, the average fertility rate required is 2.1 children per woman. This rate ensures the replacement of each parent and incorporates an allowance for child mortality.[2] In the United States, the fertility rate is approximately 1.7. Moreover, the average fertility rate across the European Union is approximately 1.53 births per woman.

Numerous studies in developed nations indicate that individuals often desire more children than they currently have. For instance, a study conducted by Ohio State University showed that young women in the U.S. reported a desire to have an average of 2.1 children. However, the actual fertility rate in the U.S. remains at 1.7.[3] The primary reason cited is the financial burden associated with raising children. Governmental benefits could provide the

1 Poole, William, and David C. Wheelock. "Aging Population: Too Few Working, Too Many Retired." Saint Louis Fed Eagle, December 9, 2021. https://www.stlouisfed.org/publications/regional-economist/april-2005/the-real-population-problem-too-few-working-too-many-retired.

2 Craig, J. "Replacement Level Fertility and Future Population Growth." Population Trends 78 (1994): 20–22.

3 Grabmeier, Jeff. "Falling Birth Rate Not Due to Less Desire to Have Children." Ohio State News, January 12, 2023. https://news.osu.edu/falling-birth-rate-not-due-to-less-desire-to-have-children/.

necessary support for those who wish to have more children but are constrained by economic factors.

Families make rational decisions about the number of children to have by weighing the costs (financial, time, opportunity costs) against the benefits (labor contributions, emotional satisfaction, security in old age) of children.[4] Central to this model is the concept of children as 'consumer durables' or 'capital goods',' providing utility or satisfaction to their parents over time. This study found a reduction in the cost of having children to increase the utility of having children, leading to a higher fertility rate.

The relationship between governmental financial incentives and fertility rates is key for a country to best manage and take care of its population. Europe presents a unique case study due to its varied social and economic landscapes and differing national family policies, as well as the general trend of declining fertility rates across its member states. This paper determined the impact of financial incentives in European countries. By using a time series regression approach, this study determined the effectiveness of financial interventions as a policy tool to mitigate the challenges posed by declining fertility rates.

The debate over pronatalist policies in industrialized countries with below-replacement fertility rates is polarized between those advocating for government intervention to encourage higher fertility rates and those opposing such measures on moral or political grounds, arguing that fertility decisions are private and skeptical of policy effectiveness.[5] Countries such as France and Luxembourg represent the former stance with explicit pronatalist policies, while the latter viewpoint is more widespread, questioning the legitimacy and impact of governmental involvement in influencing population growth. This disagreement encompasses both the effectiveness of such policies and the appropriateness of government intervention in personal fertility decisions.

The paper first discusses several academic articles written on the subject of government policies' influence on fertility. Next, I discuss the data and where they came from, followed by the regression model used. Finally, I present the results along with a discussion of what those results mean, along with some other regressions run and what they contribute.

Literature

Most fertility studies have shown an increase in fertility rates because of government programs aimed at reducing the financial cost of a child. Laroque (2008) found that an unconditional child benefit in France with a cost of 0.3% of GDP might increase total fertility by approximately 0.3 points.[6] The benefit of first-kids and third-kids was far greater than that of second-kids. The model employed was a discrete choice framework for analyzing fertility and participation decisions.

Chuard's (2021) study estimates the effect of the Baby bonus policy instituted in

4 Warren, C. Robinson. "The Economic Theory of Fertility Over Three Decades." Population Studies 51, no. 1 (1997): 63-74. DOI: 10.1080/0032472031000149736.

5 Gauthier, Anne H., and Jan Hatzius. "Family Benefits and Fertility: An Econometric Analysis." Population Studies 51, no. 3 (1997): 295–306.

6 Laroque, Guy, and Bernard Salanie. "Does Fertility Respond to Financial Incentives?" CESifo Working Paper Series No. 2339, June 2008. http://dx.doi.org/10.2139/ssrn.1157260.

2009 between Swiss cantons, which are regions within Switzerland.[7] The article uses differences in differences when comparing cantons just now instituting baby bonuses with those who already had them. They used an event study DiD to analyze changes over time to determine whether there were any long-term effects. Additionally, it facilitates the inclusion of controls for fixed effects by canton and year, as well as for trends within each canton. The article showed that the effect of the baby bonus on fertility and newborn health was positive; however, the effects declined over time until no benefit was observed today. The benefit, approximately 200-350 USD per month per child depending on the Canton, is seen as a small bonus for a rich country such as Switzerland but enough to make a difference temporarily.

A review from Israel is another example of comparing populations now receiving a benefit to those who had it previously. This study, however, highlighted how differences in culture strongly influence the fertility rate and how monetary bonuses affect each group.[8] In Israel, a child allowance was given monthly, with the monetary value increasing per child. However, for a long time, the benefit was tied to benefits for members of Israel's military. For most citizens, military service is mandatory, but for the nation's Arab population, most choose not to serve and do not receive the benefit. In 1994, the benefit was no longer tied to service; thus, Druze, Bedouin, and Muslim populations were studied to determine whether fertility rates rose once they started to receive the benefit. The difference-in-differences method was applied to the Druze and Bedouin populations, as both had a significant amount of time serving and already receiving the benefit. For the Muslim population, however, almost all are nonveterans, so a different method of comparing income levels was used. Comparing the Muslim population to the Druze population was considered but was ultimately deemed too different culturally to compare. The data used are from the State of Israel's Population Registry. This registry encompasses all Israeli citizens, documenting comprehensive details about every birth within the country and including ethnic group information. It was found that for Druze, the fertility rate for women within pregnancy increased by .12. For the Bedouin and Muslim populations, there was seemingly no impact.

Another study in Spain investigated the impact of a universal child benefit on fertility and explored the impact of both introducing and later canceling the benefit. This research focused on Spain's substantial lump-sum maternity grant, which was introduced in 2007 and subsequently abolished in 2010. This was one of the most generous baby bonuses observed at 2,500 euros to all new mothers paid at the time of birth.[9] By utilizing administrative, population-level data, the authors compile a panel dataset and use a panel regression model to determine the impact on fertility. To determine the effect of the removal of this policy on abortions, the authors used regression discontinuity, as abortions are more immediate than birth rate and can

7 Chuard, Catherine, and Philippe Chuard-Keller. "Baby Bonus in Switzerland: Effects on Fertility, Newborn Health, and Birth-Scheduling." Health Economics 30, no. 9 (2021): 2092–2123. https://doi.org/10.1002/hec.4366.

8 Frish, Roni. "The Effect of Child Allowances on Fertility in Israel." Israel Economic Review 6 (2008): 1-22. https://www.boi.org.il/media/rncp5c21/iser_1.pdf.

9 González, Libertad, and Sofia Karina Trommlerová. "Cash Transfers and Fertility: How the Introduction and Cancellation of a Child Benefit Affected Births and Abortions." Journal of Human Resources (2023). https://doi.org/10.3368/jhr.59.1.0220-10725r2.

react faster to policy.

The findings indicate that the initiation of the policy led to a 3% increase in birth rates, driven by both a decrease in abortions and an increase in conceptions. However, the cancellation of the benefit caused a 6% decline in birth rates. A detailed analysis revealed that the initial positive effect of the benefit's introduction on fertility was largely observed among highly educated parents, while the negative impact of the cancellation was more pronounced among lower-income parents.

Another study investigated the impact of a specific policy introduced in Cyprus that provided a non-means-tested monthly cash child benefit exclusively to families with at least four children. This policy aimed to financially support larger families and was studied to determine whether this policy would promote those with three kids to have a fourth child.[10] To analyze the policy's effect on fertility rates, the authors used a difference-in-differences design, using families with three children as the treated group and families with two kids as the control group. The authors believe that this model is ideal because the decision to have a third child is similar to the decision to have a fourth child, but they would not receive the benefit. The data for this analysis came from the Cyprus Family Expenditure Surveys conducted before and after the policy's introduction. These surveys provided information on household composition, income, and other demographic characteristics, facilitating a comprehensive assessment of the policy's impact on fertility. The conclusion drawn from the study is significant: the introduction of the cash child benefit led to an increase in the likelihood of families having a fourth child by approximately 5%. This effect was observed without any corresponding increase in the probability of having more than four children, suggesting that the policy specifically incentivized the transition from three to four children. However, there was no significant change in birth rate for families with fewer than three children, indicating that the policy's impact was precisely targeted.

One study investigated the impact of different family policy expenditures on fertility outcomes in 16 Western European countries.[11] Using individual-level fertility data and country policy expenditure information, the analysis employs a statistical model that integrates both personal characteristics and macroeconomic factors. The model, supplemented by Monte Carlo simulations, shows how changes in policy expenditures influence the timing of births and the total number of children, accounting for individual and country-level variations over time.

The findings reveal the distinct impacts of different family policy programs on fertility decisions. Family allowances, designed to alleviate the financial burden of child-rearing, have no significant effect on either the timing of childbirth or overall fertility rates. This suggests that simply subsidizing the direct costs of children may not be sufficient to influence fertility decisions in contexts where opportunity costs and changing gender roles are more pressing concerns. However, maternity and paren-

10 Lyssiotou, Panayiota. "Can Targeted Child Benefits Affect Fertility? Evidence from a Natural Experiment." B.E. Journal of Economic Analysis and Policy 21, no. 3 (2021): 921–965.

11 KALWIJ, ADRIAAN. "THE IMPACT OF FAMILY POLICY EXPENDITURE ON FERTILITY IN WESTERN EUROPE." Demography 47, no. 2 (2010): 503–519. http://www.jstor.org/stable/40800824.

tal leave benefits and childcare subsidies, which aim to reduce the opportunity costs associated with parenting and facilitate the combination of work and family life, demonstrate a more substantial impact on fertility.

Data

As the study looks to determine the elasticity of fertility in industrialized nations, it is restricted to European countries. To ensure access to data, only European countries in the OECD will be used given the OECD's commitment to providing reliable and comprehensive statistical data across a wide range of social and economic variables. This approach allows for a more controlled comparison, improving the ability to discern the impacts of policy and economic variables on fertility trends within a group of industrialized nations.

The OECD family database records the amount each country spends on family benefits as a percentage of GDP. This is found under the subsection "Public policies for families and children." Using family spending as a percentage of GDP standardizes cross-country comparisons by adjusting for the size of each economy. Family benefits as a percentage of GDP also reflects the government's prioritization of family and fertility-related policies within its overall economic policy framework, with a higher percentage indicating a stronger emphasis on family support. This can be further broken down into cash benefits, services provided, and tax breaks. All three of these family benefit types will be regressed separately to best determine how each benefit type affects fertility.
The OECD also has statistics that help us understand the cultural differences in how children are viewed. I will primarily use their information on the structure of families to help control for the diverse European landscape.

Eurostat has demographic data, including total fertility rates. The fertility rate, which is expressed as the total fertility rate (TFR), is a measure that quantifies the average number of children a woman is expected to have over her lifetime, given the age-specific fertility rates in a particular year. This number is found every year by dividing the total number of children born by the total number of women in their reproductive years, which are defined by the OECD as ages 15 to 49. A TFR of approximately 2.1 children per woman is considered the replacement level of fertility in most developed countries, which is the rate required to replace the population from one generation to the next.

Model

The study employs time-series linear regression with fixed effects, using each country's expenditure on family benefits, given yearly between 2003 and 2018. Several regressions will be run to test the effects of various outcomes, regression models, types of benefits, and control variables.
When looking at the effect of family policies on fertility rates across countries and years, a time series fixed effects model is proven to be the most logical approach to the data. This model was run with random effects and fixed effects, as these models allow us to draw further conclusions. There are several key benefits to this modeling framework that are essential for analysis in such a complicated environment.
There are several key benefits to this modeling framework that are essential for

analysis in such a complicated environment. By including fixed effects for every country, the model is able to approximately control for time-inflation factors such as cultural trends that may affect both the adoption of family policies and fertility rates. In addition, by including time indicators, the model can capture and adjust for time trends, ensuring that the estimated effect on fertility from family policy is not confused with general time-related trends.

$$Y_{it+1} = \beta0 + \beta1G_{it} + \beta2Z_{it} + \beta3W_{it} + \beta4V_{it} + \beta5U_{it} + \beta6T_{it} + \beta7\Delta S_{it} + \beta8R_{it} + \beta9WE + \beta10E + \beta11S + \beta12\,N + \beta13C + \varepsilon_{it}$$

The dependent variable, Y_{t+1}, represents the total fertility rate in the cycle after the cycle of benefits. In this case, the year after the year in question. β0 is the y-intercept acting as what the fertility rate would be if all the other variables were zero. The variable t represents the years since the start of the data collection, such that t=0 is the very first year of the data collection. The variable T represents the total number of years in the dataset, while t represents the numerical value of the year being tested. Variable i is the country being tested. G is the percentage of GDP used for family benefits, which will be changed several times for each benefit type and once as the total benefit. Several control variables are also used. The variable Z represents the share of children who grew up in two-parent households at time t, with β2 as its slope. W represents the female unemployment rate, and V is the male unemployment rate. Overall unemployment will also be tested with U but not in the same model as for males and females due to collinearity. T is the female labor participation rate. The variable ΔS represents the change in GDP growth from year to year. Variable C is whether a country has a history of communism. Finally, the variables WE, E, S, and N represent Western Europe, Eastern Europe, Southern Europe, and Northern/Nordic Europe, respectively. When examining the relationship between public spending on family benefits and fertility rates across EU countries, a positive correlation is anticipated, albeit with implications. Given the diversity of social, economic, and cultural landscapes across the EU, it is expected that increased public spending on family benefits, as a percentage of GDP, will be associated with improvements in fertility rates. The traditional belief on fertility is that lowering the cost through incentives will promote higher fertility rates; however, the decision to have a child is so pivotal in an individual's life that it might not have a strong correlation with government incentives.

As shown, the dependent variable is lagged by one cycle or one year. This is done to account for the time it takes from conception to birth. Thus, a one-year lag is assumed for government benefits and all the control variables. Additionally, I believe that if the control variables are significant, potential parents are more likely to look back at the previous year to decide whether the time is right to have a child rather than accurately predict the future.

Using the percentage of GDP as a measure of the scale of policy intervention allows for a more nuanced understanding of the policy's impact relative to the size of the economy. This method takes into account the economic context in which the policy operates, making the model robust to variations in policy timing and the economic landscape. By focusing on economic proportions rather than nominal values or announcement dates, the model can more accurately assess real economic commit-

ment to the policy and its potential to effect change. This approach ensures that the analysis remains relevant and accurate even in the face of delays between policy announcements and their actual implementation, providing a clearer picture of the policy's intended and actual economic footprint.

Family benefit spending encompasses public expenditures aimed at supporting families and children, covering financial assistance dedicated solely to these groups. This indicator excludes spending in other social policy domains, such as health and housing, which may benefit families but not exclusively. As mentioned, these countries are divided into three categories as defined by the OECD. Cash Transfers: This includes cash benefits to families with children, such as child allowances, which may vary based on the child's age. These factors also include aid benefits and family leave. Services for Families: Government expenditures in this category cover direct funding or subsidies for childcare and early education providers, payments to parents earmarked for childcare, support for young people and residential services, and family assistance services such as home help and center-based facilities for families in need. Tax breaks for families: This involves financial aid delivered through the tax system, including tax exemptions, child tax allowances, and child tax credits that lower the tax bill.

The unemployment control variables are expected to provide us with deep insight into the decisions of when and if to have a child. This is further broken down into male and female unemployment controls. Starting with female unemployment, the literature expects that the two will be negatively correlated such that when female unemployment increases, fertility rates will increase. This is because the opportunity cost of having a child decreases when wages are missing because childbearing is no longer a factor. However, it is also possible that the two are positively correlated. If a potential mother is concerned that she cannot find a job now, will that make her less likely to have a child with the future of finding a job uncertain post-childbirth? For male unemployment, I expect a positive relationship, as a father can continue to work, and the family would choose to have a child after obtaining a steady income. Similarly, a change in GDP is expected to have a positive effect as a proxy for a greater economy, and families are expected to have greater disposable income from using children when the economy is doing better.

One criticism of a regression such as this is that a society that values children highly will have both high fertility rates and very generous family benefits. Thus, the model seeks to control for each country's cultural attitudes toward family and child-rearing by incorporating a variable that captures cultural beliefs. The variable chosen is the percentage of children living in a household with two parents, as it serves as a proxy for societal norms around family structure and the value placed on traditional family units. This metric provides insight into how societal values may influence fertility behaviors and support family policies. Initially, the model incorporated the percentage of children raised by married parents as a metric. However, in countries such as Finland, marriage is not as prevalent despite high levels of commitment among couples. This is evident from the statistics, which show that in 2022, only 55% of Finnish households consisted of

married couples, the lowest among the surveyed countries. Consequently, Finland and similar countries offer less formal legal agreements for couples instead.

The historical political system, particularly the experience of communism, has left a lasting imprint on the demographic trends and policy frameworks of the affected countries. Communist regimes had specific family- and fertility-related policies, such as promoting gender equality in the workforce, which could have long-term effects on fertility patterns. Countries with a history of communism share this legacy, influencing their current demographics, structure, societal norms, and even the current-day government. Additionally, past literature has shown that post-communist countries, like other countries, struggle to combine motherhood and a career.[12]

Further models that will be used will incorporate interaction variables between regions and benefits to determine which benefits work best for which regions.

When studying fertility rates in a time-series panel dataset, the model finds a positive autocorrelation of fertility rates. Fertility rates tend not to change abruptly but rather evolve gradually over time. This inherent continuity in fertility behaviors among populations means that the fertility rate in one year is likely to be similar to the rate in the previous year. Factors influencing fertility, such as cultural norms, economic conditions, and social policies, also change slowly over time, further contributing to this continuity. A Prais-Winsten Estimation Panel model will be used to calculate the significance of this autocorrelation and to account for it.

Results

First, the results when using all types of family benefits aggregated with fixed effects. The control variables do not align with the preconceived narrative concerning fertility dynamics; however, none are statistically significant. Notably, for female unemployment, the belief of a positive correlation does not hold true as female unemployment increases; in fact, fertility rates decrease. This is likely due to the concern of not being able to reenter the

VARIABLES	(1) Fert_year_plus_one
GDP_Family_Total	0.0614***
	(0.0212)
TwoParent	0.00145
	(0.00298)
female_unemployment_20t64	-0.0132*
	(0.00738)
male_unemployment_rate_20t64	0.000110
	(0.00655)
women_labor_participation	-0.00272
	(0.00246)
Growth_GDP	-0.00183
	(0.00193)
Constant	1.630***
	(0.328)
Observations	310
Number of code	26
R-squared	0.3266

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) Fert_year_plus_one
GDP_Family	-0.00175
	(0.0104)
TwoParent	-0.00394***
	(0.00144)
female_unemployment_20t64	0.00886**
	(0.00361)
male_unemployment_rate_20t64	-0.0102***
	(0.00316)
women_labor_participation	-0.00287**
	(0.00118)
Growth_GDP	0.000583
	(0.000928)
Fert_	0.959***
	(0.0314)
Constant	0.607***
	(0.161)
Observations	310
Number of code	26
R-squared	0.9539

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

12 Šťastná, Anna, Jiřina Kocourková, and Branislav Šprocha. "Parental Leave Policies and Second Births: A Comparison of Czechia and Slovakia." Population Research and Policy Review 39, no. 3 (2020): 415–437. https://doi.org/10.1007/s11113-019-09546-x.

workforce after giving birth. Surprisingly, male unemployment seems to have very little impact on the decision to have a child. The employment rates in various countries may not change enough to induce a change in fertility; alternatively, people working in jobs more likely to be fired or struggle to find a job are less likely to be making a decision on whether to have a child. Now, regarding the variable of interest, the percentage of GDP spent on total family benefits is rather high. This coefficient is statistically significant at .06, suggesting that for every one-point increase in family benefits as a percentage of GDP, fertility increases by .06. This finding shows that government benefits strongly impact fertility rates and may be the best predictor of fertility rates in a given year. This model has an overall R-squared value of .3266.

For reference, consider a country such as the Czech Republic whose fertility rate in 2021 was 1.83, which is high for a European country. Were they to have spent an additional one percent of their GDP, approximately 2,818,000,000 USD, on cash benefits for families, the model predicts that the fertility rate would have been 1.89. As stated in the model, the dependent variable is a lagged indicator one year after the year of the independent variable. Thus, if we were to include the fertility status of the previous year, we would fix the autocor-

	Fert year plus one
Percent of GDP spent on Cash Benefits	0.179
	(0.034)
Eastern_Europe # Percent of GDP spent on Cash Benefits	
1	-0.190
	(0.044)
Nordic_Countries # Percent of GDP spent on Cash Benefits	
1	0.423
	(0.092)
Southern_Europe # Percent of GDP spent on Cash Benefits	
1	-0.243
	(0.108)
Percent of GDP Spent on Service Benefits	0.073
	(0.066)
Eastern_Europe # Percent of GDP Spent on Service Benefits	
1	-0.063
	(0.086)
Nordic_Countries # Percent of GDP Spent on Service Benefits	
1	-0.273
	(0.085)
Southern_Europe # Percent of GDP Spent on Service Benefits	
1	0.402
	(0.184)
Percent of GDP Spent on Tax-Breaks	0.045
	(0.062)
Eastern_Europe # Percent of GDP Spent on Tax-Breaks	
1	0.116
	(0.097)
Nordic_Countries # Percent of GDP Spent on Tax-Breaks	
1	-0.422
	(2.842)
Southern_Europe # Percent of GDP Spent on Tax-Breaks	
1	-0.034
	(0.115)
Proportion (%) of children (aged 0-17) living with two parents	0.002
	(0.003)
Female Unemployment Rate	-0.023
	(0.007)
Male Unemployment Rate	0.006
	(0.006)
Women Labor Participation	0.003
	(0.003)
(%) Change in GDP	-0.001
	(0.002)
Intercept	1.140
	(0.312)
Number of observations	306

relation problem and help us make sense of the control variables. Running this regression, we find that this variable is highly correlated with the dependent variable, with a coefficient of .96, and is extremely significant. The meaning of this coefficient is that the fertility status of the previous year is in fact the best predictor of fertility in the next year, which would make sense as fertility rates stay fairly similar from year to year within a country. The reason is that fertility rates all over Europe are declining due to other factors incorporated in this variable. However, the overall model does not incorporate this variable because the primary goal of this study is to understand the impact of other explanatory variables on fertility rates. Including the previous year's fertility rate overshadows the effects of these variables. The lagged fertility rate might absorb a significant portion of the variation in the current year's fertility rate, making it harder to detect the influence of other variables of interest.

Dividing the regressions now into those of different family benefit types, we find that cash benefits and tax breaks have a far greater effect than service benefits on increasing fertility rates. Cash and tax breaks both had a .11 increase in fertility rates for each one-point increase as a percentage of GDP, while service benefits had a negative effect on fertility. Given that it is irrational for them to have a negative impact, we assume that it is zero. Due to a lack of data, these slopes should not be taken as facts but rather should suggest that family policies involving giving money back to families affect fertility rates, while those that provide a service do not.

The differential impacts observed can be attributed to several reasons due to the nature of the benefits. First, cash benefits and tax breaks provide direct, immediate financial relief to families, making the economic burden of child-rearing substantially lighter. This direct financial enhancement is both tangible and immediately beneficial, thereby making these forms of aid more appealing and practical in the decision-making processes of families considering having children. On the other hand, service benefits, which include childcare, healthcare, and education services, while valuable, often come with barriers that can delay or complicate their utilization. Challenges such as accessibility, eligibility complexities, and possible bureaucratic entanglements could diminish the immediate perceived value of these services. There is also the leaky bucket theory, which essentially says that despite government spending more, the benefits of services may not trickle down in full as much as cash and tax breaks.

Another method employed in the model is the use of subgroups to further identify how nations that are in a similar context react to benefits. Countries sharing cultural or historical backgrounds, such as those with common societal norms regarding family size or historical policy legacies from past political regimes, can provide insights into the collective impact of such factors on fertility trends. Thus, two main methods of grouping were employed: one by region and one by history. For regions, countries are grouped into Nordic, Eastern, Western, and Southern Europe. For a historical marker, two groups were created based on whether they had communist influence from the days of the Soviet Union. It should be noted that the list of communist countries, while sharing similarities, is not the same as the list of Eastern Europe-

an countries.

Regional grouping is based on geographical proximity and often shares cultural, historical, and economic traits within regions. Countries within the same region tend to have similar socioeconomic development levels, cultural attitudes toward family and fertility, and comparable public policies regarding family support and welfare systems. For instance, conventional knowledge tells us that Nordic countries are known for their generous family welfare policies. Due to collinearity, the model omits Western Europe from the regression results.

After running the regressions for total benefit again, this time using random effects and broken into subgroups, there is strong evidence that different populations react to these benefits differently. However, given that these groups contain only a small number of countries and observations, they cannot be used for true elasticity; rather, they can be used for comparison. First, the Nordic countries have the highest fertility rates, followed by Western Europe, Eastern Europe, and Southern Europe. However, substituting region for an interaction variable of overall family benefits and fertility, it is found that benefits in Western European countries actually go further to impact the fertility rate than do benefits in other regions. The second is Nordic countries, with Eastern and Southern Europe being tied last.

This outcome could be attributed to the comprehensive welfare systems in place within these nations, where cash benefits are part of a broader suite of family-friendly policies designed to support parents financially and reduce the economic burden of child-rearing. Here, cash benefits are seen not only as financial aid but also as part of a social contract that values and supports family life and child-rearing. The effectiveness of cash benefits in these countries may also relate to high levels of trust in public institutions.

Another discovery is that former communist countries have overall lower fertility rates than countries that have no history of communism. In addition, the impacts of these countries on fertility rates are slightly less pronounced, estimated at approximately -.02.

Former communist countries in Europe generally exhibit lower fertility rates and a diminished impact of family benefits on these rates, a trend that can be attributed to several historical, economic, and social factors. Cultural legacies from the communist era, such as the emphasis on women's workforce participation and state-supported childcare, have shaped attitudes toward family sizes and dual-income norms, which persist today. Coupled with potentially lower public trust in government initiatives, these benefits have a weaker influence on fertility rates.

Now, the model conducts a much more robust regression analysis using an interaction term between each European region and each type of benefit. To reiterate, due to a lack of data, the slopes should not be taken as absolute truth but rather as a way to compare regions and benefit types. Once again, these slopes are compared to those of Western Europe. The full results are shown below, but a few interesting takeaways are that cash benefits seem to go very far in Nordic countries, but tax breaks do not. In the context of Nordic welfare

VARIABLES	(1) Fert_year_plus_one
GDP_Family_Total	0.0661*** (0.0166)
TwoParent	0.00540*** (0.00179)
female_unemployment_20t64	-0.0124** (0.00501)
male_unemployment_rate_20t64	0.00309 (0.00433)
women_labor_participation	0.0170*** (0.00271)
Growth_GDP	0.000320 (0.000848)
Constant	-0.283*** (0.0217)
Observations	284
Number of code	26

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

systems, cash benefits might be especially effective because they provide flexibility and immediate liquidity to families. This direct financial support can be used according to individual family needs, whether for childcare, education, or other child-related expenses. Tax breaks, on the other hand, are typically realized only when taxes are filed, and their benefits can be less tangible and immediate. The delayed nature of tax benefits might not align as effectively with the immediate financial considerations associated with raising children. Interestingly, Southern Europe seems to respond far better to service benefits than the other regions, despite having some of the worst fertility rates in Europe. Southern European cultures often exhibit strong familial bonds and intergenerational dependencies. In countries such as Italy, Spain, and Greece, extended family networks play a significant role in daily life, including childcare and elder care.[13] Service benefits that enhance support for families, such as subsidized child care, healthcare, and educational services, align well with these cultural norms.

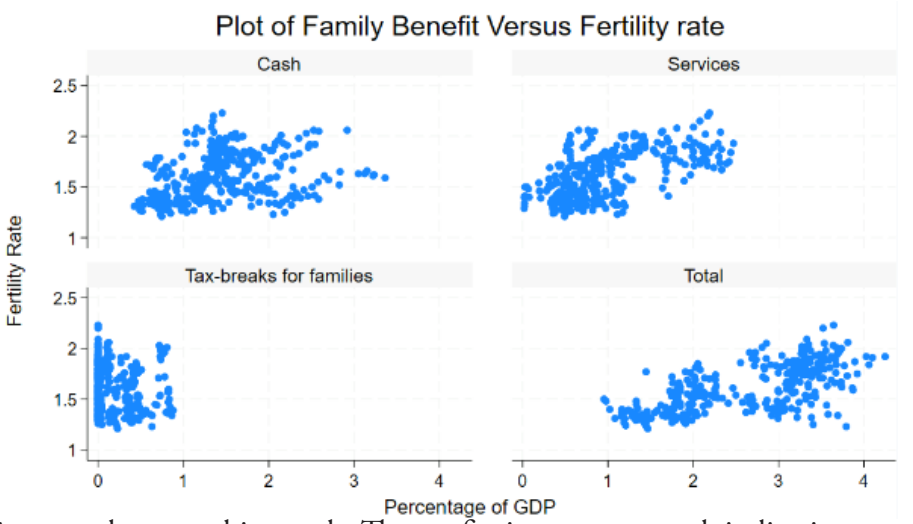
Finally, a Prais-Winsten estimation model with fixed effects was conducted to confirm the suspected autocorrelation effect. This model is specifically designed to correct for first-order autocorrelation in the error terms of a regression model. A rho_AR value of 0.88 was found, indicating a high level of autocorrelation addressed by the model. This suggests that the data points are highly correlated over time. Even so, this model still resulted in a total family benefit slope similar to the fixed effects model at approximately .066. Fertility rates are often influenced by a range of socioeconomic factors that do not change drastically over short periods. Instead, factors such as economic policies, cultural norms, and social support systems typically evolve gradually. The Prais-Winsten model effectively handles this temporal dependence by correcting for autocorrelation, thus ensuring that the influence of past fertility rates on current rates is appropriately accounted for in the analysis.

In this model, the variable concerning the share of children raised in a two-parent household is significant. As the percentage of children raised by two parents is more consistent over time than is the percentage raised by other variables (showing less variation), this effect might be obscured in a standard time series regression due to autocorrelation but revealed in the Prais-Winsten model.

Conclusion

In conclusion, this study provides valuable insights into the relationship between government financial incentives and fertility rates across European countries, leveraging extensive data from the OECD to explore how the allocation of GDP to family

13 Jappens, Maaike, and Jan Van Bavel. "Regional Family Cultures and Child Care by Grandparents in Europe." Demographic Research 27 (July 13, 2012): 85–120. https://doi.org/10.4054/demres.2012.27.4.



benefits impacts demographic trends. The findings reveal a clear positive correlation between increased spending on family benefits and higher fertility rates, with a significant increase of 0.06 points in fertility rates for every additional one percent of GDP invested in family benefits. This underscores the effectiveness of direct support services in enhancing fertility rates, thereby addressing the demographic challenges faced by the EU.

The model draws heavily on a 1997 framework developed by Gauthier, which similarly used a time series and aggregated benefits. Gauthier's study showed that a 25 percent increase in family benefits would result in a fertility level increase of 0.07.[14] My research, however, would find that such an increase would only increase fertility by approximately 0.039 given that the average percentage of GDP spent on family benefits is 2.6%.

The analysis differentiated the impacts of various types of family benefits, highlighting how their effectiveness varies across different European regions. While cash benefits and tax breaks generally promote higher fertility, the impact of service benefits is more nuanced, indicating a need for policies that are finely tuned to regional and cultural contexts.

The study's outcomes advocate for a strategic approach in designing family policies, suggesting that mere financial input in family benefits is not enough. Instead, a deeper understanding of the socioeconomic and cultural dynamics at play within different EU countries is crucial. Policies need to be contextually adapted, not only to ensure their effectiveness but also to optimize resource allocation.

Furthermore, the research illuminates the importance of considering historical and regional disparities in policy planning. For instance, while Nordic countries respond positively to cash benefits, southern European regions see greater benefits from service-oriented support. This variation necessitates a diversified policy approach tailored to meet the specific needs and preferences of each region.

While the results provide some support for a pronatalist hypothesis, they are constrained by the aggregate nature of the fertility measure and data used. Spe-

14 Ibid.

cifically, these findings do not explore the potential timing effects of family benefits or their differential impacts on various subgroups within the population. Utilizing individual-level data could yield valuable insights, and it may be helpful to examine whether family benefits influence families with different incomes or education levels in similar ways. Given that the financial burden of raising children might be relatively lower for low-income families and considering that family benefits are typically awarded irrespective of income, it is expected that these benefits would more significantly offset the costs of children in lower-income families. Additionally, investigating the potential variable impacts of family benefits based on factors such as the mother's age, marital status, and employment status would be worthwhile. Exploring the effects of other types of benefits, including means-tested benefits and those related to daycare, housing, healthcare, and education, could also provide a more comprehensive understanding of how various supports influence fertility rates.

This research suggests that European countries looking to address fertility rate declines may find greater success in implementing or extending benefits that most make sense for the culture of that country rather than relying solely on one type of family benefit. This strategic approach could provide a more effective and predictable framework for supporting families and encouraging higher fertility rates, which is essential for sustaining demographic and economic health in the face of aging populations.

***

for acknowledgments and appendixes
please visit www.tulanejournal.com

***

Corruption and the Illusion of Stability: How did Thailand and the U.S. Mask Economic Failures in 1997 and 2008?

Claudia Rogers
Boston University

Abstract

This paper examines the role of corruption in masking the economic downturns by comparing the 1997 Asian Financial Crisis and the 2008 Global Financial Crisis. Regulatory negligence and financial deception were investigated using the historical analysis of leading indicators in economic crashes to determine how corruption delayed crisis recognition. The findings suggest that corruption exacerbated economic instability and hindered policy response. This examination provides insight into the warning indicators of the financial crisis and how corruption impacts and suppresses them. The results confirm that corruption acted as a veil, preventing clarity in regulatory framework decision-making. These findings have important implications for policymakers and economic historians seeking to understand how corruption obscures the signs of an economic collapse and how to enhance future economic resilience.

Introduction

On July 2, 1997, the Thai Baht was unpegged from the United States dollar, marking the beginning of an economic crisis that would spread rapidly among South Asian countries. From there, the Baht floated, devaluing and dragging Thailand down into financial disaster, bringing the Indonesian, Malaysian, South Korean, and Philippine markets down too.[1] The interconnectedness of the Asian market contributed to a widespread economic crash. Few realized that seeds of this sudden collapse had been sown years earlier due to widespread corruption. Over ten years later in 2008, the United States felt a similar shock when Bear Stearns announced the value of its securities had plummeted.[2] The news of Bear Stearns sent shockwaves through Wall Street, revealing that the housing market was far more unstable than America had initially thought. The economies of North America and Europe were so intertwined that they both faced severe downturns starting in 2008, mirroring the

1 Park, James J. "The Asian Financial Crisis and Global Governance." East Asia Law Review 1, no. 1 (1998): 1-37. Accessed September 24, 2024. https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1039&context=ealr.

2 Lewis and Reinhardt "The U.S. Current Account Deficit." Harvard Business School Faculty & Research. Accessed September 24, 2024. https://www.hbs.edu/faculty/Pages/item.aspx?num=32498.

same patterns demonstrated in 1997.[3] Six key indicators illustrate the predictability of the 1997 Asian Financial Crisis and the framework for forecasting the 2008 Global Financial Crisis. Corruption hindered predictability and current account deficits were the most impacted by systemic corruption, making forecasting and mitigating these crises all the more challenging. This paper examines corruption in the nations where the crises originated, Thailand and the United States, exploring how corruption masked economic weaknesses and the predictability of financial crises through leading indicators. The following sections are structured as follows. Section three defines corruption for clarity, while sections four and five detail the historical context of both events. Section six presents the research, followed by section seven, which analyzes the findings, and section eight explores alternative explanations for the role of corruption. Finally, section nine concludes the paper with a discussion on recovery and implications.

Defining Corruption

Corruption is defined as the abuse of power or manipulation of systems by individuals or institutions for personal gain or to advance an agenda. These actions often occur at the expense of ethical standards, transparency, and fairness, undermining the integrity of economic, political, and social structures. Corruption masks the true state of affairs, creating illusions of stability or success. Given how prevalent corruption was in the buildup to these financial crashes, it is essential to understand the types of corruption involved and their impacts.[4] Thailand experienced cronyism, a form of corruption in which government or political figures forge deals with influential financial workers for personal gain. These mutually beneficial, corruptive relationships allowed financial institutions greater freedom while government officials profited from the appearance of a thriving economy. However, this dynamic can easily backfire when unchecked financial practices and poor oversight create vulnerabilities in the markets. These vulnerabilities include, but are not limited to, excessive debt, asset bubbles, and volatile investment patterns, leaving the economy more susceptible to shocks. Thailand's current account deficits are driven by political corruption, which created the illusion of economic prosperity but ultimately led to harsh consequences in the market. As evidence suggests, "Thailand's loss of macroeconomic stability in 1997 was an effect of political corruption."[5] Corruption obstructed proper oversight, hindering the ability to predict and address the growing financial instability before it escalated into a crisis.

The 2008 Global Financial Crisis was driven by an epidemic of misleading information in the United States. Financial institutions, real estate companies, and even regulatory bodies contributed to a culture of misinformation, obscuring the risks associated with mortgage-backed securities(MBS) and collateralized debt obligations (CDO). As described by the

3 Margesson "The Global Financial Crisis: Causes and Consequences." Congressional Research Service, 2010. Accessed September 24, 2024. sgp.fas.org/crs/misc/R40189.pdf.

4 Alonso Carrillo, María. "The Impact of Corporate Governance on Corruption Disclosure in European Listed Firms through the Implementation of Directive 2014/95/EU." Sustainability 11, no. 22 (2019): 6479.

5 Larsson, Tomas. 1998. "Democratisation, Corruption and Economic Development : the Case of Thailand." Lund University Libraries. https://lup.lub.lu.se/student-papers/search/publication/1709312.

International Monetary Fund, "widespread ingrained corruption leads to inefficiency in the allocation of resources."[6] Corruption distorted the proper functioning of economic and financial systems, preventing effective oversight and responsible investment. Widespread deception fueled reckless spending, which drove up current account deficits, simultaneously delaying corrective measures, allowing systemic issues in the market to accumulate and ultimately triggering a global market collapse.

The United States and Thailand demonstrate how corruption magnified the severity of the crises, further destabilizing already fragile financial systems. Corruption was the fan for economic systems that had gone up in flames. Both crises highlight how corruption eroded regulatory frameworks and fueled risky lending practices, ultimately causing devastating and abrupt financial collapses. Examining the effectiveness of six leading indicators for the economic crisis in 1997 will help conclude if the same indicators helped predict the 2008 crisis or if corruption hindered the indicator's effectiveness.

Origins of the Asian Financial Crisis

The 1997 Asian Financial Crisis originated in Thailand and the economic collapse was set in motion by a housing bubble and risky financial practices demonstrated in current account deficits and surmounting foreign debt, which reflected the deep rooted cronyism and inadequate oversight in the Thai financial sector. As early as 1986, ten years before the crisis, an influx of construction due to high demand and economic stability caused migration from rural to urban areas. Over time, interest rates rose and the cost of living increased, which lowered the demand for housing, but the supply still grew. The real estate bubble burst, causing investors who financed the projects to pull out, causing widespread financial instability.[7]

Simultaneously, throughout the late 1980s and early 1990s, there was a series of risky liabilities that were not backed by reserves. These liabilities and delinquent loans were made possible by the spread of political corruption, leading to increased spending.[8] To add to the level of confusion, these liabilities were tracked in foreign currency rather than the Thai Baht. When real estate investors withdrew their funding, the reserves lacked sufficient money to pay them. The liabilities were unethically and irresponsibly managed to increase the short-term payout because with lower reserves, a bank can lend more, increasing interest profit. If done too often, and without proper oversight, too few reserves will become an issue if large numbers of investors are looking to withdraw, as was the case in Thailand. During this time, political pressures on the administration drove investment banks to amass significant debt, causing unnoticed surging current account deficits. As a result, economists have linked financial crises to "higher current account imbalances"[9] and "accumulation of a larger public debt"[10], patterns that severely im-

6 Dickinson, 2000

7 Lipton, David. "What We Have Seen and Learned 20 Years After the Asian Financial Crisis." International Monetary Fund Blog, July 13, 2017. Accessed September 24, 2024. https://www.imf.org/en/Blogs/Articles/2017/07/13/what-we-have-seen-and-learned-20-years-after-the-asian-financial-crisis.

8 Leightner, Jonathan E. "Thailand's Financial Crisis: Its Causes, Consequences, and Implications." Journal of Economic Issues 41, no. 1 (2007): 61–76. http://www.jstor.org/stable/25511156.

9 Dickinson, David G., et al. Finance, Governance and Economic Performance in Pacific and South East Asia. Edward Elgar Publishing, 2000.

10 Cooray, Arusha, Ratbek Dzhumashev, and

pacted Thailand in the late 1990s.

With an economic system already teetering on the edge of disaster in the housing and financial sector, Thailand faced a third problem that would send itself, its trading partners, and neighboring nations into crisis. Investors pressured the Thai Baht, believing it to be overvalued. In a desperate attempt to preserve the currency, in July of 1997, the Thai government unpegged the Baht, allowing it to float. The value plummeted, worsening the issues in the already fragile economy.[11] The interconnectedness of the economic systems in Southeast Asia contributed to the rapidly evolving crisis. Indonesia, Malaysia, South Korea, and the Philippines quickly collapsed because Western investors pulled out of the developing Asian market after the economic failure in Thailand. With depleted reserves, shaken investor confidence, and hindered borrowing oversight, the Thai economy was exposed to crisis well before the warning signs appeared, due to economic systems infiltrated by corrupt practices.

Lead up to the 2008 Global Financial Crisis

The 2008 Global Financial Crisis began in the United States, fueled by a housing bubble and reckless lending that exposed the deception and corruption within the financial system. In 1990, the United States economy was successful and growing. The housing sector expanded rapidly to accommodate American families. As the 1990s progressed to the early 2000s, interest rates rose, and housing prices doubled. When housing prices became too high, real estate agents began selling houses to people who could not afford to pay their mortgage. The widespread distribution of subprime loans created a significant gap in the housing market, where properties were being sold, but a large number of mortgages went unpaid.[12] In 2006, foreclosures and delinquencies surged.

Meanwhile, investment banks created mortgage-backed securities (MBS), which were investment bundles that were, at their core, composed of a large number of mortgages. These investments are typically stable and predictable, as paying a mortgage is a necessary monthly expense, and millions of defaults are rare.[13] Concealing housing market issues only worsened corruption in the economy.[14] The number of subprime loans and people investing in MBS was unprecedented. During this period, collateralized debt obligations (CDO) gained popularity by taking subprime loans, dividing them into segments, and reselling them to investors. This process created bundles of loans that were fundamentally flawed. Financial institutions misled investors by packaging subprime loans as high-quality securities.[15] The American economy was based on a defaulted loan and was on the edge of collapsing because when, "debt is used inappropriately, it is possible that the debt will have a negative impact on

Friedrich Schneider. 2016. "How Does Corruption Affect Public Debt? An Empirical Analysis." Science Direct. https://www.sciencedirect.com/science/article/pii/S0305750X-16304673#s0045.

11 International Monetary Fund "Thailand's 1997 Financial Crisis: Causes and Consequences." International Monetary Fund, 2000. Accessed September 24, 2024. www.imf.org/external/pubs/ft/wp/wp0218.pdf.

12 Desilver, Drew. "Facts About the U.S. National Debt." Pew Research Center, February 14, 2023. Accessed September 24, 2024. https://www.pewresearch.org/short-reads/2023/02/14/facts-about-the-us-national-debt/.

13 Rajan, Raghuram G. Fault Lines: How Hidden Fractures Still Threaten the World Economy. Princeton, NJ: Princeton University Press, 2010.

14 Yılmaz, E.G., & Kahveci, Ş. (2024). Does the misuse of net errors and omissions arise from a corruption culture?. Journal of International Studies, https://www.jois.eu/files/9_1428_Y%C4%B1lmaz_Kahveci.pdf

15 Lewis and Reinhardt, 2005

economic growth and even threaten the country's macroeconomic stability."[16] This fact was exposed when the housing bubble popped and the value of MBS plummeted, leaving investment banks exposed.

The widespread financial instability occurred when the market for betting on the success of MBS grew in popularity. The biggest investors in MBS were firms including Bear Stearns and Lehman Brothers, whose portfolios combined held over 100 billion dollars in assets.[17] Investors worldwide were confident in these companies and the reliability of MBS, so they bet heavily on the success of the market through synthetic CDO. The market for betting on mortgage-backed securities was substantially more prominent than the underlying mortgage market, and the amount of money invested in the volatile American housing market surged rapidly in the years leading up to the financial crisis.[18]

In the U.S., corruption in the housing and financial sectors, marked by deceptive practices like subprime lending and the misrepresentation of securities, laid the groundwork for the 2008 Global Financial Crisis. As reckless investments and flawed financial mechanisms spread, economic crisis and systemic instability were seemingly inevitable.

Warning Indicators and Predictability

The warning indicators provide a clear understanding of how financial instability unfolded in both Thailand and the United States, leading up to their respective crises. One of the key signs of instability is the current account deficit as a percentage of GDP, which reflects when imports exceed exports. In 1996, Thailand's deficit reached 7.9%,[19] surpassing the recommended threshold of 4-5%, while the United States faced similar issues in 2007 with a reading of 5.3%.[20] In Thailand, the housing boom drove the rapid industrial and economic growth, increasing the demand for construction materials.[21] The United States experienced high levels of consumerism, specifically importing Chinese goods while running a large trade deficit, causing imports to exceed exports.[22] The reliance on unchecked lending practices exacerbated the trade deficit, with both nations heading toward unsustainable economic practices.[23]

Foreign debt as a percentage of GDP is another indicator of future financial issues. When this figure exceeds 30%, there are indications of negative future implications. Thailand's foreign debt reached 50.3%[24] in 1996 while the United States surpassed this benchmark at 62% in 2007.[25] In Thailand, borrowing heavily in foreign currency became unsustainable once the Baht collapsed, and they were incapable of paying back foreign debt.[26] The United States was borrowing heavily to finance homes and continuing its patterns of high consumerism, leading to increased foreign debt.[27]

16 Putra, Rommy Fernando, and Dewi Zaini Putri. 2021. "The Effect of Corruption, Democracy and Foreign Debt on Economic Growth in Asian Pacific Countries." Jambura Equilibrium Journal. https://ejurnal.ung.ac.id/index.php/equij/article/view/10272/3028.

17 Corporate Finance Institute "Lehman Brothers." Corporate Finance Institute. Accessed September 24, 2024. https://corporatefinanceinstitute.com/resources/career-map/sell-side/capital-markets/lehman-brothers/.

18 Lewis and Reinhardt, 2005

19 Park, 1998

20 Desilver, 2023

21 Kenen, Peter B. The Asian Financial Crisis: Lessons for a Resilient Asia. Washington, DC: Institute for International Economics, 2003.

22 Rajan, 2010

23 Kenen, 2003; Rajan, 2010

24 Park, 1998

25 Desilver, 2023

26 International Monetary Fund, 1998

27 Rajan, 2010

Lobbying to deregulate foreign spending encouraged the pursuit of short-term profit and high debt levels.[28]

Short-term debt as a percentage of GDP represents the amount of financial obligation to be repaid within a year; a higher percentage suggests a nation has high debt that needs to be paid off in the immediate future. In Thailand in 1996, short-term debt was 20.8%, significantly greater than the guideline for sustainable financial practices which recommends less than 10%.[29] High-risk borrowing was left unchecked by Thai financial leaders, contributing to short-term debt. Coupled with decreasing investor confidence, Thailand scrambled to pay short-term debt. The United States fails to provide data on short-term debt for 2007, however it followed similar patterns of unchecked borrowing.[30]

Another critical measure is the difference between current account deficit and foreign investments as a percentage of GDP. A negative figure indicated future financial turmoil as seen in Thailand who endured a measure of -5.6% in 1996.[31] The United States appeared much more stable at 3.1% in 2007, indicating low risk for financial instability. Thailand's alarming figure reveals that the government mismanaged investments, leading to a lack of stable long term funding to cover its deficits.[32] The United States displayed a positive number and attracted foreign investments from China and Japan in mortgage-backed securities (MBS), which made the economy appear healthier than it was.[33] This misleading claim inflated the value of assets, dragging down international investors.

The next indicator is foreign debt as a percentage of export earnings, which provides insight into the amount owed to foreign creditors relative to the income generated from exporting goods and services. In this case, both the United States and Thailand sit at 3.6% and 11.5%, respectively remaining below the threshold of 30%.[34] Thailand's robust exports in manufacturing were able to hold foreign debt as a percentage of export earnings to safe levels leading up to 1997. The United States, the world's largest economy, has an expansive export base, which is why its figure is much lower.[35]

Finally, the number of months a nation can finance with international reserves represents how long a country can sustain its import expenses using its available foreign currency reserves. Thailand's reserves were sufficient to cover seven months of imports in 1996.[36] The United States was only able to finance one month in 2007.[37] Any number of less than three months would suggest low sustainability in foreign markets. Here, the United States shows a warning sign that financial instability is impacting its presence in the global economy, thus prompting future problems. The United States is in a unique position to have the privilege of being able to trade worldwide in its own currency, so despite having low

28 Lipton, 2017
29 Park, 1998
30 International Monetary Fund, 2000
31 Park, 1998
32 International Monetary Fund "IMF Staff." Finance & Development. International Monetary Fund, June 1998. Accessed September 24, 2024. https://www.imf.org/external/pubs/ft/fandd/1998/06/imfstaff.htm.
33 Rajan, 2010
34 World Bank "Total Reserves (Includes Gold, Current US$)." World Bank. Accessed September 24, 2024. https://data.worldbank.org/indicator/FI.RES.TOTL.MO?locations=US.
35 Park, 1998
36 Park, 1998
37 World Bank, 2024

international reserves to finance only one month of imports, this is less alarming than it would be in other nations. Lack of oversight and manipulation in America led people to perceive this as much less of an issue. It still highlighted moderate financial strain and vulnerabilities in the market. In Thailand, rapid economic growth provided a buffer in international reserves at times, which gave investors a false sense of security in the economy, which was failing underneath.

Analyzing the Evidence in Historical Research

Corruption obscured financial instability in both Thailand and the United States. Thailand displayed four of six leading indicators of future financial instability yet its government ignored, concealed or downplayed the warning signs. The United States exhibited three of these indicators and had very little information and few guidelines for recognizing economic ruin in the era of development and globalization from Thailand. Similar to Thailand ten years prior, the United States made it impossible to address a crisis it refused to acknowledge, which allowed corruption to obscure economic vulnerabilities. The American economy had insufficient information regarding the warning signs of financial difficulties in Thailand, and the information it did possess was deliberately disregarded.

In Thailand prior to the 1997 Asian Financial Crisis, political mismanagement led to unsustainable lending practices and poor financial oversight. Current account deficits ballooned to unmanageable levels as a result of overreliance on foreign capital inflow. These warning signs were ignored by political and financial professionals, with continued confidence in the Thai Baht despite its obvious strain on the world economic scale. It would have been nearly impossible to spot signs of corruption in Thailand by examining the mismatched balance sheets leading up to 1997, as risky lending was encouraged and even required to maintain involvement in the financial game of Southeast Asia's developing economies. The Thai government and financial sector downplayed the risks of over-relying on short-term capital and liabilities, instead emphasizing that increased liabilities generated money through interest, a message that made the Asian Financial Crisis devastatingly sudden as opposed to gradual and preventable.

Thailand's corruption set the stage for a weak regulatory framework, where government officials and financial leaders prioritized short term gain over long term stability. By failing to address the surmounting economic instability associated with risky lending practices, political leaders remained complacent in manipulating finances and misreporting, which prevented many corrective measures from being implemented. The fragile Thai economy provided a playbook for the American financial system, detailing how to mishandle overspending and corruption.

Only a decade later, with the effects of the Asian Financial Crisis freshly in the mind of the world economy, America's financial institutions fell into a similar trap. The United States ran up impressive current account deficits, fueled by massive consumerism and spending. Just as Thailand's real estate sector endured a housing boom and stock market improvements, the American housing market, based on risky lending

and defaulted loans, created the illusion of an economy that was not only strong but gaining strength as well. Deliberate irresponsibility is what drove the United States further into debt. With no way for borrowers to repay the debt, these abusive lending practices impacted the balance of payments across the private and public sectors, resulting in the irresponsible allocation of financial resources. Spending to maintain the illusion of a booming housing market fostered a culture of cheating to survive in the volatile American economy. Mortgage-backed securities were the backbone of the illusion of stability, with funds pouring into investments that ultimately plummeted, dragging the economy down with them. These actions distorted the true economic picture, making it harder for anyone to predict the coming collapse. Prosperity in the United States and Thailand was the curated effect of the financial and government sectors because cronyism and lobbying manipulated financial instruments to conceal risk.[38]

The American economy had limited information from the warning signs of financial struggles in Thailand, and the information it did have was willfully ignored. United States policymakers and financial institutions concealed vital details from consumers, investors, and government bodies, perpetuating the corrupt culture that began with the 1997 Asian Financial Crisis. By denying and overlooking the few warning signs, they missed the opportunity to correct the looming financial crisis before reaching a worldwide crash.

In both nations, corruption clouded decision making. These fragile economies seemed strong, with rising real estate developments, booming stock markets, and increasing foreign investments; factors that were manipulated or exaggerated to maintain the illusion of prosperity.[39] Despite financial indicators signaling trouble, corruption and denial from government officials and businesses, fostered widespread complacency. In 2006, William Poole, Chair of the Federal Reserve in St. Louis, eased the minds of readers when he said, "There is no inherent reason that such changes [in adjustment of the current account] would lead to a financial market crisis,"[40] falsely easing the minds of nervous Americans. Financial elites in both countries ignored or concealed risks to protect their facade of an ever-growing economy. Ultimately, this mix of corruption and the failure to acknowledge mounting vulnerabilities directly led to the unforeseen crashes. Seemingly thriving economies were fragile, collapsing under dishonesty and corruption.

Alternative Approaches

Corruption played a significant role in the failures of Thailand and the United States during their recessions, but other economic and political factors, such as market failures and globalization, also played a major role. These factors suggest that corruption was more of a symptom of deeper economic issues rather than the root cause of the Asian Financial Crisis and the Global Financial Crisis.

38 Mishra, Prachi, et al. "Lobbying and the financial crisis." CEPR, 27 January 2010, https://cepr.org/voxeu/columns/lobbying-and-financial-crisis. Accessed 3 March 2025.

39 Carson, Michael, and John Clark. "Asian Financial Crisis." Federal Reserve History, https://www.federalreservehistory.org/essays/asian-financial-crisis. Accessed 3 March 2025.

40 Poole, William, "How Dangerous is the U.S. Current Account Deficit?" (2006). Center for Applied Economics. 33.
https://digitalcommons.lindenwood.edu/cee/33

One key factor was the impact of global capital flows. Thailand's economy relied heavily on foreign investment, while the United States, as a major player in global financial markets, had significant influence over international capital movement. This heavy dependence on global capital flows left both countries vulnerable. Globalization, the growing interconnectedness of financial markets, allowed for greater cross-border investments, making these economies more susceptible to instability. Both Thailand and the United States relied too heavily on these global trade and financial systems, accumulating substantial amounts of debt that ultimately worsened their economic crises.[41]

Market failures, particularly in the housing market, also contributed to the crises. In both nations, overvalued housing assets, irrational behavior from both consumers and investors, and technological advancements led to rapid housing market expansion. This expansion resulted in a massive oversupply of housing and insufficient demand. Real estate prices soared without enough consumers willing or able to pay the inflated prices, creating a bubble. In both Thailand and the United States, the housing market grew out of sync with actual consumer demand, leading to a situation where housing prices far exceeded the true market value.[42]

While alternative explanations certainly played a role, the mechanisms of corruption in both cases demonstrate how deeply embedded practices of mismanagement and greed within financial and political systems were at the heart of these crises.

Conclusion

Both the Asian Financial Crisis and the Global Financial Crisis exposed deep-rooted systemic corruption. Thailand's economy suffered from cronyism and a lack of oversight, and its recovery involved multifaceted reforms aimed at preventing future corruption. The International Monetary Fund provided a $17.2 billion bailout, which was contingent upon structural and financial changes. By the early 2000s, Thailand had largely recovered, with improved GDP growth, reduced inflation, and a restructured financial sector designed to curb corrupt practices.[43]

In the United States, recovery focused on addressing the deception and misleading information that led to financial overexposure, marked by high current account deficits and public debt. Programs like the Troubled Asset Relief Program (TARP) helped stabilize banks, while the Dodd-Frank Act implemented stricter regulations to prevent future crises. It imposes more stringent standards, including tougher requirements for capital, leverage, risk management, mergers and acquisitions, and stress testing, on banks and other financial firms whose failure could threaten the stability of the American financial system.[44]

Both crises offer valuable lessons for predicting future economic downturns in the United States and Thailand, where cor-

41 Liang, Yan. 2012. "Global Imbalances and Financial Crisis: Financial Globalization as a Common Cause." Journal of Economic Issues 46 (2): 353–62. doi:10.2753/JEI0021-3624460210.

42 Arestis, Phillip, Santonu Basu and Sushanta Mallick. 2005. "Financial Globalization: The Need for a Single Currency and a Global Central Bank." Journal of Post Keynesian Economics 27 (3): 507–31. doi:10.1080/01603477.2005.11051444.

43 International Monetary Fund, 2000

44 U.S. Congress. Dodd-Frank Wall Street Reform and Consumer Protection Act. H.R. 4173. 111th Congress (2010). https://www.congress.gov/bill/111th-congress/house-bill/4173/text.

ruption remains a pressing issue. Without addressing these vulnerabilities, financial instability will continue to threaten the global economy. This research and analysis highlights the underlying issues of undetected corruption, where complacency and deception fueled financial collapse. To prevent future crises, transparency, accountability, and robust regulatory frameworks are essential. It is vital to prioritize reform and encourage vigilance in these economic systems to build a more alert financial system that safeguards against corruption.

***

for acknowledgments and appendixes
please visit www.tulanejournal.com

***

COVID-19 and Education Inequality: How Demographics and School Closures Affected Learning

Nathaniel Nelson
University of Minnesota - Twin Cities

Abstract

In this paper, I explore how racial and socioeconomic test score gaps have changed since the onset of the COVID-19 pandemic and to what extent teaching modality was responsible for the growth in score gaps. Additionally, I analyze how test score gaps evolved from 2022 to 2023. This paper addresses the differences between individual-level score gaps and district-level gaps. Given that past literature finds growth in test score gaps due to the COVID-19 pandemic, I attempt to identify whether high-minority and high-poverty school districts are falling behind as a whole or if the widening in test score gaps is due to minority and economically disadvantaged students falling behind peers in their district. I use a triple-difference (DDD) model to determine whether in-person learning affected the growth of test score gaps and to identify interactions between district- and individual-level effects. I find that (1) From 2019 to 2022, the test score gap between high-poverty and low-poverty school districts widened significantly, as did the gap between high-minority and low-minority districts; (2) Districts that spent more of the 2020–2021 school year in person saw smaller drops in test scores, particularly in math; (3) In-person learning reduced score gaps between high- and low-poverty school districts in math but may have widened Black-white score gaps in reading; (4) Within districts, economically disadvantaged students have seen slower recovery from 2022 to 2023 than their classmates; and (5) All significant COVID-related score gaps that appeared between 2019 and 2022 persisted in 2023.

Introduction

Students have been falling behind since the start of the pandemic. A January 2023 meta-analysis of studies from 15 countries estimated that the average student experienced learning loss equivalent to 35% of a school year and that learning deficits have been relatively stable since the early stages of the pandemic.[1] Kuhfeld, Soland, and Lewis[2] find that learning loss in math continued throughout the first two years

1 Bastian A. Betthäuser, Gil Hen, and Noam Angrist, "A Systematic Review and Meta-Analysis of the Evidence on Learning During the COVID-19 Pandemic," Nature Human Behaviour 7, no. 1 (2023): 375-385, https://doi.org/10.1038/s41562-022-01506-4.

2 Megan Kuhfeld, James Soland, and Karyn Lewis, "Test Score Patterns Across Three COVID-19-Impacted School Years," Educational Researcher 51, no. 7 (October 2022): 500–506, https://doi.org/10.3102/0013189X221113022.

of the pandemic, whereas learning loss in reading occurred mostly between fall 2020 and fall 2021.

Learning loss continued into the 2022–2023 school year. A report by the NWEA Center for School and Student Progress finds that achievement gains for the 2022–2023 school year were smaller than in pre-pandemic years in all but the youngest grade tested (3rd grade) across all race/ethnicity groups. At the end of the school year, the average student was 4.1 months behind in reading and 4.5 months behind in math. Growth in reading achievement for grades 6-8 has been especially slow, lagging behind pre-COVID gains by nearly 20%.[3]

My analysis shows that differences between districts drove the widening of test score gaps from 2019 to 2022 and that differences within districts are a major driver of gaps in 2023. My findings add to the literature by (1) presenting a triple differences model for test score gaps through 2023; (2) studying interactions between modality and district demographics; (3) including COVID-19 death rates in a model of test score gaps; and (4) examining the interaction of socioeconomic and racial subgroups with overall district composition to gain insight into the mechanisms driving test score gaps.

I use data on test scores, school modality, and demographic information reported at the local educational agency (LEA) level. Each school district is an LEA, but other educational agencies such as charter schools also comprise their own LEA.[4] Data on COVID cases are available at the county level and are linked to score, modality, and demographic data, as described in Section 2.2. Using these data, I estimate the effect of the pandemic on test scores. I use a difference-in-differences model to determine how the test scores of different subgroups changed from 2019 to 2023. I estimate the effect of modality on test scores by comparing the change in scores for districts with a high percentage of the 2020–2021 school year spent in-person to districts with less in-person instruction. I interact modality with race and socioeconomic status to estimate the effect of modality on particular group differences. A causal interpretation of these comparisons is valid under the assumption of parallel trends—that score differences between districts and groups would have been stable if not for the COVID-19 pandemic. As there is non-negligible deviation from parallel trends for some of these tests, I utilize additional techniques described in section 3.1.

Before the COVID-19 pandemic, research on learning loss focused on summer breaks. Atteberry and McEachin[5] estimated that students lose between 17% and 28% of the previous school year's learning in reading/language arts (RLA) and between 25% and 34% in math during summer break.

There is also evidence to suggest that summer learning loss leads to a growing score gap. Hayes and Grether[6] examined test

3 Karyn Lewis and Megan Kuhfeld, Education's Long COVID: 2022-23 Achievement Data Reveal Stalled Progress Toward Pandemic Recovery, ED630208 (Portland, OR: Center for School and Student Progress at NWEA, July 2023), 12, https://eric.ed.gov/?id=ED630208.

4 Throughout this paper I will refer to local education agencies as "districts"particularly in discussions of "district- level effects".

5 Allison Atteberry and Andrew McEachin, "School's Out: The Role of Summers in Understanding Achievement Disparities," American Educational Research Journal 58, no. 2 (July 8, 2020): 239-282. https://doi.org/10.3102/0002831220937285.

6 Donald P. Hayes and Judith Grether, "The School

scores of students in New York City and concluded that upward of 80% of the difference in progress between students in the wealthiest and whitest schools and students in the poorest, less white schools was associated with summer months when children were not in school. A more recent study of test score data from Baltimore showed that the socioeconomic test score gap did not grow throughout the school year but grew significantly over the summer, suggesting that the growth in score gaps from 1st to 9th grade is attributable to unequal summer learning loss.[7] Other studies, however, have estimated that approximately 4% of summer learning loss is explained by race and socioeconomic status (SES), suggesting that there are other more important factors.[8] Regardless, this research suggests that test score gaps grow when students are not attending school.

In line with this finding, experts expected test score gaps to increase during the COVID-19 pandemic, when children were not attending school in person. In a November 2020 survey, education researchers predicted that socioeconomic gaps would grow by 0.3 standard deviations for math and 0.25 standard deviations for reading by 2021. This is approximately equivalent to a relative achievement loss of 5 months in math and 6 months in reading.[9] Analysis in June 2020 estimated that the average student could fall behind by seven months during the pandemic. Projected learning losses were 10.3 months for Black students, 9.2 months for Hispanic students, and over a year for low-income students. They estimated that white students could see a 1.6 percent reduction in lifetime earnings due to this learning loss; for Black and Hispanic students, the reduction in earnings could be more than 3 percent. Furthermore, they estimate a drop in GDP of $173 billion to $271 billion per year as a result.[10]

One of the most important education policy issues in 2020 and 2021 was determining when schools should return to fully in-person learning. During the height of the pandemic, experts suggested that in-person school would be better for learning than remote or hybrid education. Organizations such as the American Academy of Pediatrics recommended a safe return to in-person learning as early as the summer of 2020.[11] The US Department of Education claimed that in-person education improves academic outcomes and recommended returning to in-person schooling as public health would allow. To facilitate this, the American Rescue Plan Act of 2021 provided almost $122 billion to schools through Elementary and Secondary School Emergency Relief (ESSER) funds, which could be used to improve the safety of in-person instruction.[12]

Year and Vacations: When Do Students Learn," (Paper presented at the Eastern Sociological Association Convention, New York, NY, April 19, 1969): 1–22, https://eric.ed.gov/?id=ED037322.

7 Karl L. Alexander, Doris R. Entwisle, and Linda Steffel Olson, "Lasting Consequences of the Summer Learning Gap," American Sociological Review 72, no. 2 (April 2007): 167–180, https://doi.org/10.1177/000312240707200202.

8 Atteberry and McEachin, "School's Out: The Role of Summers in Understanding Achievement Disparities."

9 Drew H. Bailey et al., "Achievement Gaps in the Wake of COVID-19," Educational Researcher 50, no. 5 (June 1, 2021): 266–275, https://doi.org/10.3102/0013189X211011237.

10 Emma Dorn et al., "COVID-19 and learning loss—disparities grow and students need help," McKinsey & Company, December 8, 2020, https://www.mckinsey.com/industries/public- sector/our-insights/ covid-19-and-learning-loss-disparities-grow-and-students-need-help.

11 American Academy of Pediatrics, "COVID-19 Guidance for Safe Schools," COVID-19 Interim Guidance, September 8, 2022, https://www.aap.org/en/pages/2019-novel-coronavirus-covid-19-infections/clinical-guidance/covid-19-planning-considerations-return-to-in-person-education-in-schools/.

12 Department of Education, Supporting Students During the COVID-19 Pandemic: Maximizing In-Person Learning and Implementing Effective Practices for Students in Quarantine and Isolation (Washington, DC: Department of Education, September, 2021), https://www.ed.gov/coronavirus/

Ex post studies suggest that the effects of pandemic learning loss have been unequal. Schools with higher poverty rates are further behind their counterparts;[13][14] this gap is more pronounced in math scores than in reading scores.[15] Learning loss in both reading and math was worse in high-poverty school districts, with the disparity mainly increasing in the 2020–2021 academic year.[16] Furthermore, there is evidence that younger students who were already performing worse on tests before the pandemic have experienced greater learning loss than high-achieving students.[17] There is already a trend of increasing inequality in education. Owens, Reardon, and Jencks[18] noted an increasing socioeconomic score gap in the US, finding that "between-district income segregation of families with children enrolled in public school increased by over 15% from 1990 to 2010."

It is also well-established that in-person instruction led to better educational outcomes,[19][20] particularly for younger students.[21][22] Although the test score data I use are not separated by grade, it is important to consider that policies affect children differently depending on their age—a consideration that is unfortunately obscured in my analysis. Despite the initial differences by teaching modality, it does not appear that school districts that spent more time in person experienced better recovery after the initial drop in test scores. Halloran et al.[23] examine "recovery rates," defined as "the share of the 2019–2021 test score decline that is recovered by Spring 2022." They find that recovery rates varied widely by state and district, but not by the rate of in-person schooling in the 2020–2021 school year. This suggests that the effect of teaching modality may have been limited to the 2020–2021 school year, when learning formats varied. See Figures C1, C2, C3, C4. This analysis explores the change in scores between 2019 and 2022, which captures both the initial drop and the initial recovery but does not enable me to disentangle the unique effects of each period. In addition, I examine the recovery from 2022 to 2023. I find evidence for the modality-related test score gap that Halloran and colleagues attribute to the 2020–

supporting-students- during-covid-19-pandemic.

13 Erin Fahle et al., School District and Community Factors Associated with Learning Loss during the COVID-19 Pandemic, Stanford Education Data Archive, (Cambridge, MA: Harvard University, May 2023), 1-64, https://cepr.harvard.edu/sites/hwpi.harvard.edu/files/cepr/files/explaining_covid_losses_5.23.pdf.

14 Betthäuser, Bach-Mortensen, and Engzell, "A Systematic Review and Meta-Analysis of the Evidence on Learning During the COVID-19 Pandemic."

15 Dan Goldhaber et al., "A Comprehensive Picture of Achievement Across the COVID-19 Pandemic Years: Examining Variation in Test Levels and Growth Across Districts, Schools, Grades, and Students" (Working Paper No. 266-0522, CALDER, 2022), 1-95, https://caldercenter.org/sites/default/files/2024-11/CALDER%20Working%20Paper%20266-0522_0.pdf

16 Kuhfeld, Soland, and Lewis, "Test Score Patterns Across Three COVID-19-Impacted School Years."

17 Véronique Irwin et al., Report on the Condition of Education 2023, NCES 2023-144, (Washington, DC: National Center for Education Statistics, May 2023) 1-44, https://nces.ed.gov/pubs2023/2023144.pdf.

18 Ann Owens, Sean F. Reardon, and Christopher Jencks, "Income Segregation Between Schools and School Districts," American Educational Research Journal 53, no. 4 (August, 2016): 1159–1197, https://doi.org/10.3102/0002831216652722.

19 Martin R. West, How Much Have Students Missed Academically Because of the Pandemic? A Review of the Evidence to Date, ED614130, (Bothell, WA: Center on Reinventing Public Education, July 2021), 1-18, https://files.eric.ed.gov/fulltext/ED614130.pdf.

20 E. M. Fahle et al., School District and Community Factors Associated With Learning Loss During the COVID-19 Pandemic.

21 Rebecca Jack et al., "Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts," American Economic Review: Insights 5, no. 2 (June 1, 2023): 173–190, https://doi.org/10.1257/aeri.20210748

22 Svenja Hammerstein et al., "Effects of COVID-19-Related School Closures on Student Achievement-A Systematic Review," Frontiers in Psychology 12 (September 16, 2021), https://doi.org/10.3389/fpsyg.2021.746289.

23 Clare Halloran et al., Post COVID-19 Test Score Recovery: Initial Evidence from State Testing Data, w31113, (Cambridge, MA: National Bureau of Economic Research, April 2023), https://doi.org/10.3386/w31113.

2021 school year and show that it had not closed by spring 2023.

Regarding the effects of teaching modality on score gaps, E. M. Fahle et al.[24] found no significant differences between the effect of in-person learning for different racial and socioeconomic subgroups and could not identify the mechanism by which high-poverty districts performed worse after the pandemic. However, the aforementioned study and Goldhaber et al.[25] concluded that remote and hybrid instruction were significantly more harmful in mid- and high-poverty schools and school districts. Thus, it seems that modality had an effect at the district-level and/or the school-level but, within each school, there is no evidence that students from different backgrounds were impacted significantly differently. Despite this, test score gaps within districts have widened from 2022 to 2023. E. Fahle et al.[26] report that high-income students are recovering roughly twice as fast as low-income students. My findings support these conclusions.

Data: Data Sources

Test score data were retrieved from the Stanford Education Data Archive (SEDA). These data sets are constructed using the United States Department of Education EDFacts data system. I used administrative district-level data from SEDA2023. The data set is extensive, as states are required by federal law to administer standardized tests in math and reading/language arts (RLA) to every student in grades 3 to 8 and report aggregated test score data to EDFacts. Data have been suppressed for several reasons including low participation rate, incomplete data by subgroup, alternate assessments, scores only falling in the top or bottom proficiency category, or cells failing to meet minimum statistical estimation requirements. Data are also suppressed if an estimate is based on too few observations or is too imprecise. In particular, many subgroup estimates are removed at the district level. Because I use these subgroup estimates, my data set is not nationally representative.

I use score estimates from the Year Standardized (YS) scale. These estimates represent standardized averages across all students in grades 3-8. Per the SEDA 2023 documentation: "we standardize the estimates to the 2019 national average in each grade and subject. In this scale, each unit is equivalent to a 2019 national standard deviation in the same subject and grade." I use the Empirical Bayes (EB) estimates reported by SEDA. Although these estimates account for differences in testing practices between states, it is still important to note that tests vary in both content and timing.

SEDA also provides socioeconomic, demographic, and segregation data from the American Community Survey (ACS), the Common Core of Data (CCD), and the Civil Rights Data Collection (CRDC). An in-depth description of the data collection and cleaning can be found in the documentation for SEDA 5.1[27] and SEDA

24 E. M. Fahle et al., School District and Community Factors Associated With Learning Loss During the COVID-19 Pandemic.

25 Dan Goldhaber et al., "The Consequences of Remote and Hybrid Instruction During the Pandemic." (Working Paper No.30010, NBER, 2022), http://www.nber.org/papers/w3001.

26 Fahle, Erin, Thomas J. Kane, Sean Reardon and Douglas Staiger, The First Year of Pandemic Recovery: A District-Level Analysis, Center for Education Policy Research Brief, (Cambridge, MA: Harvard University, January 2024), https://educationrecoveryscorecard.org/wp-content/uploads/2024/01/ERS-Report-Final-1.31.pdf.

27 Erin Fahale et al., Stanford Education Data Archive

2023.[28]

Data on COVID-19 deaths are from the Johns Hopkins Coronavirus Resource Center.[29] Total deaths from COVID-19 are estimated for each county and identified by FIPS code. Estimates were reported daily from January 22, 2020 to March 3, 2023. Data were gathered from many sources such as the World Health Organization, the European Centre for Disease Prevention and Control, and data published by individual states and counties. Population estimates are also reported for each county.

I retrieved data on school modality from HealthData.gov, a website of the United States Department of Health and Human Services. I use their definitions for different modalities:

"In-Person: All schools within the district offer face-to-face instruction 5 days per week to all students at all available grade levels."

“Remote: Schools within the district do not offer face-to-face instruction; all learning is conducted online/remotely to all students at all available grade levels."

“Hybrid: Schools within the district offer a combination of in-person and remote learning; face-to-face instruction is offered less than 5 days per week, or only to a subset of students.”[30]

Data are reported weekly, with each school district classified as in-person, remote, or hybrid. Figure C1 shows the percentage of school districts classified in each modality throughout the 2020–2021 school year. Figures C2, C3, C4 show the average share of the school year that each state spent in each modality.

Data: Data Cleaning

I use different data sets to examine the Black-white test score gap, the socioeconomic status (SES) test score gap, and general trends for all students. Due to data cleaning and suppression by SEDA, many relevant estimates are missing; to maximize the number of observations included in my sample, I use different restrictions as appropriate. Six different data sets are used in my analysis, which are referred to as Data Sets A - F. Each of these data sets is a subset of all available data. A description of each group is reported in Appendix A and sample composition is shown in Tables A1-6. For each sample, I filter out any school district missing score estimates for the relevant subgroup(s) in 2016, 2017, 2018, 2019, or 2022 so that there is a balanced panel of school districts across years. I repeat this analysis including 2023 scores. Fewer school districts have data reported in 2023, so this analysis is based on a smaller sample, as seen in the sample composition tables in the appendix.

Districts that saw large changes in demographics or enrollment between 2019 and

Technical Documentation: Version 5.0, Stanford Education Data Archive, (Cambridge, MA: Harvard University, 2024), https://purl.stanford.edu/cs829jn7849.

28 Erin Fahale et al., Stanford Education Data Archive Technical Documentation SEDA2023, Stanford Education Data Archive, (Cambridge, MA: Harvard University, January 2024), https://edopportunity.org/docs/seda2023_documentation_20240130.pdf

29 Ensheng Dong et al., “The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: Data Collection Process, Challenges Faced, and Lessons Learned,” The Lancet Infectious Diseases 22, no. 12 (December, 2022): 1474-4457, https://doi.org/10.1016/ S1473-3099(22)00434-0.

30 Centers for Disease Control and Prevention (School Learning Modalities, 2020-2021; accessed March 17, 2024) https://healthdata.gov/National/School-Learning-Modalities-2020-2021/a8v3-a3m3/about_data

2022 are flagged in the SEDA 2023 data set. I chose to keep these districts in the sample, which could be a source of bias in my results.

To estimate COVID severity, I calculated the average death rate between September 1, 2020 and June 30, 2021 to capture the whole school year. The rate is:

$$\frac{(deaths_{6.30.21} - deaths_{9.1.20}) \times 1000}{10 \times Population}$$

where $deaths_d$ is the sum of deaths in an area from the onset of the pandemic to date d. It is therefore an average monthly death rate per 1000 people. I use the z-score of this estimate for easier interpretation.

COVID-19 deaths were reported at the county level, but many school districts are not contained within one county. I used data from the National Center for Education Statistics (NCES) to select one county for each local educational agency.[31] For every school district within multiple counties, I selected the county containing the plurality of the school district's land area.

I used weekly modality data from HealthData.gov to calculate the percentage of the 2020–2021 school year that each school district spent fully in-person. Data are reported from September 6, 2020 to May 30, 2021. The minimum number of school days required by any US state is 160.[32] Accordingly, I removed data for school districts that reported modality data for fewer than 32 weeks. This ensures that the averages reflect the majority of the school year and avoids bias from temporal differences in modality. I calculated the percentage of the school year that districts were reported as in-person, hybrid, or remote, respectively.

Methods: All Students

I use difference-in-differences analysis to explore changes in trends from 2019 to 2022 and 2023. Using 2019 as a base year, I explore how districts diverged from their 2019 scores and how they recovered from 2022 to 2023. I include district-level fixed effects in all regressions to account for the preexisting differences in scores between districts and state-year dummy variables to account for states' different policy responses to the pandemic, beyond school closures. I also weight my estimates by the number of students in each district taking each test (math or reading).

Traditional difference-in-differences analysis relies on the assumption of parallel trends; a "treated" and "untreated" group must have had the same trends prior to an intervention. This does not hold for all groups that I analyze. I use methods introduced by Rambachan and Roth[33] to check for robustness to these differential trends. Essentially, I test the significance of my estimates assuming that the differential trends continued into the post-treatment period and were amplified by up to M times the maximum pre-treatment deviation from parallel trends. As stated by

31 Douglas E. Geverdt, Education Demographic and Geographic Estimates Program (EDGE): School District Geographic Relationship Files User's Manual, NCES 2018-076 (Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2019). http://nces.ed.gov/pubsearch.

32 Gerardo Silva-Padron and Meghan McCann, "50-State Comparison: Instructional Time Policies," Education Commission of the States, February 6, 2023, https://www.ecs.org/50-state-comparison-instructional-time-policies-2023/.

33 Ashesh Rambachan and Jonathan Roth, "A More Credible Approach to Parallel Trends," Review of Economic Studies 90, no. 5 (September 5, 2023): 2555–2591, doi:10.1093/restud/rdad018

Rambachan and Roth:

$$\Delta^{RM} = \{\delta : \forall t \geq 0, |\delta_{t+1} - \delta_t| \leq M \cdot \max_{t<0} |\delta_{t+1} - \delta_t|$$

Here δ is a vector of differential trends for the pre- and post-period. Note that this is complicated by the fact that data are not available for 2020 and 2021; for t=2019, I treat t+1=2022. In regression tables throughout, I indicate results robust to a deviation from parallel trends defined by M=2 in bold.

I begin by verifying findings mentioned in the introduction, using Data Set A. I use a difference-in-differences model to explore the change in test score gaps from 2019 to 2022:[34]

$$score_{d,s,t} = \beta_1 year22_t + \beta_2 year22_t \times pctecd_d + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,s,t}$$

$$score_{d,s,t} = \beta_1 year22_t + \beta_2 year22_t \times pctblk_d + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,s,t}$$

$$score_{d,s,t} = \beta_1 year22_t + \beta_2 year22_t \times pcthsp_d + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,s,t}$$

where $score_{d,s,t}$ is the average standardized score in district d in year t in state s, $year22_t$ is a dummy variable that equals 1 for observations from 2022, and $pctecd_d$, $pctblk_d$, $pcthsp_d$ represent the percent of students in district d who are economically disadvantaged, Black, or Hispanic, respectively. α_d are district fixed effects.

In these equations, a significant negative value of β_2 is evidence that districts with a higher percentage of students within the relevant subgroup are performing worse in 2022 than in 2019.

34 Throughout this section, I write equations using "score" as the dependent variable. Since I am examining both mathematics and reading/language arts scores, each equation represents two regressions: one using math scores as the dependent variable and the other using reading/language arts scores. To avoid redundancy, I provide general equations. Separate regressions are shown in the results section.

$$score_{d,s,t} = \beta_1 year22_t + \beta_2 pctecd_d + \beta_3 inperson_d + \beta_4 year22_t \times pctecd_d + \beta_5 year22_t \times inperson_d + \beta_6 pctecd_d \times inperson_d + \beta_7 year22_t \times pctecd_d \times inperson_d + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,s,t}$$

I use a triple difference (DDD) estimator as introduced by Gruber (1994)[35] to estimate the effect of in-person learning on the change in district test score gaps.

where variables are defined as above and $inperson_d$ takes values in [0, 1] and represents the percentage of the 2020–2021 school year that district d was fully in-person.

Here β_7 is a DDD estimator representing the difference in the change in test score gaps between districts based on whether schools were in person. A positive value would indicate that economically disadvantaged school districts that spent more of the 2020–2021 school year in person lost less ground than similar districts that spent more time in remote or hybrid learning.

$$score_{d,s,t} = \beta_1 year_t + \beta_2 pctecd_d + \beta_3 year_t \times pctecd_d + \alpha_s + \epsilon_{d,s,t}$$

$$score_{d,s,t} = \beta_1 year_t + \beta_2 pctecd_d + \beta_3 inperson_d + \beta_4 year_t \times pctecd_d + \beta_5 year_t \times inperson_d + \beta_6 pctecd_d \times inperson_d + \beta_7 year22_t \times pctecd_d \times inperson_d + \alpha_s + \epsilon_{d,s,t}$$

I use similar models on Data Set B to explore trends through Spring 2023:
Here $year_t$ has 3 levels: 2019 (base year), 2022, and 2023. Regression output shows different coefficients for 2022, 2023, and interactions with both years. I simplify the remainder of this section by writing one equation to represent models including and excluding 2023 data as the equations are analogous.

Methods: Subgroup Analysis

35 Gruber, Jonathan, "The Incidence of Mandated Maternity Benefits," The American Economic Review 84, no. 3 (1994): 622–41, http://www.jstor.org/stable/2118071.

$$score_{d,g,s,t} = \beta_1 year_t + \beta_2 grp_g + \beta_3 year_t \times grp_g + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,g,s,t}$$

Data from SEDA include average scores for racial and socioeconomic subgroups. I use these averages to examine score gaps at the subgroup level. I explore several different subgroups and provide similar analysis for each. First, I examine the growth in score gaps from 2019 to 2022 of the form

$$score_{d,g,s,t} = \beta_1 year_t + \beta_2 grp_g + \beta_3 pcteed_d + \beta_4 year_t \times grp_g + \beta_5 year_t \times pcteed_d + \beta_6 grp_g \times pcteed_d + \beta_7 year_t \times grp_g \times pcteed_d + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,g,s,t}$$

and the analogous equations for Black and Hispanic percentages where scored,g,s,t is the average standardized score in district i in year t for subgroup g and grpg is a dummy variable equal to 1 for observations in the subgroup of interest (g). Thus, β_3 is an estimate of the change in a test score gap from 2019 to 2022. I fit the model:

$$score_{d,g,s,t} = \beta_1 year_t + \beta_2 grp_g + \beta_3 inperson_d + \beta_4 year_t \times grp_g + \beta_5 year_t \times inperson_d + \beta_6 grp_g \times inperson_d + \beta_7 year_t \times grp_g \times inperson_d + \delta state_s \times year22_t + \alpha_d + \epsilon_{d,g,s,t}$$

As above, β_7 is a DDD estimator representing the difference in the change in test score gaps based on the percentage of the year schools were in person.

I also provide a fully interacted model including average COVID deaths within each district and a fully interacted model including the difference between the average district-wide score in 2019 and the average district-wide score in 2018 (separated by subgroup). I refer to the latter as pre-trends because I use it to test whether the trends before the pandemic can explain the difference in trends during and after the school closures. These models are included to check the robustness of the model in equation (8).

Finally, I explore a DDD model with year, subgroup, modality, and percent of district economically disadvantaged. This model can provide insight into the mechanisms behind test score gaps. If economically disadvantaged districts and students are performing worse, optimal policy response depends on whether the trend is occurring mostly at the district level or the individual level. The regression equation is below:

A negative coefficient on β_4 suggests that the effect is on an individual level—regardless of the average socioeconomic status of the school district, children in that subgroup have been performing worse after the pandemic. A negative coefficient on β_6 suggests that the effect is larger at the district level—disadvantaged districts as a whole have been performing worse, which can explain the differences between subgroup means. The interaction term β_7 can provide insight into how disadvantaged students are performing in disadvantaged districts.

Results: District Level Effects

Consistent with E. M. Fahle et al.,[36] I find that school districts with a higher share of economically disadvantaged students saw significantly larger learning loss. Furthermore, this newly created gap has not closed from 2022 to 2023. The basic model (see Table 1 columns (1) and (2)) shows no evidence that school districts with few economically disadvantaged students experienced any learning loss in reading. There is evidence that math scores fell regardless

36 E. M. Fahle et al., School District and Community Factors Associated With Learning Loss During the COVID-19 Pandemic.

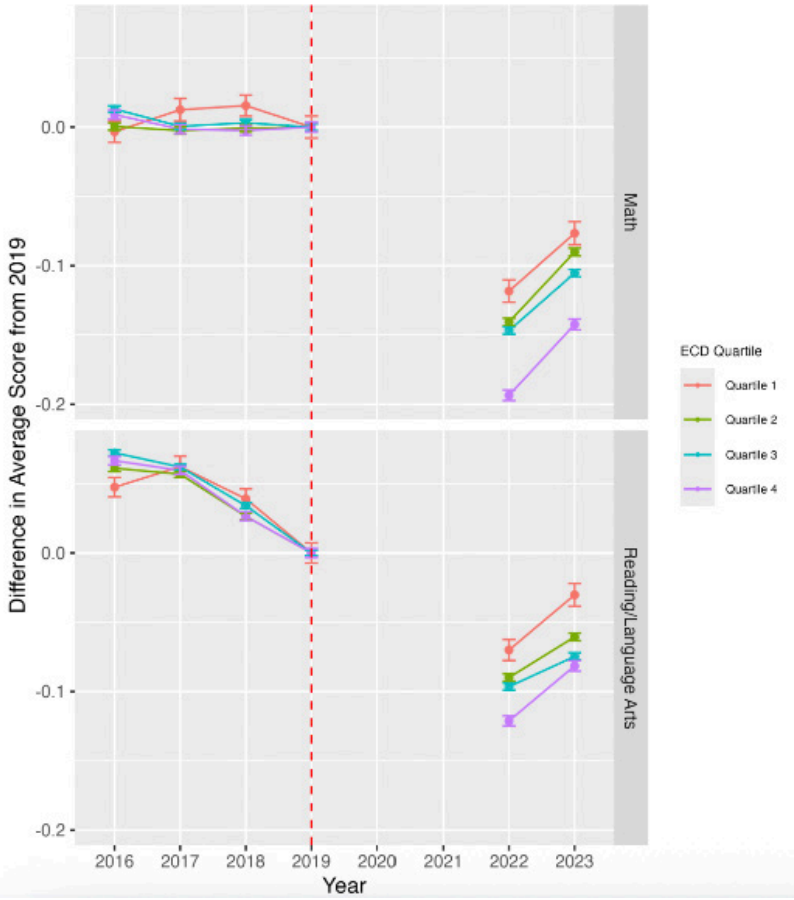


Figure 1: Average scores by subject and percent of students in district classified as economically disadvantaged (Quartile) - data set B

of district demographics, but this impact appears to be short-term, with near full recovery by 2023. However, for school districts with a high share of economically disadvantaged students, the drop in test scores was much more extreme in both math and reading. This gap is persistent; while less disadvantaged school districts have recovered, economically disadvantaged school districts are still far behind pre-COVID achievement levels. It appears that the gap in math scores has shrunk slightly, while the gap in reading scores has not.

Figure 1 shows the difference in average scores from the 2019 level for 2016 through 2023. The data have been split into quartiles based on the percentage of the student body that is economically disadvantaged. The plot shows that all quartiles had similar trends prior to the onset of the pandemic, but districts with more economically disadvantaged students saw a greater drop in scores from 2019 to 2022. Although all quartiles show evidence of recovery from 2022 to 2023, the new gap between low-poverty and high-poverty schools did not close. Note that this plot obscures the score gap that existed before the pandemic.

Columns (3) and (4) in Table 1 show a regression of modality on test scores. I find evidence that schools districts that spent more of the 2020–2021 school year in person had higher average math and reading scores in both 2022 and 2023. The overall effect of in-person learning was larger for math scores than reading scores. Although,

	Dependent variable:					
	Math	RLA	Math	RLA	Math	RLA
	(1)	(2)	(3)	(4)	(5)	(6)
2022	−0.029***	−0.001	−0.179***	−0.097***	−0.064***	−0.018***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)
2023	−0.008	0.012**	−0.133***	−0.084***	−0.034***	−0.004
	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)
2022 × % ECD	**−0.164*****	**−0.117*****			**−0.173*****	−0.112***
	(0.007)	(0.006)			(0.008)	(0.007)
2023 × % ECD	**−0.142*****	**−0.121*****			**−0.150*****	−0.111***
	(0.007)	(0.006)			(0.008)	(0.007)
2022 × In Person			**0.079*****	0.036***	0.028***	0.029***
			(0.005)	(0.005)	(0.011)	(0.010)
2023 × In Person			**0.061*****	0.032***	0.017	0.034***
			(0.005)	(0.005)	(0.011)	(0.010)
2022 × % ECD × In Person					0.075***	−0.006
					(0.018)	(0.016)
2023 × % ECD × In Person					0.064***	−0.024
					(0.018)	(0.017)
Observations	10,302	10,302	10,302	10,302	10,302	10,302
R^2	0.522	0.481	0.519	0.467	0.535	0.485
Adjusted R^2	0.278	0.217	0.273	0.195	0.297	0.222
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Note:					*p<0.1; **p<0.05; ***p<0.01	

Table 1: DD regression of score on percent of district ECD (2023 included) - base year 2019

as seen in Figure C6, trends leading up to the pandemic are not parallel, the finding is robust to this difference in parallel trends in math scores, using methods of Rambachan and Roth. Columns (5) and (6) show district economic demographics interacted with the percentage of the 2020–2021 school year a district spent in person. This interaction reveals that in-person learning had a unique positive effect for economically disadvantaged school districts' math scores, but I find no unique effect on reading scores. The difference in the drop in math scores between low- and high-poverty school districts is larger for schools that spent less time in person. Figure C7 shows trends in scores by the share of a school district that is economically disadvantaged. Separate graphs are provided to show trends in schools that were never in person in the 2020–2021 school year and fully in-person schools. There are differential trends for schools that would be fully in-person, so this estimate is not robust with respect to the methods described above.

Table 2 shows similar regressions but examines the racial demographics of schools (and excludes the regression focused solely on modality).

Columns (1) and (2) show that districts with a higher percentage of Black students saw larger drops in both math and reading scores than districts with fewer Black students. The gap has shrunk from 2022 to 2023, but as of 2023, there was still a significantly larger gap than in 2019. Similarly, columns (3) and (4) show that

	Dependent variable:							
	Math	RLA	Math	RLA	Math	RLA	Math	RLA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2022	−0.054***	−0.032***	−0.107***	−0.064***	−0.087***	−0.045***	−0.147***	−0.082***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)
2023	−0.039***	−0.029***	−0.071***	−0.052***	−0.066***	−0.042***	−0.100***	−0.067***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)
2022 × % Black	**−0.227*****	**−0.126*****			**−0.202*****	**−0.104*****		
	(0.008)	(0.008)			(0.010)	(0.009)		
2023 × % Black	**−0.170*****	**−0.105*****			**−0.139*****	**−0.072*****		
	(0.008)	(0.008)			(0.010)	(0.010)		
2022 × % Hispanic			**0.173*****	0.081***			**0.162*****	0.074***
			(0.010)	(0.009)			(0.012)	(0.011)
2023 × % Hispanic			**−0.176*****	−0.092***			**−0.164*****	−0.075***
			(0.010)	(0.009)			(0.012)	(0.011)
2022 × In Person					0.050***	0.026***	**0.061*****	0.029***
					(0.006)	(0.006)	(0.006)	(0.006)
2023 × In Person					0.046***	0.032***	**0.047*****	0.032***
					(0.006)	(0.006)	(0.006)	(0.006)
2022 × % Black × In Person					−0.015	−0.055***		
					(0.021)	(0.019)		
2023 × % Black × In Person					−0.052**	−0.098***		
					(0.021)	(0.020)		
2022 × % Hispanic × In Person							0.066*	0.013
							(0.036)	(0.033)
2023 × % Hispanic × In Person							0.019	−0.061*
							(0.036)	(0.033)
Observations	10,302	10,302	10,302	10,302	10,302	10,302	10,302	10,302
R^2	0.527	0.472	0.522	0.468	0.533	0.475	0.532	0.472
Adjusted R^2	0.285	0.203	0.278	0.197	0.295	0.207	0.293	0.202
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Note:							*p<0.1; **p<0.05; ***p<0.01	

Table 2: DD regression of score on percent of district black and hispanic on (2023 included) - base year 2019

districts with a higher percentage of Hispanic students have also seen larger drops in reading and math scores. Here, I do not find evidence that the gap is closing.

Examining the interactions of racial demographics with teaching modality confirms that school districts that spent more of the 2020–2021 school year in person saw a smaller drop in average math and reading scores. Columns (5) and (6) show the interaction between the percentage of the school district that is Black and teaching modality. I find evidence that school districts with more Black students saw smaller gains associated with in-person learning. In 2022, districts with a high percentage of Black students that spent more of the 2020–2021 school year in person had lower average reading scores than those that spent less time in person. Looking at 2023 test scores, this gap has only grown. In 2022, there was no noticeable difference in math scores between modalities within majority-Black school districts. By 2023, a significant gap appeared, with in-person school districts again performing relatively worse. Another way to interpret this is that, among schools that spent more of the 2020–2021 school year in person, those with a higher share of Black students saw lower returns to in-person learning, with returns perhaps even negative in the long run.

As shown in Figures C8 and C9, trends leading up to 2019 are not parallel, so these results should not be interpreted as causal. Accordingly, the findings are not robust to a violation in parallel trends. How-

ever, the fact that the estimates are increasing in magnitude suggests that this trend warrants further investigation. In particular, among fully in-person school districts (top panel of C9), districts in "Quartile 1" (i.e., districts with the fewest Hispanic students) saw a much greater recovery from 2022 to 2023 than other districts.

Subgroup Analysis: Economically Disadvantaged

Analysis of scores by socioeconomic status (using data set D) shows a relatively small growth in the math test score gap within school districts from 2019 to 2022 and only weak evidence of growth in a reading score gap (see Table 3, columns (1) & (2)). This lack of a score gap contrasts with the much larger growth in test score gaps between districts, suggesting that the effect of the pandemic acted at the district level, rather than the individual level. Figure 2 shows score trends from 2016 to 2023 for both subgroups. The figure shows that, while the initial drop in scores was similar for both subgroups, scores of higher income students increased more from 2022 to 2023 than scores of economically disadvantaged students. Table 3 also shows this, with the negative value on the interaction term for year and economically disadvantaged larger in magnitude for 2023 than 2022. Overall, this is evidence of an unequal recovery, where economically disadvantaged students are not recovering from the pandemic learning loss as quickly as their higher income classmates.

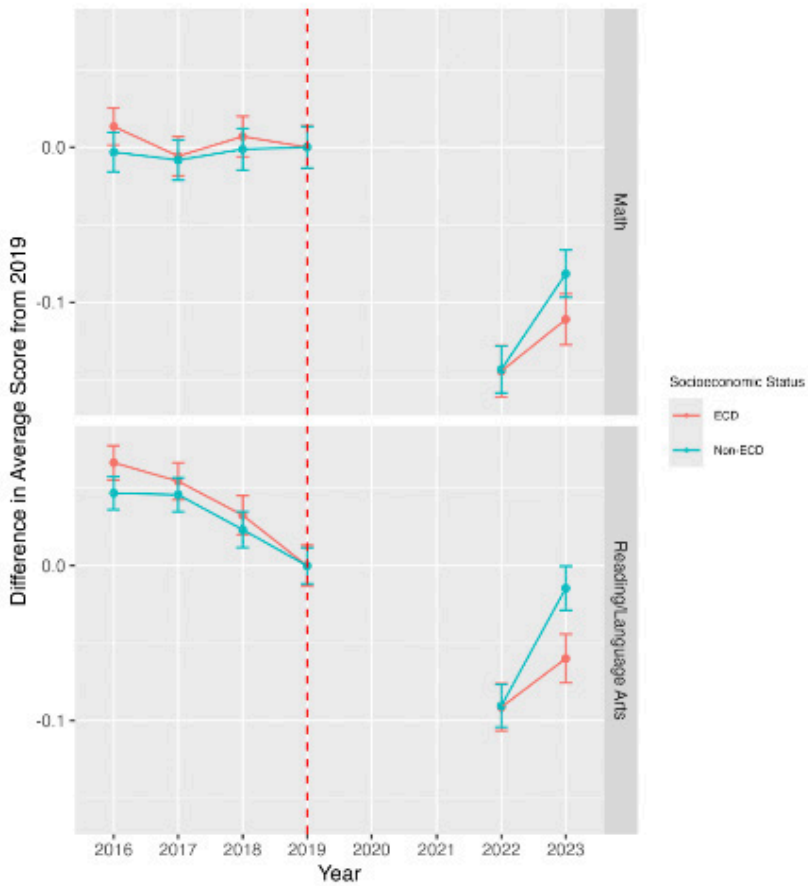


Figure 2: Average scores by subject and socioeconomic status - data set D

The DDD regression (Table 3, columns (3) & (4)) does not provide evidence for differential returns to in-person learning by socioeconomic status. However, there is evidence that in-person learning during the 2020–2021 school year is associated with higher scores, at least for higher income students.

Columns (5) and (6) of Table 3 include the average COVID-19 death rate over the 2020–2021 school year. Schools in counties with more COVID-19 deaths saw a larger drop in average reading scores but no change in average math scores. Furthermore, controlling for COVID-19 deaths reveals a negative interaction between the economically disadvantaged subgroup and in-person modality on reading scores, most significant in 2023 (see the row labeled 2023 × ECD × In Person). This contrasts with the finding shown in Table 1, that districts with a higher percentage of economically disadvantaged students had a positive relationship between in-person learning and math scores. In-person learning appears to have reduced the growth in inequality in math scores but may have widened gaps in reading scores. Due to violations in parallel trends, neither of these findings can be interpreted as causal. Finally, the model with COVID deaths suggests that the effect of in-person learning cannot be explained by death rates, nor can the slower recovery of economically disadvantaged students.

To check whether results are due to pre-COVID trends, I include a model with the change in score from 2018 to 2019 fully interacted in Table B5. The results are similar, suggesting that these post-COVID trends cannot be explained by districts' pre-COVID trends. Despite this, analysis based on Rambachan and Roth shows a lack of robustness, so these results should be interpreted with caution.
Table B3 shows the regressions discussed above using data set C and excluding 2023 test score data. This makes use of a larger data set and shows similar results, suggesting that the results are not due to a biased sample.

Table 4 shows the regression with individual socioeconomic status and district socioeconomic status specified in Equation 9. This demonstrates that, on average, students in districts with a high percentage of economically disadvantaged students performed worse in all years, regardless of their own economic status. Conversely, controlling for the demographics of a school district, it still appears that economically disadvantaged students are falling behind in the recovery from 2022 to 2023. This finding is robust to a violation of parallel trends. The coefficient on the interaction term between individual SES and district SES is positive, although this finding is less robust. These scores are standardized, meaning group averages can only reasonably get so far from zero; this positive interaction may be partially due to this. There is good evidence that, in 2023, the COVID-related SES score gap is smaller in districts with more economically disadvantaged students. Essentially, the pandemic caused a disproportionate drop in test scores in high-poverty school districts from 2019 to 2022, which continued into 2023. These effects were felt relatively equally among all students in these school districts. Furthermore, in low-poverty school districts, economically disadvantaged students have seen less recovery in test scores from 2022 to 2023. This trend does not hold in

	Dependent variable:					
	Math	RLA	Math	RLA	Math	RLA
	(1)	(2)	(3)	(4)	(5)	(6)
2022	−0.102***	−0.065***	−0.117***	−0.074***	−0.117***	−0.081***
	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
2023	−0.071***	0.072***	−0.081***	0.064***	−0.084***	0.054***
	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
ECD	−0.591***	−0.573***	−0.603***	−0.586***	−0.565***	−0.555***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
2022 × ECD	−0.017**	−0.017**	−0.011	−0.009	−0.009	−0.003
	(0.007)	(0.007)	(0.009)	(0.008)	(0.009)	(0.009)
2023 × ECD	−0.035***	−0.043***	−0.033***	−0.035***	−0.026***	−0.023**
	(0.007)	(0.007)	(0.009)	(0.008)	(0.009)	(0.009)
2022 × In Person			0.085***	0.045***	0.085***	0.053***
			(0.017)	(0.016)	(0.017)	(0.016)
2023 × In Person			0.061***	0.039**	0.065***	0.051***
			(0.017)	(0.016)	(0.017)	(0.016)
ECD × In Person			0.049***	0.056***	−0.002	0.014
			(0.016)	(0.015)	(0.016)	(0.015)
2022 × COVID					0.001	−0.015**
					(0.008)	(0.007)
2023 × COVID					−0.002	−0.020***
					(0.008)	(0.007)
ECD × COVID					0.095***	0.083***
					(0.007)	(0.007)
2022 × ECD × In Person			−0.022	−0.032	−0.024	−0.038*
			(0.023)	(0.022)	(0.023)	(0.022)
2023 × ECD × In Person			−0.004	−0.034	−0.016	−0.052**
			(0.023)	(0.022)	(0.023)	(0.022)
2022 × ECD × COVID					−0.0003	0.004
					(0.010)	(0.010)
2023 × ECD × In Person					0.012	0.022**
					(0.010)	(0.010)
2022 × In Person × COVID					−0.014	0.002
					(0.018)	(0.017)
2023 × In Person × COVID					−0.014	0.004
					(0.018)	(0.017)
ECD × In Person × COVID					−0.088***	−0.076***
					(0.018)	(0.017)
2022 × ECD × In Person × COVID					0.014	0.007
					(0.025)	(0.024)
2023 × ECD × In Person × COVID					0.020	0.003
					(0.025)	(0.024)
Observations	9,414	9,414	9,414	9,414	9,414	9,414
R^2	0.860	0.867	0.861	0.867	0.862	0.867
Adjusted R^2	0.831	0.840	0.832	0.840	0.834	0.839
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 3: Regression of scores on socioeconomic status (2023 included) - base year 2019

high-poverty school districts.

Table B4 shows the same regression (specified in Equation 9) using data set C (and thus excluding 2023). The positive coefficient on the interaction term between 2022 and the ECD subgroup for math scores suggests that, during the pandemic, economically disadvantaged students gained ground in low-poverty districts. The negative DDD term suggests that this trend did not hold to the same degree in high-poverty districts. Since the data are not nationally representative, these results should be interpreted with some skepticism, but they support the conclusion that students in high-poverty schools were most affected by the pandemic, regardless of their own socioeconomic status. This is consistent with the findings of E. M. Fahle

	Dependent variable:	
	Math	RLA
	(1)	(2)
2022	−0.031**	0.006
	(0.013)	(0.012)
2023	0.006	0.156***
	(0.013)	(0.012)
ECD	−0.746***	−0.636***
	(0.012)	(0.012)
2022 × ECD	0.017	−0.014
	(0.018)	(0.018)
2023 × ECD	−0.045**	−0.097***
	(0.018)	(0.018)
2022 × % District ECD	**−0.186*** **	**−0.188*** **
	(0.024)	(0.023)
2023 × % District ECD	**−0.201*** **	**−0.225*** **
	(0.024)	(0.024)
ECD × % District ECD	0.286***	0.098***
	(0.023)	(0.023)
2022 × ECD × % District ECD	−0.009	0.048
	(0.034)	(0.034)
2023 × ECD × % District ECD	0.075**	0.160***
	(0.034)	(0.034)
Observations	9,414	9,414
R^2	0.874	0.874
Adjusted R^2	0.848	0.849
District Fixed Effects	Yes	Yes
State-Year Dummies	Yes	Yes
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 4: Regression for interaction of ECD and percent of district ECD - base year 2019

et al.,[37] that the mechanisms that drove learning loss affected students within a district to a similar degree, regardless of each individual student's background.

Subgroup Analysis: Race

A DD regression with race as subgroups as in equation (7) is shown in columns (1) and (2) of Table 5 (using data set E). I find no significant growth in the Black-white test score gap from 2019 to 2022, but the white-Hispanic score gap grew in both math and reading. This is not robust to differential pre-trends. Figure C5 shows score trends from 2016 to 2023 for all three subgroups. The DDD regression specified in equation (9) shows no evidence of differential returns to in-person learning by race. When modality interaction terms are added, the growth in the white-Hispanic gap is no longer significant at the 95% level. Columns (5) and (6) include COVID-19 death rates and, similarly to their inclusion in the SES regressions, do not significantly alter the results.

Table B6 shows columns (1)-(4) using data set F and including 2023 scores. Like non-ECD students, white students have recovered significantly from 2022 to 2023, at least in reading scores. There is also notable growth in the white-Hispanic score gap that does not shrink from 2022 to 2023. There are no significant interactions with modality. The uneven recovery can also be

37 E. M. Fahle et al., School District and Community Factors Associated With Learning Loss During the COVID-19 Pandemic.

seen in Figure C5. While trends are similar before 2020, Hispanic students saw the largest drop in test scores in both reading and math and are still further behind 2019 levels than Black or white students. Table B8 includes the difference in 2019 and 2018 scores by district (pre-trends). The interactions with the pre-trends are not significant, suggesting previous trends within school districts do not explain my results. Interactions between race and the percent of the district that is economically disadvantaged are shown in table B7. The interaction terms are not significant in 2022 or 2023.

Discussion

My findings suggest that the growth in test score gaps from 2019 to 2022 is largely

	Dependent variable:					
	Math	RLA	Math	RLA	Math	RLA
	(1)	(2)	(3)	(4)	(5)	(6)
2022	−0.103***	−0.034	−0.111***	−0.038	−0.112***	−0.041
	(0.030)	(0.032)	(0.031)	(0.033)	(0.030)	(0.032)
Black	−0.777***	−0.725***	−0.812***	−0.750***	−0.790***	−0.731***
	(0.013)	(0.013)	(0.017)	(0.017)	(0.017)	(0.017)
Hispanic	−0.520***	−0.567***	−0.545***	−0.586***	−0.505***	−0.544***
	(0.012)	(0.013)	(0.017)	(0.018)	(0.017)	(0.018)
2022 × Black	−0.028	−0.013	−0.012	−0.005	−0.009	−0.004
	(0.019)	(0.019)	(0.024)	(0.026)	(0.024)	(0.026)
2022 × Hispanic	−0.047**	−0.042**	−0.043*	−0.037	−0.040	−0.036
	(0.018)	(0.019)	(0.025)	(0.026)	(0.026)	(0.027)
2022 × In Person			0.060	0.028	0.063	0.036
			(0.041)	(0.043)	(0.039)	(0.041)
Black × In Person			0.128***	0.091**	0.100**	0.066
			(0.041)	(0.043)	(0.040)	(0.043)
Hispanic × In Person			0.078**	0.062	0.027	0.008
			(0.039)	(0.041)	(0.038)	(0.041)
2022 × COVID					−0.006	−0.015
					(0.018)	(0.019)
Black × COVID					0.090***	0.073***
					(0.019)	(0.020)
Hispanic × COVID					0.148***	0.154***
					(0.021)	(0.023)
2022 × Black × In-Person			−0.053	−0.027	−0.058	−0.029
			(0.061)	(0.064)	(0.059)	(0.063)
2022 × Hispanic × In-Person			−0.010	−0.013	−0.011	−0.015
			(0.057)	(0.059)	(0.056)	(0.060)
2022 × Black × COVID					0.012	0.004
					(0.028)	(0.030)
2022 × Hispanic × COVID					0.009	0.008
					(0.031)	(0.033)
2022 × In-Person × COVID					0.010	0.033
					(0.038)	(0.040)
Black × In-Person × COVID					−0.054	−0.065
					(0.043)	(0.046)
Hispanic × In-Person × COVID					−0.109***	−0.122***
					(0.042)	(0.045)
2022 × Black × In-Person × COVID					−0.020	−0.010
					(0.064)	(0.067)
2022 × Hispanic × In-Person × COVID					−0.001	−0.020
					(0.062)	(0.066)
Observations	1,488	1,488	1,488	1,488	1,488	1,488
R^2	0.854	0.834	0.855	0.833	0.865	0.842
Adjusted R^2	0.823	0.797	0.823	0.796	0.833	0.804
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5: Regression of score on race - base year 2019

attributable to district-level differences; growth in gaps between high-poverty and low-poverty schools is more pronounced than growth in gaps between economically disadvantaged students and higher income students. This growth in score gaps is larger in math than in reading. Results regarding race and ethnicity (Black and Hispanic) are similar; districts with more Black or Hispanic students saw a larger drop in test scores from 2019 to 2022 and this unequal drop has persisted into 2023.

Recall that E. M. Fahle et al.[38] and Goldhaber et al.[39] both find that high-poverty districts were more affected by remote and hybrid learning. My findings support this claim and suggest that the effect is stronger on math scores than on reading scores. I extend this analysis to 2023 test scores and find that high-poverty districts that spent more of the 2020–2021 school year in-person still have higher math scores than high-poverty districts that were remote or hybrid. Despite this equalizing effect of in-person learning, I find evidence that school districts with more Black students saw a smaller effect of in-person learning and perhaps even a negative impact on reading scores. This is unexpected and difficult to explain, so more research should be done to draw accurate conclusions from this finding. This result could also be due to differential trends before the pandemic. Overall, in-person instruction in 2020–2021 appears to have supported math learning more than reading learning. To the extent that in-person instruction uniquely helped disadvantaged school districts, the impact was on math scores.

This does not constitute an argument that schools should have been reopened sooner. There is reason to believe[40] that school reopenings led to an increase in COVID transmission.[41][42][43] Rather, cogent analysis requires consideration of all aspects of a policy; this paper contributes to the larger discussion by addressing the specific areas of learning loss and test score gaps.

Overall, my results suggest that high-poverty districts were harmed more by the pandemic than low-poverty districts. Low-poverty districts had the most noticeable and consistent drop in test scores and there is no indication that the gap has been closing.

Furthermore, economically disadvantaged students in low- or mid-poverty school districts have seen slower recovery from 2022 to 2023 than their peers. Thus, the effect of the pandemic on education can be split into two main periods: the initial drop (2019–2022) where high-poverty and high-minority school districts saw a larger drop in test scores than low-poverty or majority-white districts, and the recovery (2022–2023) where high-poverty and high-minority districts are still lagging behind and, additionally, economically dis-

38 E. M. Fahle et al., School District and Community Factors Associated With Learning Loss During the COVID-19 Pandemic.

39 Goldhaber et al., "The Consequences of Remote and Hybrid Instruction During the Pandemic."

40 Isphording, Lipfert, and Pestel 2021 find no association between schools reopening and increased transmission rates in Germany. This paper does not attempt to evaluate the conflicting results.

41 Emanuele Amodio et al., "Schools Opening and COVID-19 Diffusion: Evidence From Geolocalized Microdata," European Economic Review 143, (April 2022): 2914-2921. https://doi.org/10.1016/j.Euroecorev.2021.104003.

42 Victor Chernozhukov, Hiroyuki Kasahara, and Paul Schrimpf, "The association of opening K–12 schools with the spread of COVID-19 in the United States: County-level panel data analysis," Proceedings of the National Academy of Sciences of the United States of America 118, no. 42 (October 19, 2021), https://doi.org/ 10.1073/pnas.2103420118.

43 Dan Goldhaber et al.,"To What Extent Does In-Person Schooling Contribute to the Spread of COVID-19? Evidence from Michigan and Washington," (Working Paper No. 28455, NBER, 2021), http://www.nber.org/papers/w28455

advantaged students in low-poverty school districts are experiencing slower recovery than their peers.

The present paper uses school districts as the smallest unit of analysis. This obscures the important differences between schools within a school district and children within a school. More analysis at the school level could help determine where inequalities are growing. Analysis using panel data could also be insightful.

Another limitation is that test score data were averaged across all students in grades 3 through 8. It is reasonable to believe that school closures and other aspects of the pandemic affected children in 3rd grade differently than those in 8th grade. Furthermore, analysis by gender could potentially show interesting differences. While this analysis highlights important trends, analysis with more specific subgroups could be compelling.

Alternatively, access to more data would improve this analysis. As seen in the sample composition tables, my analysis included test scores from millions of students, but there are many states, school districts, and students missing. While this analysis suggests national trends, a nationally representative study would be ideal.

Future research should examine how districts can address these inequalities. Specifically, it would be useful to see how much funding is needed to close score gaps and how that funding should be allocated in order to be most effective. During the pandemic, the American Rescue Plan Elementary and Secondary School Emergency Relief Fund allocated 122 billion dollars to state education agencies. Halloran et al.[44] find no correlation between this spending and recovery, but they note that impacts may not be visible yet.

***

for acknowledgments and appendixes please visit www.tulanejournal.com

***

44 Halloran et al., Post COVID-19 Test Score Recovery.

Grocery Cart Choices and Health Outcomes: A State-Level Analysis of Dietary Patterns and Socio-Economic Factors in the U.S.

Maria Laura Peña, Rafael Nicastro, Juan Manueal Peña
The University of Texas at Austin

Abstract

This paper explores the relationship between grocery consumption patterns and health outcomes across U.S. states. By analyzing data ranging from 2021 to 2023, we investigate how the composition of an average grocery cart correlates with health metrics such as obesity, diabetes, heart disease, and life expectancy. We focus on understanding whether differences in consumption are driven by price, education, or other socio-economic variables and how these factors contribute to health disparities. Utilizing tools like Python, R, and Stata, we employ multivariate regression, clustering, and decision tree models to identify key trends. Our findings reveal that wealthier regions such as the District of Columbia, where prices for healthier foods like fruits and vegetables tend to be higher, are associated with better health outcomes as well as proof that increased consumption of processed foods is linked to negative health factors. Socio-economic factors, including income, education, and tobacco consumption, play a critical role in shaping health outcomes. This research contributes to policy discussions on food accessibility, pricing strategies, and health interventions aimed at improving dietary habits and reducing health inequalities.

Introduction

The relationship between dietary habits and health outcomes is a critical area of study, both in public health and economics. As governments, policymakers, and health organizations increasingly focus on improving public health, understanding the factors influencing food consumption patterns across different regions becomes paramount. The question of how the composition of grocery purchases affects health is multifaceted, involving variables such as income, education, food prices, habits, and cultural beliefs. The significance of this research lies in the potential for actionable insights that could inform policy decisions on subsidies, taxation, and educational programs to promote healthier dietary choices. For example, Mississippi and West Virginia, which have some of the highest poverty rates in the U.S., also report higher rates of obesity[1] and diabetes,[2] likely

1 County Health Rankings & Roadmaps. "Adult Obesity." Last modified 2024. https://www.countyhealthrankings.org/health-data/health-factors/health-behaviors/diet-and-exercise/adult-obesity?year=2024.

2 County Health Rankings & Roadmaps. "Dia-

influenced by a reliance on lower-cost, processed foods. In contrast, states like California and Massachusetts, where median incomes are higher,[3] tend to have greater access to fresh produce and report lower rates of diet-related illnesses.[4]

In this study, we seek to explore how average grocery baskets—representing the types and quantities of alimentary goods consumed—correlate with health outcomes like obesity, diabetes, heart disease, and life expectancy. We aim to answer three key questions: (1) Does the composition of an average grocery cart influence health outcomes? (2) How do grocery baskets differ across U.S. states, and are these differences related to health disparities? (3) Are the variations in grocery baskets driven primarily by food prices, education, or some other factor?

We leverage datasets from the United States, focusing on state-level data to analyze regional dietary patterns and health outcomes. The data spans the years 2021 to 2023, offering a comprehensive view of recent consumption patterns and health outcomes. Our analysis covers a range of food categories, from fruits, vegetables, and meat to processed foods, dairy, and oils, providing a holistic view of dietary patterns. The focus of this paper is on our results from the US states' data.

To address our research questions, we employed a combination of exploratory data analysis (EDA), multivariable regression, and machine learning techniques such as clustering and decision trees. This multifaceted approach allows us to not only assess correlations between food consumption and health, but to also delve into the underlying factors—like price per unit (PPU), average revenue per capita (ARPC), education, and income—that may explain regional differences in grocery carts. Moreover, by examining the price sensitivity of certain food categories, we can better understand how affordability impacts dietary choices and, consequently, health outcomes.

The implications of this research are wide-reaching. By identifying the factors most closely associated with healthier grocery carts and better health outcomes, we can provide valuable insights for policymakers. For instance, targeted subsidies on healthier food items, increased education on nutrition, or adjustments to taxation policies could help bridge the health gap across regions. Moreover, by later expanding our analysis to include multiple countries, we hope to uncover global trends that could lead to broader, more effective public health interventions as well as methods to increase health benefits, thereby decreasing negative health factors without substantially reducing demand for products and harming business owners.

This study aims to contribute to the growing body of literature linking diet and health by offering a detailed, data-driven perspective on the complex interplay between food prices, consumption habits, and health outcomes. In doing so, we seek to not only uncover the drivers behind regional differences in dietary patterns but

betes Prevalence." Last modified 2024. https://www.countyhealthrankings.org/health-data/health-outcomes/quality-of-life/diabetes-prevalence?year=2024.

3 Federal Reserve Bank of St. Louis. "2023, Release Tables: Per Capita Personal Income by State, Annual | FRED | St. Louis Fed." https://fred.stlouisfed.org/release/tables?rid=110&eid=257197&od=#.

4 Centers for Disease Control and Prevention. "Heart Disease Mortality." Last reviewed February 25, 2022. https://www.cdc.gov/nchs/pressroom/sosmap/heart_disease_mortality/heart_disease.htm.

also provide concrete suggestions for how to improve public health through informed policy decisions.

Literature Review

The relationship between food prices, dietary choices, and health outcomes has been extensively studied across various disciplines, from economics and public health to nutrition and food policy. Despite the growing body of literature, a significant debate remains concerning the nature of these relationships and the factors that drive food choices. Understanding the influence of food prices and accessibility on health outcomes is crucial for informing policy decisions aimed at improving public health and reducing inequalities. This review summarizes key findings from existing studies and outlines the gaps that our research seeks to address.

According to the Dietary Guidelines for Americans, 2020–2025, a healthy dietary pattern should consist of nutrient-dense foods and beverages across all food groups, with a focus on portion control and limiting added sugars, saturated fats, sodium, and alcohol.[5] While these guidelines provide a framework for what constitutes a healthy diet, the accessibility of such foods varies widely across regions. Food deserts—areas with limited access to nutritious, but affordable food—are a persistent issue in the United States.[6] Alan Dangour of the Leverhulme Centre for Integrative Research on Agriculture and Health emphasizes that food availability, particularly fresh fruits and vegetables, is a key determinant of health outcomes.[7] Low income and rural areas are disproportionately affected by these food deserts, which limit access to healthier food options and contribute to dietary inequalities.

The role of food prices is another critical factor in shaping dietary habits. A study published in Environment and Urbanisation by Marc Cohen and James Garrett called "The food price crisis and urban food (in)security" suggests that food demand is highly sensitive to price changes, particularly among lower-income groups and larger families. As prices rise, consumers often shift to cheaper, less healthy alternatives, exacerbating health disparities among socio-economically disadvantaged populations.[8] This price sensitivity is particularly relevant when considering the affordability of healthier food options, such as fruits, vegetables, and lean proteins, which are often priced higher than processed and sugary foods. This economic reality contributes to the prevalence of diet-related health issues like obesity and diabetes in low-income communities.[9]

The relationship between income and obesity prevalence further underscores these disparities. A 2018 study published in Nature found that in U.S. states where median household incomes were below $45,000 per year, over 35% of the popu-

5 U.S. Department of Agriculture and U.S. Department of Health and Human Services, Dietary Guidelines for Americans, 2020-2025, 9th ed. (2020).

6 Ploeg, Michele V. 2010. "Access to Affordable, Nutritious Food Is Limited in "Food Deserts" | Economic Research Service." USDA Economic Research Service. https://www.ers.usda.gov/amber-waves/2010/march/access-to-affordable-nutritious-food-is-limited-in-food-deserts.

7 Alan D. Dangour, Georgina Mace, and Bhavani Shankar, "Food systems, nutrition, health and the environment," The Lancet Planetary Health 1, no. 1 (2017): e8 - e9, doi: 10.1016/S2542-5196(17)30004-9.

8 Marc J. Cohen and James L. Garrett, "The food price crisis and urban food (in)security," Environment and Urbanization 22, no. 2 (2010): 467-482, https://doi.org/10.1177/0956247810380375.

9 Centers for Disease Control and Prevention. "BRFSS Prevalence & Trends Data: Explore by Topic." Accessed August 10th, 2024. https://www.cdc.gov/brfss/brfssprevalence/.

lation was obese. In contrast, obesity rates were less than 25% in states where median incomes were above $65,000.[10] This correlation suggests that economic constraints directly impact dietary quality, leading to higher rates of obesity in lower-income states. Such disparities highlight how income levels not only affect access to food but also the quality of dietary choices, reinforcing the need for policy interventions that improve affordability and accessibility of healthier foods.

Studies focusing on the intersection of food prices and health outcomes have revealed important trends. A comprehensive analysis of food pricing dynamics highlights the role of agricultural costs, transportation, and market conditions in shaping food affordability.[11] Variations in food production due to climate, technological advancements, and economic policies create price volatility, which in turn affects dietary choices. Low-income communities are particularly vulnerable to price fluctuations, as they often rely on affordable but less nutritious foods.[12] Moreover, targeted price interventions, such as subsidies on healthier foods or taxes on unhealthy items, have been shown to influence consumer behavior and dietary choices. A systematic review of subsidy interventions found that nearly all studies reported significant increases in the purchase and consumption of subsidized healthier foods, such as fruits, vegetables, and whole grains.[13]

While there is extensive research on food prices, dietary choices, and health outcomes, key gaps remain in how these factors interact at the regional level and how they vary based on socio-economic disparities. Much of the existing literature treats these factors separately, either examining the impact of food prices on dietary habits or the effect of dietary choices on health outcomes, but not integrating these perspectives holistically.

Additionally, most studies focus on national or individual-level data, whereas there is limited research at the state level, which is crucial for understanding localized food policies and health disparities within a country as diverse as the U.S. Our study aims to fill this gap by using state-level data across multiple years (2021-2023) to analyze how food prices, education, and socio-economic status shape dietary choices and health outcomes on a regional scale. Furthermore, few studies employ machine learning techniques such as clustering and decision trees to capture non-linear interactions between variables, despite dietary behaviors being shaped by complex, multi-dimensional factors. By applying multivariable regression, clustering, and decision tree models, this study provides a more robust analytical framework to uncover underlying patterns that traditional statistical approaches may overlook.

In doing so, this research contributes to both economic and public health discussions by providing region-specific insights

10 R. Alexander Bentley, Paul Ormerod, and Damian J. Ruck, "Recent origin and evolution of obesity-income correlation across the United States," Palgrave Communications 4, no. 146 (2018), https://doi.org/10.1057/s41599-018-0201-x.

11 GGI Insights, "Food Prices: Understanding and Navigating Market Dynamics," Gray Group International, October 1, 2024, https://www.graygroupintl.com/blog/food-prices.

12 Drewnowski, A., & Specter, S. E. (2004). Poverty and obesity: The role of energy density and energy costs. American Journal of Clinical Nutrition, 79(1), 6–16. https://doi.org/10.1093/ajcn/79.1.6

13 Ruopeng An, "Effectiveness of Subsidies in Promoting Healthy Food Purchases and Consumption: A Review of Field Experiments," Public Health Nutrition 16, no.7 (2013): 1215-28, https://doi.org/10.1017/S1368980012004715.

for policymakers, which can inform targeted interventions such as subsidies for healthier foods, improvements in food accessibility, and education programs to promote healthier dietary habits.

In summary, the existing literature underscores the complexity of the relationships between food prices, dietary choices, and health outcomes. While economic and accessibility factors play significant roles, cultural beliefs and education also influence consumption patterns. Our study aims to build on these insights by using a multi-state, multi-variable approach that incorporates economic modeling to reach conclusions.

Methods and Data

This study focuses on analyzing the relationship between grocery consumption patterns and health outcomes across U.S. states from 2021 to 2023. Our main objective is to understand how the composition of items commonly found in a grocery cart correlates with key health indicators, such as obesity rates, diabetes prevalence, physical inactivity, heart disease mortality, and life expectancy. To achieve this, we employ various statistical techniques and machine learning models, drawing from a rich dataset that covers a wide range of food categories, prices, and socio-economic variables.

Data

The data used in this study is drawn from multiple sources and consists of consumption, price, and health outcome variables aggregated at the state level. Our analysis spans 50 U.S. states (excluding the District of Columbia for visualization purposes), focusing on three primary categories of variables:

1. Grocery Consumption Data: We include data on the average consumption (in kilograms) per capita of various food products. These are further divided into subsegments, such as:

Fruits & Nuts: Including fresh, processed, and frozen fruits, as well as nuts.
Meat: Including fresh meat, processed meat, and meat substitutes.
Dairy Products & Eggs: Including milk, eggs, yogurt, cheese, and other dairy products.
Vegetables: Including fresh, processed, and frozen vegetables.
Convenience Foods: Including ready-to-eat meals and soups.
Other categories: Bread & cereal products, confectionery & snacks, oils & fats, sauces & spices, spreads & sweeteners, and pet food.

2. Price Data: The dataset also includes the Price Per Unit (PPU) for each of the food categories in U.S. dollars, representing the average price consumers pay for these products. This allows us to analyze how food prices may influence consumption patterns and health outcomes.

3. Health Outcome Data: Health outcomes for each state are measured using several key indicators, including:
Obesity Rate (% of adults)
Diabetes Prevalence (% of adults, age-adjusted)
Physical Inactivity (% of adults, age-adjusted)
Heart Disease Mortality Rate (number of deaths per 100,000 people)
Life Expectancy (average life expectancy in

years)

4. Socio-Economic Variables: To account for the broader socio-economic context, we include data on:

Income Per Capita (in U.S. dollars)
obacco Use (% of adults)
Educational Attainment (% of adults with a four-year degree)

Methods

1. Exploratory Data Analysis (EDA)
The first step in our analysis involved visualizing the data to uncover preliminary trends and patterns. We created bar graphs to compare the average consumption of different food categories across states, heatmaps to identify correlations between food consumption and health outcomes, and scatter plots to explore the relationship between prices and consumption levels.

2. Multivariate Regression
To quantify the relationship between grocery consumption and health outcomes, we employed multiple linear regression models. The dependent variable in our models was the Health Index, which we constructed by combining the obesity rate, diabetes prevalence, heart disease mortality, and life expectancy measurements into a single score. Independent variables included the average consumption levels of various food categories, Price Per Unit (PPU) for each category, and control variables such as income, tobacco use, and education levels.

We ran several models to test different combinations of independent variables. One model included both PPU and Average Revenue Per Capita (ARPC) as explanatory variables, while another excluded these price-related variables to assess the direct effect of food consumption on health outcomes.

3. Clustering Analysis
While multivariate regression is effective for examining linear relationships, it may miss non-linear patterns and latent groupings in the data. To address this, we used clustering techniques, specifically k-means clustering. We categorized states into clusters based on their consumption of key food categories like fruits, vegetables, meat, and processed foods. This method allowed us to group states based on similarities in grocery consumption patterns, which helped us to identify dietary clusters across the U.S. and examine whether states with similar grocery baskets exhibit common health outcomes.

Unlike regression, clustering does not impose a predefined relationship between variables, making it especially useful for detecting emerging consumption patterns that might not be evident through traditional statistical methods.

4. Decision Trees and Random Forests
To better capture non-linear relationships and interaction effects, we employed decision tree-based models, including random forests. These models allow for a hierarchical segmentation of the data, identifying threshold effects where specific levels of food consumption or pricing lead to distinct changes in health outcomes.

Unlike regression, which assumes a continuous and additive relationship between variables, decision trees divide the data into decision-based branches, making them more suitable for complex interactions. Additionally, feature importance scores from

random forests help identify the strongest predictors of health outcomes, highlighting the variables that influence public health the most.

By integrating regression, clustering, and tree-based models, our approach provides a holistic analysis, capturing both broad trends and subgroup patterns that may otherwise go unnoticed.

5. Ridge Regression
To address potential multicollinearity among highly correlated food categories and pricing variables, we applied Ridge Regression, a penalized regression technique. Using cross-validation, we selected an optimal regularization parameter (alpha = 0.0572), which stabilized coefficient estimates by reducing the influence of collinear predictors while preserving interpretability. All predictor variables were standardized to ensure proper regularization.

Model Evaluation

For each model, we evaluated performance using metrics such as the R-squared value, which indicates the proportion of variance in the health outcomes explained by the model, and the Mean Squared Error (MSE), which measures the accuracy of our predictions. We also checked for the statistical significance of individual variables, focusing on p-values to determine which factors most strongly influenced health outcomes.

Data Limitations

While our dataset is comprehensive for U.S. states, there are several limitations to consider. First, our analysis is based on aggregate state-level data, which may obscure individual-level variations in consumption and health. Second, the data covers only a three-year period (2021-2023), which may not fully capture long-term trends in dietary habits and health outcomes. Additionally, there were missing data points for certain key variables, such as heart disease and tobacco use, in 2023. To address this, we used the average of the 2021 and 2022 data to fill these gaps. Furthermore, while we account for variables like income and education, other potential confounders—such as healthcare access and local policy differences—are not included in our models. Lastly, the remaining impact of the COVID-19 pandemic during this period may have influenced both food purchasing behaviors and health outcomes, potentially skewing some of our findings. Despite these limitations, our study provides valuable insights into how food consumption patterns correlate with health outcomes at the state level.

Results

Our analysis of grocery consumption patterns and health outcomes across U.S. states from 2021 to 2023 revealed several key findings. We identified significant relationships between food consumption, food prices, and health outcomes such as obesity rates, diabetes prevalence, heart disease, and life expectancy. In this section, we highlight the most important results, supported by data visualization, and offer insights into the broader implications of our findings.

State-Level Consumption Patterns

The bar graphs above show significant regional differences in food consumption patterns across U.S. states. For example,

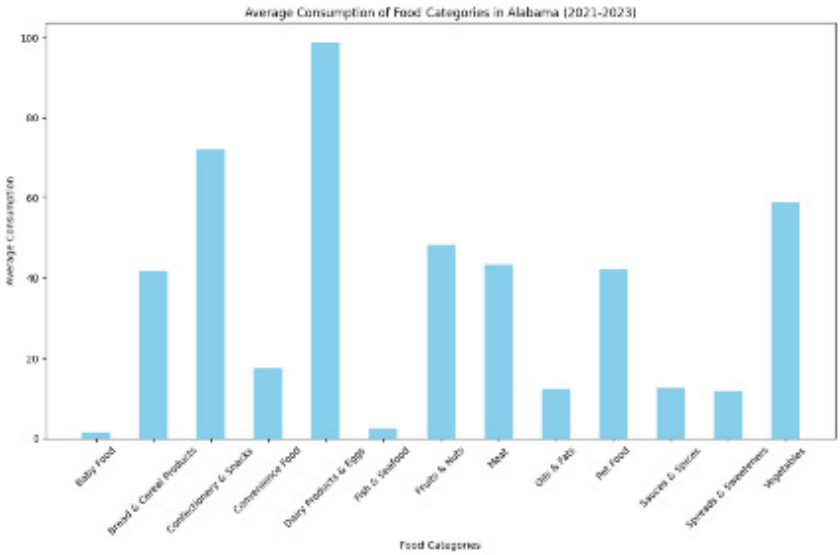


Fig. 1 - Bar Graph of Average Consumption of Food Categories, Alabama (2021-2023)

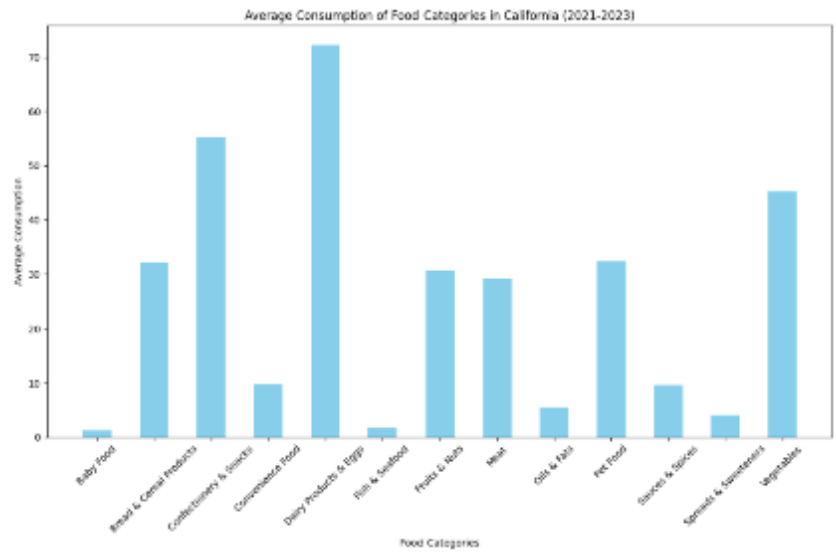


Fig. 2 - Bar Graph of Average Consumption of Food Categories, California (2021-2023)

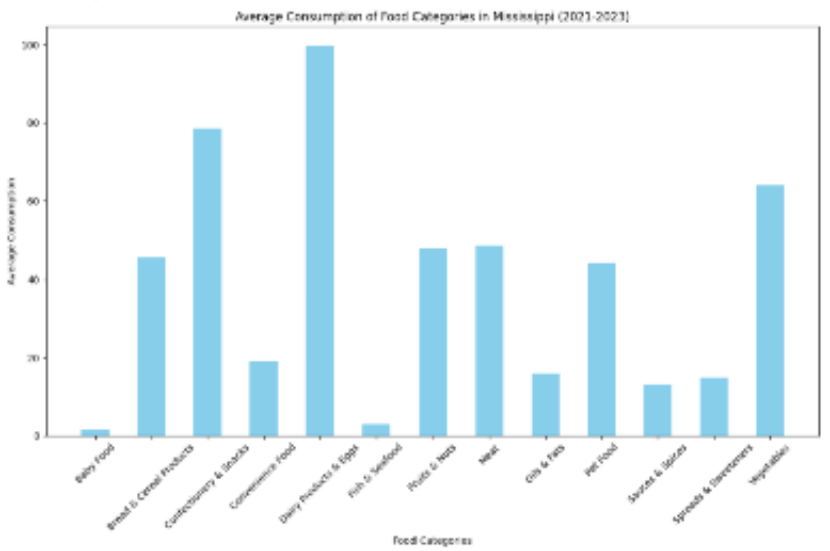


Fig. 3 - Bar Graph of Average Consumption of Food Categories, Mississippi (2021-2023)

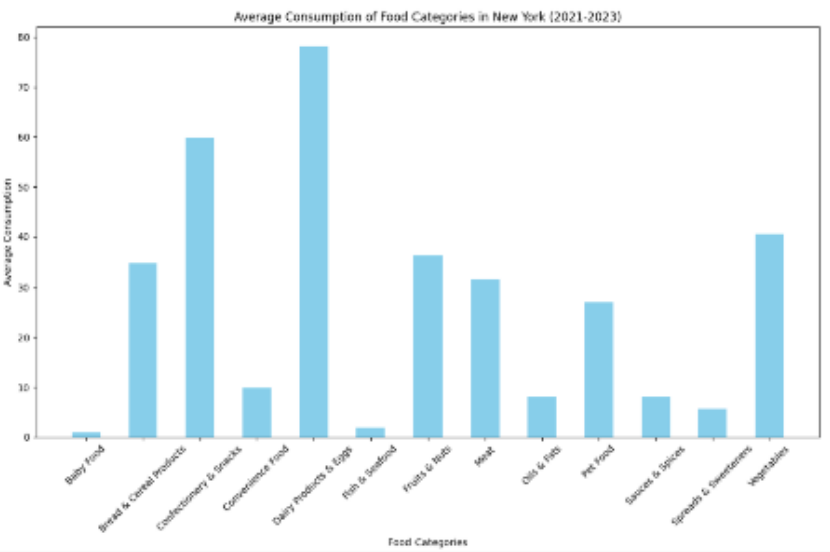


Fig. 4 - Bar Graph of Average Consumption of Food Categories, New York (2021-2023)

states such as California (Fig. 2) and New York (Fig. 4) tend to have higher consumption levels of fruits, vegetables, and fish, while states like Mississippi (Fig. 3) and Alabama (Fig. 1) consume more processed and convenience foods.

For example, California's strong agricultural industry supports greater availability of fresh produce,[14] contributing to higher consumption levels. Meanwhile, Mississippi, where food deserts are more prevalent, shows a greater reliance on processed and convenience foods, which aligns with higher obesity rates in the region.[15]

Explanation: This visualization highlights the disparities in dietary habits across states. Wealthier states with higher average incomes tend to consume healthier food

14 California Association of Food Banks. "Farm To Family: See How It Works." n.d. Accessed March 8, 2025. https://www.cafoodbanks.org/farm-to-family-see-how-it-works/ .

15 PBS news. "Mississippi 'Food Deserts' Fuel Obesity Epidemic." 2010. Accessed March 8, 2025. https://www.pbs.org/newshour/show/mississippi-food-deserts-fuel-obesity-epidemic .

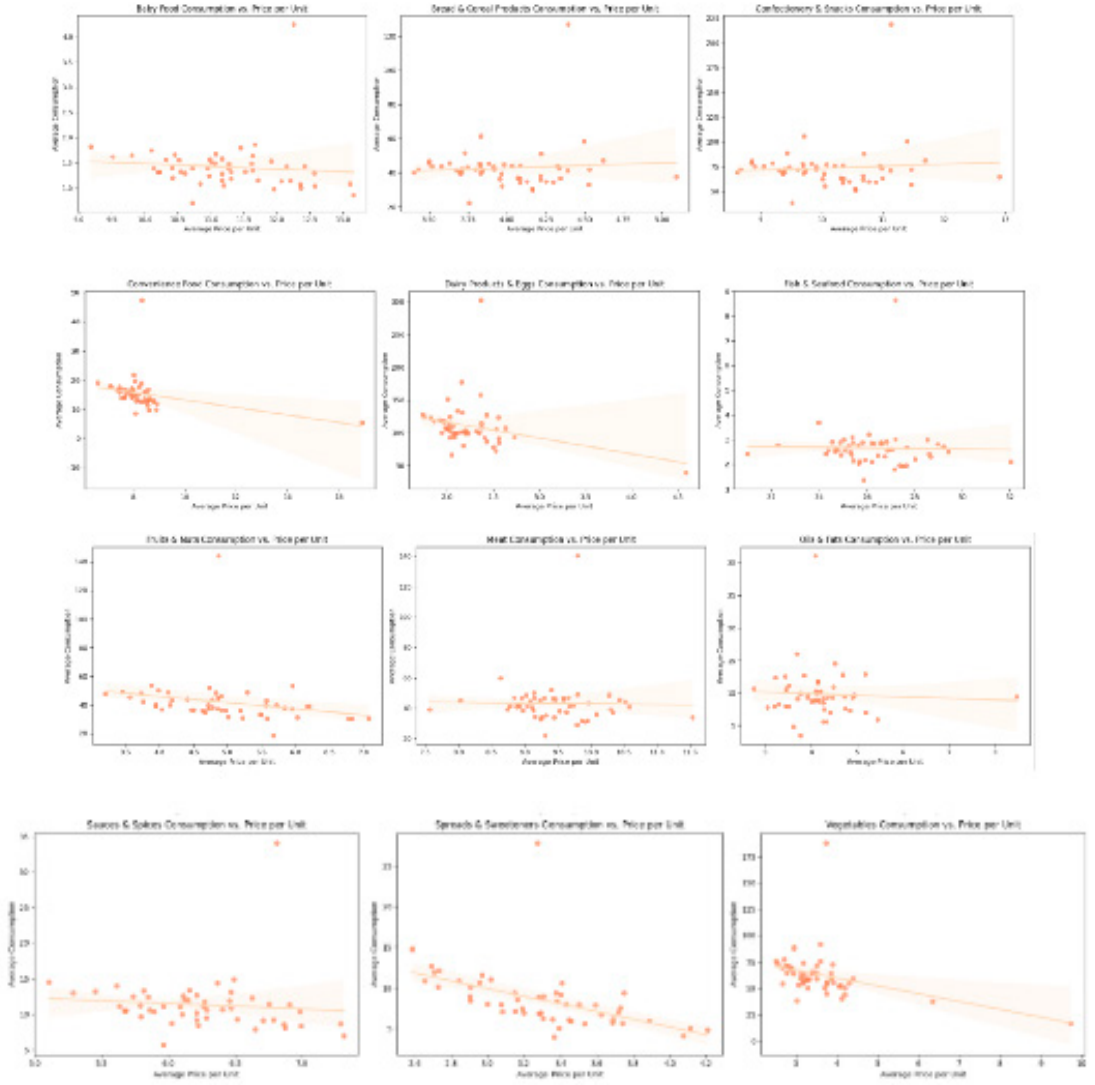


Fig. 5 - Scatter Plots of Price Per Unit vs. Consumption Levels for Key Food Products

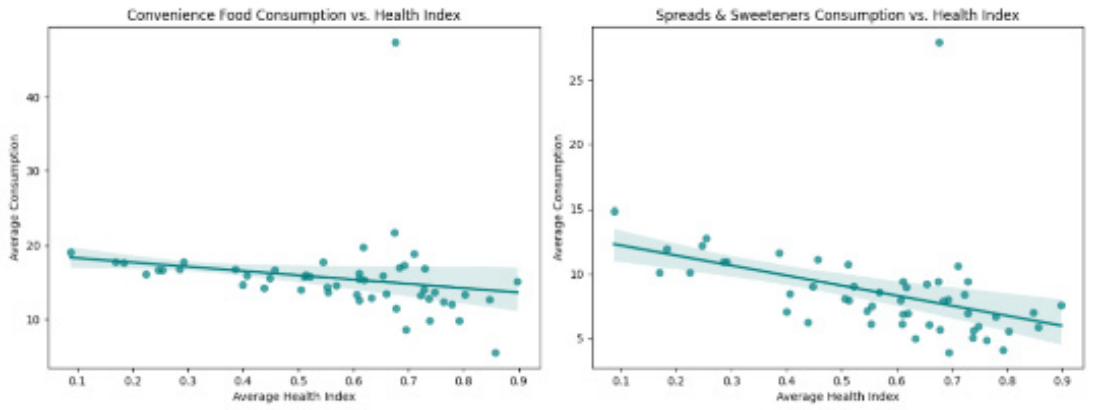


Fig. 6 - Scatter Plots of Consumption vs. Health Index

items such as fruits, vegetables, and seafood. In contrast, states with lower income levels consume more processed and convenience foods, likely due to affordability and availability issues. This variation in consumption patterns may be linked to regional health disparities, as states with healthier consumption habits tend to report better health outcomes.

2. Correlation Between Food Prices and Health Outcomes

The correlation analysis reveals significant relationships between food consumption categories, health outcomes, and socio-economic factors such as income and education. Key findings include:

Higher consumption of fruits, nuts, and vegetables is strongly associated with better health outcomes, including lower obesity rates, reduced heart disease prevalence, and longer life expectancy.

Processed and convenience foods show a negative correlation with health, with increased consumption linked to higher obesity and diabetes rates.

Income and education levels play a crucial role in shaping dietary habits and health outcomes, as wealthier states tend to have healthier grocery baskets and correspondingly better health indicators.

While a broad heatmap across all states shows mixed or weaker correlations, more distinct patterns emerge when analyzed on a state-by-state basis. A full breakdown of these correlations by state, along with detailed heatmaps, can be found in Appendix B.

3. Price Sensitivity and Consumption Patterns

Scatter plots illustrate the relationship between the price per unit (PPU) of various food products and their consumption levels across U.S. states. Generally, higher prices for healthier foods, such as fruits, nuts, and vegetables, are associated with lower consumption. Conversely, lower prices for convenience and processed foods often correspond with higher consumption levels, which may contribute to poorer health indicators (Fig. 6) . These trends highlight the role of affordability and purchasing power in shaping dietary choices, though variations exist across income levels and regions.

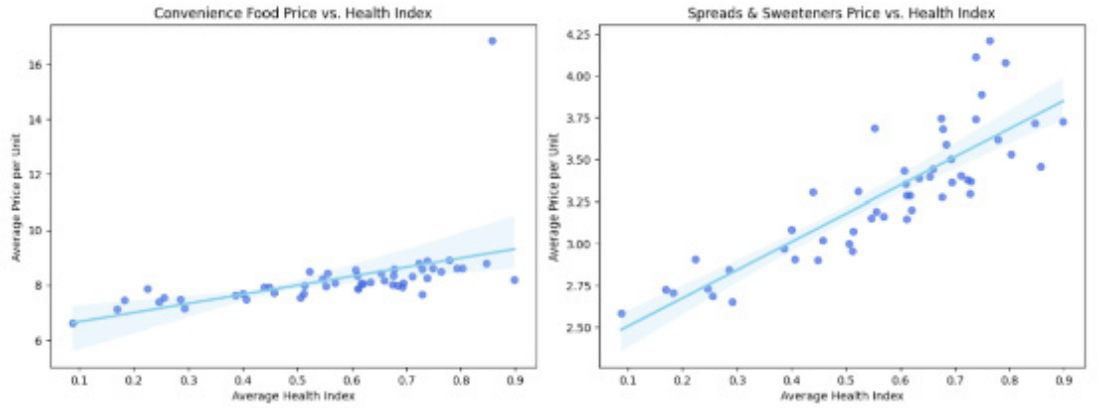


Fig. 7 - Scatter Plots of Price Per Unit vs. Health Index

Scatterplots show the relationship between spreads & sweeteners consumption and the Health Index, as well as the relationship between convenience foods consumption and the Health Index.

Scatterplots in this figure show the price per unit for spreads & sweeteners and convenience foods, and how these prices correlate with the Health Index. Additionally, the figure shows a scatter plot of the PPU of fruits and nuts and it's correlation with the Health Index.

The scatterplots in Figure 6 indicate that higher consumption of spreads & sweeteners and convenience foods is associated with a lower Health Index, suggesting that these food categories contribute to poorer health outcomes.

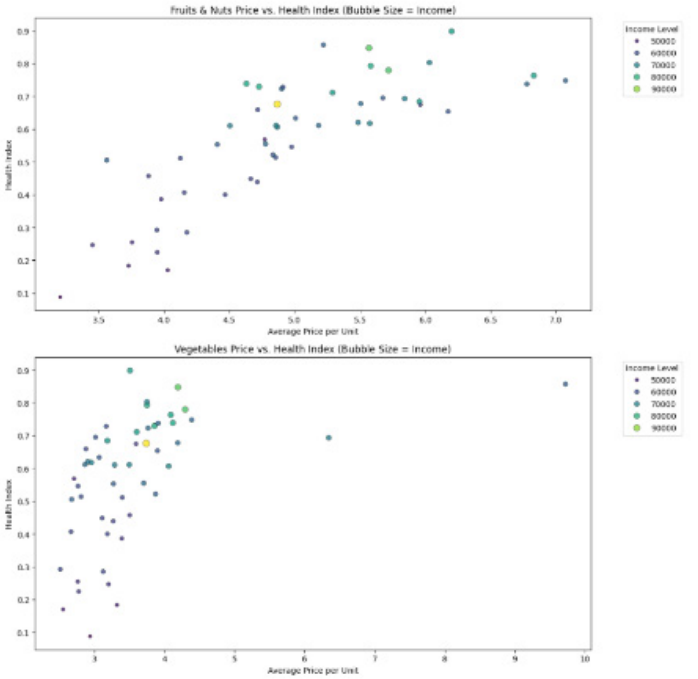


Fig. 8 - Scatterplots illustrate the relationship between the price per unit of fruits & nuts and vegetables with the Health Index, with bubble size representing state income levels.

The Health Index was constructed by combining four key health indicators: obesity rate, diabetes rate, heart disease rate, and life expectancy. Since higher values of obesity, diabetes, and heart disease indicate worse health outcomes, these variables were normalized and inverted, ensuring that a higher index value represents better overall health. Life expectancy, which directly correlates with good health, remained unchanged. The final Health Index is the average of these adjusted values.

Looking at Figure 7, we observe that higher prices for spreads & sweeteners and convenience foods correlate with a higher Health Index. This suggests that in states where these products are more expensive, people may consume them less, potentially leading to better overall health outcomes. This trend highlights the potential role of food pricing in shaping dietary choices and public health.

Our analysis suggests that in some states, higher prices for healthier foods (e.g., fruits, nuts, vegetables) are associated with better health outcomes. This trend is particularly evident in wealthier states, where higher income levels may allow greater access to nutritious foods and better healthcare. However, this relationship is not universal—some lower-income states with relatively high healthy food prices do not exhibit the same positive correlation with health. This suggests that other socioeconomic, policy, or accessibility factors may influence health outcomes beyond just food pricing and income levels.

For instance, Washington D.C., despite having high prices for healthy foods,[16] also boasts some of the lowest obesity rates.[17] It is worth noting however, that D.C. is an outlier in many regards. Meanwhile, in states like Arkansas and Louisiana, where incomes are lower, high prices for fresh produce may act as a barrier, reinforcing greater consumption of inexpensive processed foods.

4. Regression Analysis Results

Our multiple linear regression models explain about 97% of the variance in the Health Index across states (R-squared = 0.970, Adjusted R-squared = 0.957). Key findings from the regression analysis include:

Positive Correlations:

Higher consumption of fruits and vegetables tends to correlate with better health outcomes, though the strength of this relationship varies across states.

Income and education remain significant predictors, with wealthier and more educated populations generally exhibiting lower rates of obesity, diabetes, and heart disease. However, exceptions exist, particularly in states where other factors—such as healthcare access or food availability—moderate these effects.

Negative Correlations:

Processed and convenience foods, such as spreads and sweeteners, show negative correlations with health outcomes. States with higher consumption of these items tend to report worse health indices.

16 Statista. "Food Market in the United States." Accessed March 2, 2025. https://www.statista.com/outlook/cmo/food/united-states.

17 County Health Rankings & Roadmaps. "Adult Obesity." Last modified 2024. https://www.countyhealthrankings.org/health-data/health-factors/health-behaviors/diet-and-exercise/adult-obesity?year=2024.

Explanation: The regression results confirm the strong relationships between dietary patterns, socio-economic factors, and health outcomes. Income and education play crucial roles in shaping both dietary choices and health. The consumption of healthier foods like fruits and vegetables is associated with better health while the consumption of processed and convenience foods is detrimental to public health.

Ridge Regression:

To improve model stability while retaining all relevant predictors, we applied Ridge Regression with an optimal regularization parameter of $\alpha = 0.0572$. This relatively low penalty confirms that some features were redundant but not excessively so. The final model achieved a Mean Squared Error (MSE) of 0.00259, indicating that Ridge Regression effectively stabilized the estimates without sacrificing accuracy.

The Ridge Regression model reveals key factors influencing health outcomes across U.S. states. Higher consumption of Fish & Seafood (+0.151552), Fruits & Nuts (+0.056096), and Vegetables (+0.024587) is associated with better health outcomes, reinforcing the well-documented benefits of nutrient-rich foods. Additionally, higher education levels (+0.051752) and income (+0.021962) correlate positively with health, suggesting that socio-economic status plays a crucial role in dietary quality and overall well-being.

Conversely, greater intake of Convenience Food (-0.204704) and Spreads & Sweeteners (-0.177838) is linked to poorer health outcomes, reflecting the negative impact of processed and high-sugar diets. Lifestyle factors also play a role, with higher inactivity rates (-0.013634) and tobacco use

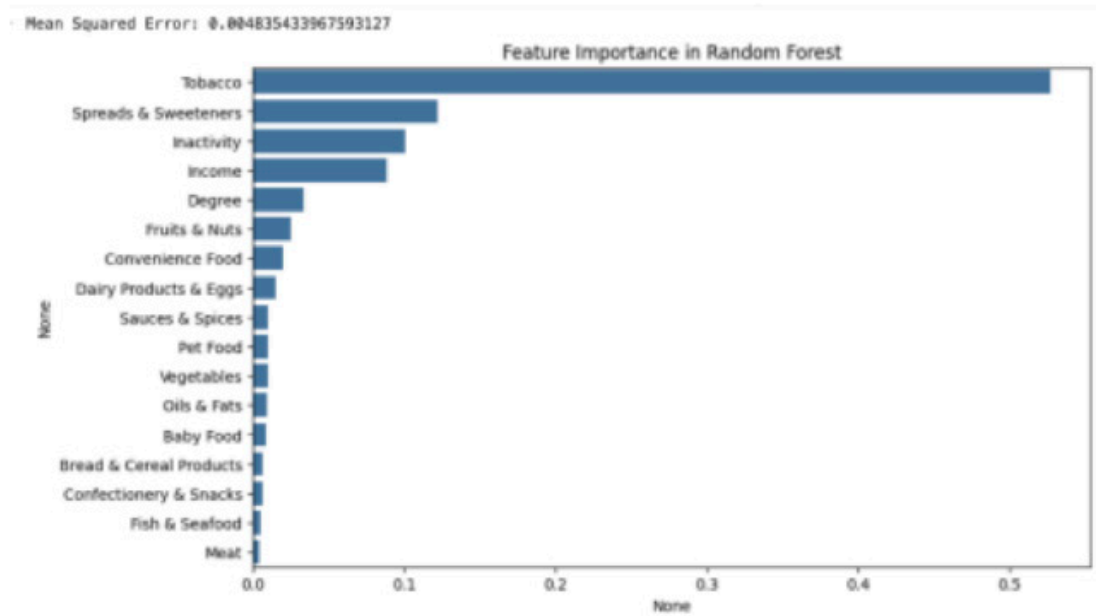


Fig. 9 - Random Forest MSE and feature importance without PPUs and ARPCs. Random forest feature importance analysis reveals the strongest predictors of health outcomes. Tobacco consumption is the most significant factor, followed by spreads & sweeteners, physical inactivity, income, and educational attainment.

(-0.025734) contributing to adverse health indicators.

Price sensitivity further shapes health outcomes. Higher prices for Fish & Seafood (+0.161793) and Fruits & Nuts (+0.068846) are associated with better health, suggesting that wealthier states continue to afford these nutritious options despite higher costs. In contrast, lower prices for Convenience Food (-0.020634) correlate with worse health, implying that the affordability of processed foods may encourage unhealthy consumption patterns.

Notably, Ridge Regression largely shrank the coefficients of ARPC (Average Revenue Per Capita) variables, indicating that they were redundant with PPU (Price Per Unit) and direct consumption values. This suggests that PPU and direct consumption levels provide more meaningful insights into dietary impacts on health than revenue-based spending measures.

These findings highlight the complex interplay between food pricing, consumption habits, and socio-economic factors in shaping public health. The full Ridge Regression coefficients are reported in Appendix C.

5. Random Forest Model Insights

Our random forest model further highlights the importance of food prices and consumption patterns in predicting health outcomes. The top food-related features influencing the Health Index include:

PPU Fruits & Nuts: The price per unit of fruits and nuts emerged as the most important factor, indicating that higher prices for these healthier items are strongly associated with better health outcomes in wealthier states.

PPU Spreads & Sweeteners: Higher prices for spreads and sweeteners are linked to better health outcomes, suggesting that when these items are pricier, they are consumed more judiciously, contributing to better health.

Explanation: The random forest model results suggest that price sensitivity plays a significant role in food consumption choices, particularly for healthier items like fruits and nuts. In states where these items are more expensive, wealthier populations can afford to consume them, leading to better health outcomes. On the other hand, states with lower food prices for less nutritious items may experience higher consumption of those products, which could contribute to increased public health challenges related to diet and nutrition.

Discussion

Our study offers valuable insights into the relationships between grocery consumption patterns, food prices, socio-economic factors, and health outcomes across the United States. The results show that consumption of healthier foods, such as fruits, vegetables, and seafood, is associated with better health outcomes, including lower rates of obesity, diabetes, and heart disease, as well as higher life expectancy. Conversely, the consumption of processed and convenience foods is often associated with higher rates of obesity and poor health outcomes, which reflects the negative impact of unhealthy dietary habits.

By applying Ridge Regression, we addressed redundancy among highly correlated variables while preserving the interpretability of the results. The model confirmed that food consumption patterns and pricing structures play significant roles in shaping public health, with clear distinctions between the impact of nutrient-rich versus processed foods. The findings reinforce the importance of both affordability and accessibility in determining dietary habits across socio-economic groups.

One of the key takeaways from our research is the strong role that socio-economic factors, such as income and education, play in determining dietary habits and, consequently, health outcomes. States with higher income levels and better-educated populations tend to have healthier grocery baskets, which in turn leads to better health. For instance, Vermont has one of the highest rates of organic food consumption, reflecting both higher incomes and a strong local food movement. Conversely, Kentucky and West Virginia, with lower median household incomes, see greater consumption of processed and high-calorie foods, contributing to their higher rates of obesity and heart disease.

These findings highlight the importance of addressing socio-economic disparities when designing public health interventions aimed at improving diet and reducing health inequalities. Wealthier states can afford to consume healthier, more expensive food options like fruits, nuts, and fish, whereas lower-income states may rely more heavily on cheaper, processed alternatives, contributing to adverse health effects. Food prices emerged as a significant factor in our analysis, though their impact on health outcomes varies by state. In some higher-income states, pricier healthy food options coincide with better health outcomes, likely due to greater affordability among wealthier populations. However, this pattern does not hold uniformly across all states, suggesting that additional factors—such as food accessibility and cultural dietary habits—also influence these relationships. Conversely, in lower-income states, where affordability of healthy options is a challenge, higher consumption of cheaper processed foods is more common, contributing to the public health issues observed in those areas. Our findings align with the existing literature, which emphasizes the importance of price sensitivity, education, and socio-economic status in shaping dietary behaviors. However, our analysis extends these insights by offering a more granular view of how these factors interact at the state level, providing policymakers with region-specific data to inform future interventions. Strategies like subsidies for healthier foods and educational campaigns promoting better dietary habits may be particularly effective in states with lower income levels and/or higher obesity rates.

Beyond economic interventions, additional non-economic strategies could further enhance food accessibility and improve public health outcomes:

Nutrition Education & Public Awareness Campaigns: Many dietary choices are influenced by knowledge gaps rather than price alone. Expanding school-based nutrition programs, community workshops, and digital health campaigns could help individuals make more informed food choices.

Urban Planning and Zoning Policies to

Reduce Food Deserts: Food accessibility remains a challenge in many lower-income and rural areas. Zoning policies that encourage grocery store development, farmers' markets, and mobile fresh food programs could help mitigate food deserts and improve access to healthier food options.

Public-Private Partnerships for Food Accessibility: Collaborations between governments, grocery retailers, non-profits, and local farms could improve the distribution of affordable, fresh food. Initiatives such as grocery store incentives, farm-to-table programs, and subsidized food delivery services could help bridge accessibility gaps.

Community-Led Initiatives: Grassroots programs such as urban gardening projects, community food cooperatives, and local farmers' markets empower residents to take control of their food supply and promote sustainable eating habits.

Integrating these non-economic strategies alongside financial incentives could help create a more comprehensive, long-term approach to reducing dietary disparities and improving public health. Future policy recommendations should consider a multi-faceted framework that combines economic, educational, and structural interventions to address food accessibility challenges at both the state and community level.

Limitations

Despite the robustness of our analysis, there are several limitations that must be considered when interpreting the results.
1. State-Level Aggregation: Our analysis is based on aggregate data at the state level, which may obscure important variations at the individual or community level. While this approach allows us to identify broad trends across states, it does not capture the full diversity of food consumption patterns or health outcomes that may exist within states. Future research could benefit from using more granular, individual-level data to explore these relationships in greater detail.

2. Causal Inference: While our multivariate regression and random forest models provide strong associations between grocery consumption and health outcomes, these methods cannot definitively establish causality. For example, it is possible that poor health outcomes lead to changes in dietary habits rather than the reverse. To better understand causal relationships, future research could employ more sophisticated econometric techniques, such as instrumental variable (IV) estimation or regression discontinuity (RD) design, which may help isolate the effect of specific variables on health outcomes.

3. Limited Timeframe: Our data covers only three years (2021-2023), which may not fully capture longer-term trends in grocery consumption or health outcomes. Additionally, the COVID-19 pandemic during this period could have influenced food purchasing behaviors, either temporarily or permanently, thus skewing the results. Extending the analysis over a longer time horizon would provide a more comprehensive view of the relationships we examined.

4. Food Price Data: Although we accounted for food prices in our analysis, the price per unit (PPU) data does not reflect regional differences in food accessibility or the impact of food deserts, which could

further explain variations in grocery consumption patterns and health outcomes. For instance, in some areas, higher-priced, healthier food items may be more readily available than in others. Incorporating data on food deserts and local food accessibility would enrich the analysis.

5. Control Variables: While we included important socio-economic variables like income, education, and tobacco use, other factors that influence health outcomes, such as access to healthcare, local government policies, and environmental factors, among others, were not incorporated into our models. These omitted variables may introduce bias into our estimates, making it important to interpret the results with caution.

6. Intra-State Variability: Our analysis is conducted at the state level, which means it does not capture intra-state disparities in grocery consumption and health outcomes. Large and socio-economically diverse states, such as California and Texas, encompass significant variations between urban and rural areas, as well as between high-income and low-income communities. Urban areas tend to have greater access to fresh produce and grocery stores, whereas rural regions may experience food deserts, leading to differences in dietary patterns. Additionally, cultural, political, and demographic differences within states could further influence food consumption and health outcomes in ways that are not captured in our aggregated data. Future research could explore county-level or metropolitan-area analyses to provide a more granular view of these relationships. Texas could be a good place to start as it demonstrates significant intra-state variability: affluent urban areas like Austin have high availability of organic and fresh foods, whereas rural regions and border towns experience more limited access, leading to differences in dietary patterns within the same state.

7. Omitted Variables and Potential Bias: While our study includes key socio-economic factors such as income, education, and food prices, it does not account for several other important variables that may influence dietary habits and health outcomes. Notably, we do not incorporate healthcare access, which plays a crucial role in determining health outcomes beyond dietary patterns. States with better healthcare infrastructure may report lower rates of chronic illnesses, independent of dietary habits, which could bias our results. Additionally, we do not include food environment indices, such as the proximity of grocery stores, farmers' markets, or food deserts. These factors significantly impact food accessibility, particularly in low-income and rural areas, where the availability of fresh and healthy food options may be limited. Lastly, cultural dietary preferences are not explicitly accounted for in our analysis. Regional and ethnic variations in diet, such as differences in staple foods or traditional cooking practices, could introduce additional complexity in the relationship between grocery cart composition and health outcomes. By not including these variables, our model may underestimate the full range of factors that shape food consumption patterns and public health disparities. Future research should aim to integrate these additional dimensions to provide a more holistic understanding of dietary behavior and its effects on health.

Future Work

Looking ahead, we plan to enhance our U.S.-focused analysis by incorporating additional socio-economic variables such as healthcare access and food environment indices. Expanding the dataset to include more granular geographic divisions (e.g., county-level data) could provide further insights into dietary disparities and health outcomes.

Additionally, while this study focuses on state-level patterns in the U.S., a future avenue of research could involve exploring similar analyses in international contexts. Comparing findings across different economic and cultural environments could provide broader insights into how food policies and economic factors shape dietary habits and health outcomes. However, such an extension would require careful methodological adjustments to ensure cross-country comparability.

In summary, while our study provides important insights into how food consumption patterns, prices, and socio-economic factors interact to influence health outcomes in the U.S., there remains much to explore in terms of causal relationships and regional nuances. Addressing these limitations in future work will enable a more comprehensive understanding of the complex links between diet, socio-economic status, and health, ultimately contributing to more effective public health strategies.

Conclusions

This study explored the relationship between grocery consumption patterns, food prices, socio-economic factors, and health outcomes across the United States from 2021 to 2023. Our analysis revealed that healthier food consumption, such as consumption of fruits, vegetables, and seafood, is associated with better health outcomes while higher consumption of processed and convenience foods correlates with worse health indicators, including obesity, diabetes, and heart disease.

However, beyond these statistical relationships, our study highlights structural inequalities that shape dietary habits. Socio-economic factors like income and education were found to be critical determinants of dietary habits and health, highlighting the disparities between wealthier and lower-income states. Overall, income has an indirect impact on health status, as it enables individuals to either afford healthier food options or resort to cheaper processed and convenience foods, depending on their financial situation. The fact that healthier foods tend to be consumed more in wealthier states raises equity concerns, sincem low-income populations face both economic and geographic barriers in accessing nutritious options. This underscores the need for policy interventions that go beyond pricing strategies, incorporating zoning laws, food accessibility programs, and community-led health initiatives to ensure that healthier foods are not only affordable but also available and culturally accessible.

While our analysis provides robust evidence of correlations, it does not establish causality, which remains a key limitation. Future research should explore longitudinal data or employ causal inference techniques (e.g., instrumental variables, natural experiments) to better isolate the effects of food pricing and consumption patterns on health outcomes. Additionally, more granular county-level or urban-rural comparisons could offer deeper insights into

intra-state disparities that our state-level approach could not fully capture.

Policymakers should consider a multi-faceted approach to addressing dietary disparities, integrating economic measures (e.g., implementing policies such as taxing demerit goods or establishing price ceilings on healthier foods in regions where they are under-consumed) with structural and educational strategies. For instance, food accessibility programs targeting rural food deserts, school-based nutrition initiatives, and urban planning policies promoting grocery store placement could create sustainable improvements in public health.

In conclusion, our research highlights the urgent need for a more comprehensive strategy in tackling dietary health disparities. While pricing policies play a role, they must be accompanied by accessibility initiatives, education programs, and infrastructure investments to create a food system that promotes health equity across all income levels and geographic regions. Future studies should continue refining these insights by exploring causal mechanisms and micro-level dietary behaviors, further strengthening the evidence base for policy reform.

***

for acknowledgments and appendixes
please visit www.tulanejournal.com

***

Sources of Discontent and Distrust with the Chinese Communist Party

Sam Borne
Bowdoin College

Introduction

The Chinese Communist Party (CCP) is popular. As repeated surveys have shown, satisfaction with the CCP began high in the early 2000s and has remained consistently high through the late 2010s.[1] However, when the 2000s began, almost no one thought the Chinese Communist Party would still be in power nearly a quarter century later. American President Bill Clinton, in 2000, argued that when China joined the World Trade Organization, it would spark a political revolution. "The more China liberalizes its economy, the more fully it will liberate the potential of its people – their initiative, their imagination, their remarkable spirit of enterprise."[2] Clinton, academics, and the business community believed that once the nation began to liberalize economically, political liberalization would follow as surely as summer follows spring. However, as time has proven, they were wrong.

Andrew Nathan, a political scientist who studies China, concluded that "China specialists and democracy theorists—myself among them—expected the regime to fall to democratization … this particular authoritarian system, however, has proven resilient."[3] In 2018, The New York Times published a series of articles titled "What the West Got Wrong About China" and reflected on the belief that political liberalization would follow economic liberalization just as the moon follows the sun. The New York Times concluded, "the West was sure the Chinese approach would not work. It just had to wait. It's still waiting."[4] This point is worth repeating to emphasize how wrong the West was. The West believed, beyond a shadow of a doubt, that China's economic and political system would be doomed to inevitable failure. However, China has not just grown; it has prospered. The West believed economic liberalization would turn China into a trading partner, creating a rising tide that would lift all boats and an inseparable bond between the two nations. Peaceful cooperation and coexistence existed for some time, and then a conflagration began, which started a trade war.[5] However harmful a trade war might

1 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.;Edelman. "Edelman Trust Barometer." 2024. https://www.edelman.com/sites/g/files/aatuss191/files/2024-01/2024%20Edelman%20Trust%20Barometer%20Global%20Report_0.pdf.; Li, Lianjiang. "Assessing Popular Support for the Communist Party of China." China: An International Journal 21, no. 2 (2023): 51–71. Project MUSE. https://doi.org/10.1353/chn.2023.a898341.

2 Clinton, Bill. "Remarks to Paul H. Nitze School of Advanced International Studies of the Johns Hopkins University." Transcribed by the Federal News Service, March 9, 2000. https://www.nytimes.com/2000/03/09/world/clinton-s-words-on-china-trade-is-the-smart-thing.html.

3 Nathan, Andrew J. "China's Changing of the Guard: Authoritarian Resilience." Journal of Democracy 14, no. 1 (2003): 6–17. https://doi.org/10.1353/jod.2003.0019.

4 Pan, Phillip. "What the West Got Wrong About China, The Land That Failed." New York Times, November 18, 2018. https://www.nytimes.com/interactive/2018/11/18/world/asia/china-rules.html.

5 Siripurapu, Anshu, and Noah Berman. "The Contentious U.S.-China Trade Relationship." Council on Foreign Relations, 2024. https://www.cfr.org/backgrounder/contentious-us-china-trade-relationship.

be, a real war between the US and China, two superpowers, would be worse. While a real war seems unlikely, the probability has risen alarmingly and dramatically in recent years, prompting the US government to plan for such an eventuality.[6] When Bill Clinton welcomed China into the World Trade Organization in 2000, few even imagined a mere quarter century later, a war might be on the horizon. So where did Clinton and others go wrong? To answer that question, one must first ask a more fundamental question: what makes a citizen trust or distrust the government?

This paper rigorously and statistically examines empirical evidence and political theory to illuminate the sources that cause trust and distrust within the national government of the Chinese Communist Party. With an improved understanding of these basic citizen and state relations, political scientists can better understand the Chinese Communist Party's resilience and potential vulnerabilities.

Background: Political and Economic Theory of the CCP's Reliance

Performance Legitimacy:

From 2003 to 2016, the Ash Center for Democratic Governance and Innovation conducted regular surveys to study citizens' perceptions of the Chinese government. The survey asked respondents to rate their satisfaction with the government on a scale from 1 to 4: one representing "very dissatisfied," two "dissatisfied," three, "satisfied," and four, "very satisfied."[7] Beginning in 2003, the average satisfaction with the Central Government stood at 3.16, meaning the average respondent reported being slightly above satisfied, and that number rose to 3.3 by 2016. The report concluded that "across a wide variety of metrics, by 2016, the Chinese government was more popular than at any point during the previous two decades."[8] The researchers believe that increased satisfaction rates with the government occurred because the government helped generate a broad-based increase in the quality of life.[9] The government provides for its people; therefore, the people support the government. This theory that the government derives its authority and legitimacy from its ability to provide economic growth and social stability is known as performance legacy.

The theory of performance legitimacy originated around the rise of the Western Zhou Dynasty in 1050 BCE, which used performance legitimacy to justify their claim to rule.[10] Performance legitimacy is a modification of the Mandate of Heaven, a system where the government rules as long as it can provide peace and economic growth. If the government fails to uphold its end of the bargain due to corruption, invasion, or natural disasters, the government loses its mandate to rule, and a new government usurps its place.

6 Beckley, Michael. "How Primed for War Is China?" Foreign Policy, February 4, 2024. https://foreignpolicy.com/2024/02/04/china-war-military-taiwan-us-asia-xi-escalation-crisis/.
; Hirsh, Michael. "The Pentagon Is Freaking Out About a Potential War With China." Politico, June 9, 2023. https://www.politico.com/news/magazine/2023/06/09/america-weapons-china-00100373.

7 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.

8 Ibid.

9 Ibid.

10 Zhao, Dingxin. "The Mandate of Heaven and Performance Legitimation in Historical and Contemporary China." American Behavioral Scientist 53, no. 3 (2009): 416–433. https://doi.org/10.1177/0002764209338800.

The CCP has overseen a period of profound economic growth over the past half-century. The Chinese economy's real annual GDP has grown from 1979 to 2018 at an average yearly rate of 9.5%, meaning the economy doubled in size about every eight years.[11] This profound economic growth has lifted 800 million people out of poverty while the per capita GDP has risen by 6000% since 1978.[12] As the Ash Center explains, transformative economic growth means "that the government's provision of healthcare, welfare, and other essential public services was far better and more equitable."[13] The Chinese Communist Party gained its people's approval by overseeing profound economic growth and improving its citizens' quality of life.

Performance legitimacy offers a chance for the CCP to ground its mandate to rule in tangible and quantifiable goals such as economic growth and social stability.[14] In the early days of Mao, the CCP claimed its legitimacy came from its communist ideology and Marxism-Leninism-Mao Zedong Thought. However, under Deng Xiaoping, the economy began to reform and liberalize, abandoning the fanatical ideological belief in the sort of communism that governed the nation under Mao. Due to Deng's reforms, the government could no longer claim the moral righteousness of its political ideology as its basis for legitimacy. Similarly, in an atheist state, the divine rights of rulers or any other religious model remained impossible. Legitimacy based on religion would only make the government appear utterly hypocritical after years of heavy suppression and persecution of religious communities. Consent of the governed, which provided the basis for legitimacy in democratic nations, remained a laughable notion, as the CCP had no intent of distributing its power among the people. Without better alternatives, Deng and others in the post-Mao era reoriented the CCP's claim to legitimacy from ideology to its economic performance.[15] Deng assumed that a government that could improve its people's economic situation would increase its citizens' satisfaction and happiness, which would lead to increased levels of trust in the government. Scholars examining data from the World Value Survey have confirmed this, finding that increases in life satisfaction and happiness are positive predictors of a person's trust in the central government.[16] Meanwhile, to "promote regime legitimacy," scholars have concluded the government ought to attempt to crack down on corruption while reducing poverty through improvement in the material economic conditions of its citizens.[17]

The CCP's ability to provide social stability and basic social services is a critical tenet of its legitimacy. Increases in spending on health care, education, and social welfare correlate with increased trust in the central government.[18] The more the government spends on public goods that improve

11 Morrison, Wayne M. China's Economic Rise: History, Trends, Challenges, and Implications for the United States. Congressional Research Service, RL3353, 2019. https://crsreports.congress.gov.

12 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.

13 Ibid.

14 Zhu, Yuchao. "'Performance Legitimacy' and China's Political Adaptation Strategy." Journal of Chinese Political Science16, no. 2 (2011): 123–140. ProQuest.

15 Ibid.

16 Hu, Dora, and Teresa Wright. "Developments in China's Public Opinion from Hu to Xi: Corruption, Activism and Regime Legitimacy." The China Quarterly 256 (2023): 1068–1084.

17 Ibid.

18 Dickson, Bruce J., Pierre F. Landry, Mingming Shen, and Jie Yan. "Public Goods and Regime Support in Urban China." The China Quarterly 228 (2016): 859–880.

citizens' quality of life, the more citizens trust the government. Furthermore, providing public goods may help build regime support, even when citizens are dissatisfied with local officials. Studies have shown that improving the quality of public service can compensate for dissatisfaction with officials due to corruption or disagreements over policy goals and even overcome this dissatisfaction.[19] Scholars have also argued that the Chinese government must provide these basic services to its citizens precisely because its legitimacy depends on providing these goods. The Chinese government must act responsibly to meet the demands and desires of the public, for should it fail to do so, it would lose its mandate to govern.[20] However, the Chinese government has embraced a double-sided sword by staking its legitimacy on its performance. As long as the economy continues to grow and the people feel safe, satisfied, and content, the government will remain legitimate in the eyes of its people. However, should the government fail to provide economic growth and social stability, it would lose its current justification for existing, potentially endangering the regime. Nevertheless, performance legitimacy seems to best explain the current high levels of support for the government, given the massive economic growth throughout the past half-century.

The Difficulties of Performance Legitimacy: A Double-Edged Sword

Performance Legitimacy assumes that the government can provide social stability and economic growth. However, should the government fail to provide for its citizens, they may begin to blame the government.[21] Current predictions estimate that by 2035, the Chinese economy will slow down to a growth rate of around 2%.[22] China has grown tremendously as the economy has developed from an agricultural to a manufacturing society. However, despite China's remarkable growth, it can not defy the laws of economic gravity, and its growth rate will inevitably slow.[23] Slowing economic growth means that should an economic crisis occur, the effects will be more pronounced. If the economy doubles in size every eight years, with a 9% growth rate, a 10% recession in terms of real GDP would all but disappear in a year. However, if the economy grew slower, say 2%, and took approximately 35 years to double, the impact of a 10% recession on real GDP would take years to recover from.[24] Economic growth also comes with downsides, such as changing social values, increasing corruption, economic inequality, and environmental degradation.[25] Scholars believe that as China grows wealthier, it will become more likely to encounter these challenges. However, determining the

19 Ibid.

20 Yang, Heng, and Dingxin Zhao. "Performance Legitimacy, State Autonomy and China's Economic Miracle." Journal of Contemporary China 24, no. 91 (2015): 64–82. https://doi.org/10.1080/10670564.2014.918403.

21 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.;Zhao, Dingxin. "The Mandate of Heaven and Performance Legitimation in Historical and Contemporary China." American Behavioral Scientist 53, no. 3 (2009): 416–433. https://doi.org/10.1177/0002764209338800.

22 Morrison, Wayne M. China's Economic Rise: History, Trends, Challenges, and Implications for the United States. Congressional Research Service, RL3353, 2019. https://crsreports.congress.gov.

23 Ibid.

24 Wise, Carol. "China and Latin America's Emerging Economies: New Realities Amid Old Challenges." Latin American Policy 7 (2016): 26–51. https://doi.org/10.1111/lamp.12087.; Zheng, Zhaohui. "The Concept and Essence of Middle-Income Trap." In Middle-Income Trap. Singapore: Palgrave Macmillan, 2020. https://doi.org/10.1007/978-981-15-7401-6_1.

25 Zeng, Jinghan. "Introduction." In The Chinese Communist Party's Capacity to Rule, Critical Studies of the Asia-Pacific. London: Palgrave Macmillan, 2016. https://doi.org/10.1007/978-1-137-53368-5_1.

impact of economic growth on citizens' stratification can be difficult.

Performance legitimacy constitutes a double-edged sword, for it provides the government with an easy way to gain legitimacy, but by basing their CCP's legitimacy on its ability to provide economic growth, it rejects alternative bases of legitimacy such as the divine right of the kings. Critically, by embracing performance legitimacy the CCP has created a political system where as long as the economy grows, it can remain legitimate, however, should it fail to sustain robust economic growth, its legitimacy will begin to waver.

Embracing Performance Legitimacy Nevertheless

While many scholars contend that performance legitimacy explains the CCP's continued and even growth in support, this belief is far from universal. Some scholars contend that an apparent causal link exists between economic growth and happiness; when a person's economic situation improves, so does their happiness.[26] As happiness increases, citizens trust the national government more because they created the economic situation. However, as Heike Holbig and Bruce Gilley argue, "the relationship between growth and regime legitimacy is not an obvious one." Holbig and Gilley refute the assumption that increased real income leads to increased happiness because economic satisfaction constitutes a "highly abstract notion."[27] Holbig and Gilley contend that perceived well-being is not an objective measure but rather a subjective one. Consider the following hypothetical: if the average personal income of an entire village constituted 10 million USD a year, and one villager made only 1 million USD per year, that villager would be poor relative to that economic situation. However, should the villager move to a village where the average personal income of an entire village stands at $10,000, the original village would be rich relative to that situation. The hypothetical example reveals that feelings of rich or poor depend on the relative "intertemporal, interpersonal, interregional, and international comparison[s]."[28] Holbig and Gilley do not dispute the claim that economic growth can lead to happiness but introduce an intermediate step, where economic growth must cause citizens to perceive positive economic changes in order to lead to increased happiness.

The slowdown in the economic growth rate reveals a more nuanced picture of Chinese legitimacy than simply performance legitimacy. After the Great Recession of 2008, China's economic growth rates slowed down from regular double digits to around 6-8% per year, but instead of a crisis of legitimacy, the Chinese government plodded along.[29] Scholars had previously claimed that a slowdown in economic growth could trigger a legitimacy crisis in the government.[30] However, that did not happen, and by some accounts, trust in the government

26 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.

27 Holbig, Heike, and Bruce Gilley. "Reclaiming Legitimacy in China." Politics & Policy 38, no. 3 (2010): 395+. Gale Academic OneFile. https://link.gale.com/apps/doc/A233048048/AONE?u=brun62796&sid=bookmark-AONE&xid=1ca64f53.

28 Ibid.

29 Morrison, Wayne M. China's Economic Rise: History, Trends, Challenges, and Implications for the United States. Congressional Research Service, RL3353, 2019. https://crsreports.congress.gov.

30 Zhao, Dingxin. "The Mandate of Heaven and Performance Legitimation in Historical and Contemporary China." American Behavioral Scientist 53, no. 3 (2009): 416–433. https://doi.org/10.1177/0002764209338800.

even increased after the financial crisis.[31]

Scholars argue that framing the economic circumstances impacts perceptions of the economy. As Heike Holbig in a later article argues, economic legitimacy and growth "can be experienced only through linguistic constructions of social reality, or more precisely, by framings – schemata of interpretation that help us to navigate our experiential universe, whether in the form of statistics, policy briefs, news, propaganda slogans, films or hearsay."[32] Put in another way, economic "performance is not a tangible thing, despite the associated evidence of money, rising living standards and material wealth."[33] Holbig argues that the government's ability to convince its people of its economic growth matters more than tangible economic growth. In other words, what people believe matters more than the factual basis of reality. Holbig argues that in some ways, perceptions of reality matter more than the reality itself.

Scholars have also argued that low expectations for the regime allow its citizens to treat its performance as satisfactory. Literature suggests that in some cases, such as healthcare, extraordinarily low explications mean that citizens are satisfied if the government provides almost any service.[34] Meanwhile, some scholars have argued that corruption campaigns increase citizens' trust as the government takes visible actions to improve the nation. However, initiating a corruption investigation might reveal and confirm to the public the existence of widespread corruption.[35] Yuha Wang, and Bruce Dickon, write that when it came to fighting corruption, no easy solution exists. "Our findings reveal a bitter irony: corruption negatively impacts regime support, but fighting corruption is no panacea. Regarding public opinion, the cure may be as bad as the disease.[36]

While at first, performance legitimacy may offer a straightforward explanation for what causes regime support, under closer examination, the nuances and complexities reveal themselves. However, all theories contain nuance, and performance legitimacy has value due to its compelling explanation for the continuance in the high levels of trust with the CCP. Scholars create models to simplify the real world and examine the impact of variables that might be impossible to isolate and study on a larger scale. While performance legitimacy has faults as a theory, nevertheless, it is worth examining as a theoretical framework to help explain the macro-level trends in the consistently high levels of trust in the Chinese Communist Party, and the reason for the few who distrust the government.

Hypothesis: The theory of performance legitimacy explains trust and distrust in the Chinese Communist Party. Therefore, when individuals believe the government has failed to provide economic growth or

31 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.

32 Holbig, Heike. "Whose New Normal? Framing the Economic Slowdown Under Xi Jinping." Journal of Chinese Political Science 23 (2018): 341–363. https://doi.org/10.1007/s11366-018-9539-6.

33 Ibid.

34 Ratigan, Kirsten. "Riding the Tiger of Performance Legitimacy? Chinese Villagers' Satisfaction with State Healthcare Provision." International Political Science Review 43, no. 2 (2022): 259–278. https://doi.org/10.1177/0192512120927482.

35 Gong, Ting, and Wenyan Tu. "Fighting Corruption in China: Trajectory, Dynamics, and Impact." China Review 22, no. 2 (2022): 1–19. ProQuest.

36 Wang, Yuhua, and Bruce J. Dickson. "How Corruption Investigations Undermine Regime Support: Evidence from China." Political Science Research and Methods 10, no. 1 (2022): 33–48.

social stability they become more likely to distrust the government.
Hypothesis 1a: Perceived poor economic conditions lead to distrust of the national government.
Hypothesis 1b: Perceived widespread corruption in the national government leads to distrust in the national government.
Hypothesis 1c: Lack of access to healthcare leads to distrust in the national government.

Research Design

Introduction:

Given high levels of trust in the CCP, what makes citizens distrust the government? To study this question, I used surveys from the Asian Barometer, which regularly surveys Asian nation's citizens about their trust in government, institutions, traditionalism, ideological beliefs, and other various socioeconomic topics. The Asian Barometer began in the early 2000s, with the first wave occurring in China from March 2002 to June 2002, and surveyed approximately 3000 randomly selected individuals. In the later waves, responses fluctuated between 3500 and 5000 unique participants. The total number of participants counts the number of people who participated in the survey and does not include those randomly selected and those who declined to participate. Those who chose to participate did so voluntarily, without any reward, and after being informed that all responses were anonymized to prevent identification. Wave two occurred during 2008, wave three during 2011, wave four from 2014-2016, and wave five during 2019. Results from wave six have yet to be released at the time of this paper.

Dependent Variable: Trust in the National Government:

For this paper, I analyzed the impact of economic dissatisfaction, perceived levels of corruption at the national level, and the ability to access health care on trust in the national government. Participants rated their trust in the national government on a scale from one to four, with one being "A Great Deal of Trust'', two as "Quite a Lot of Trust'', three as "Not Very Much Trust," and four as "None at All''. Although not prompted, the survey also recorded if participants did not understand the question, could not choose, or declined to answer. I then recoded the responses into a dummy variable, with trust responses recoded as 0, and do not trust responses as 1. I also recoded the nonresponse options as 1, standing for distrust of the government. I did this as I assumed that if citizens trusted the government, they would feel comfortable stating this. However, if a citizen did not feel comfortable, they may attempt to simply evade, or pretend not to comprehend the question in order to avoid answering the question. During the second wave, the trust of national government questions was in reverse order, with one standing for "None at All'', and four as "A Great Deal of Trust''. In the fifth wave, the options for trust in the government were expanded to a scale from 1 to 6, with 1 representing "Trust fully" and 6 representing "distrust fully."

Independent Variables:

Economic Satisfaction:

In order to measure economic satisfaction, respondents were asked, "How would you rate the overall economic condition

of our country today?" and then asked to choose an option ranging from 1 to 5, with 1 being "very good", 2 as "good", 3 as "neither good nor bad", 4 as "bad" and 5 "very bad". The surveyors also recorded if the respondents did not provide an answer due to an inability to choose, understand the question, or declined to answer. Those variables were classified as missing and not included in the analysis.

Corruption: Perceived Levels of in National Government:

In order to measure perceived levels of corruption, respondents were asked, "How widespread do you think corruption and bribe-taking are in the national government in the capital city" and then asked to choose an option on a scale from 1 to 4, with 1 being "Hardly anyone is involved" 2 as "Not a lot of officials are corrupt" 3 as "Most officials are corrupt" and 4 as "Almost everyone". When respondents failed to answer due to an inability to choose, understand the question, or declined to answer, those responses were classified as missing and not included in the analysis.

Healthcare Access:

In order to measure access to healthcare, respondents were asked, "How easy is it to receive medical treatment at a nearby clinic?" No question relating to healthcare access was asked in the first wave. In waves two and three, respondents were given the options, ranging from 1-4, with 1 as "very difficult," 2 as "difficult," 3 as "easy," and 4 as "very easy." In waves four and five, the question was modified to ask, "As far as you know, how easy or difficult is it to obtain the following services?" with healthcare access in a list of services such as roads, running water, help from the police, and internet access. The order of the prompts was also reversed, with 1 as "very easy" and 4 as "very difficult". Similar to the previous questions, when respondents failed to answer, due to an ability to choose, understand the question, or declined to answer, those variables were classified as missing and not included in the analysis.

Controls in Regression:

For this experiment, I used five separate demographic controls, gender, age, education level, income level, and subjective social status, to remove variance in the data that could occur due to different demographic factors. I decided to control for gender, age, and subjective social status as men, older citizens, and those with higher subjective social status have been shown to trust the state more than women, younger citizens, and those with lower subjective social status.[37] Meanwhile, I also chose to control education, as those with higher levels of education had less trust than those with less education.[38] I also controlled for income level, as income impacts the ability to access education and subjective social status.

I also used three other variables, the trustworthiness of strangers, perceived levels of crime, and deference to government leaders, to control for other factors impacting trust/distrust in the government. I controlled for crime as it has previously been shown to cause distrust in the government.[39]Scholars have also found that communi-

37 Heurlin, Christopher. "NGO Governance and Management in China." In NGO Governance and Management in China, edited by Reza Hasmath and Jennifer Hsu. Routledge, 2016.

38 Ibid.

39 Sun, I.Y., Z. Han, Y. Wu, and R. Guo. "Trust in the Police in Rural China: A Comparison Between Villagers

ties with higher levels of social trust within networks and relationships can impact the perception of the national government, leading me to control the perceived trustworthiness of strangers.[40] Others have found that those who have an "authoritarian bias," meaning a tendency to trust government figures, are unsurprisingly more likely to trust the government, leading me to control for deference to government leaders as a proxy.[41] These studies led to control for the selected variables and then ran a linear regression, using the data and these select controls to examine the impact of the independent variables on trust in the national government.

Discussion

As the descriptive statistics reveal, the number of individuals distrusting the national government throughout the various waves has remained consistently small. In the first wave in 2002, only 1.2% of respondents reported distrust of the national government, while over 93% trusted the government, and approximately 5% did not answer. By the time the fifth wave occurred in 2019, those numbers had barely changed, with 1.4% of citizens reporting distrust in the government, over 94% of citizens trusted the government, and only 4% did not respond. The nonresponse numbers barely impacted the results, given that over 85% of citizens reported trusting the government throughout each wave.

However, while the number of citizens trusting the government has remained constant, the strength of the trust has

and Local Officials." Asian Journal of Criminology 14 (2019): 241–258. https://doi.org/10.1007/s11417-019-09290-3.

40 Zhu, Ling, and Dongmin Kong. "Corruption and Privatization: Evidence from a Natural Experiment in China." Economics of Transition and Institutional Change 31, no. 1 (2023): 217–239. https://doi.org/10.1111/ecot.12331.

41 Li, Hui. "Chinese Students' Perceptions of the 'Good Citizen': Obedience to an Authoritarian Regime." Citizenship Teaching & Learning 13, no. 2 (2018): 177.

Sources of Distrust in the National Government

	Wave One	Wave Two	Wave Three	Wave Four	Wave Five
Gender	.726	.041	-.295	.063	.398
	(.557)	(.776)	(.383)	(.230)	(.314)
Age	.019	.023	.003	-.006	.011
	(.022)	(.032)	(.014)	(.009)	(.011)
Education	-.222*	.095	.118	-.063	.003
	(.120)	(.224)	(.103)	(.050)	(.065)
Income	.180	-.200	-.029	.045	-.145
	(.195)	(.314)	(.100)	(.088)	(.142)
Subjective Social Status	-.104	-.408	-.159	.028	-.053
	(.278)	(.186)	(.149)	(.057)	(.099)
Overall Economic Condition	.721***	.526**	.760***	.552***	.121
	(.260)	(.409)	(.187)	(.109)	(.189)
Trustworthiness of Strangers	-.294	.429	.137	.422*	.086
	(.564)	(.831)	(.400)	(.251)	(.320)
Access to Healthcare		-.481	-.524**	.585***	.420*
		(.710)	(.263)	(.155)	(.233)
Level of Crime	.605**		.393	.309*	.051
	(.269)		(.287)	(.179)	(.337)
Perceived National Corruption	1.039 ***	1.855***	1.532***	.596***	1.607***
	(.410)	(.545)	(.247)	(.202)	(.292)
Deference to Government Leaders	-.672	-.986	.659**	.314***	.132
	(.459)	(.849)	(.276)	(.081)	(.085)

***p<0.01, **p<0.05, *p <0.1

Table One: Regression Results, Dependent Variable: Trust in Government

Trust in National Government

	Wave One	Wave Two	Wave Three	Wave Four	Wave Five
None At All	0.3%	0.1%	0.5%	0.4%	0.2%
Not Very Much Trust	0.9%	0.8%	2.4%	3.8%	1.4%
Quite a Lot of Trust	5.8%	27.0%	42.1%	51.2%	48.4%
A Great Deal of Trust	87.3%	65.0%	51%	35.4%	45.4%
Nonresponse	5.6%	7.1%	4.9%	9.2%	4.6%
Total Number	3183	5098	3473	4068	4961

Table Two: Trust in National Government

Economic Satisfaction

	Wave One	Wave Two	Wave Three	Wave Four	Wave Five
Very Bad	0.5%	0.2%	0.6%	1.8%	0.3%
Bad	9.8%	3.8%	7.1%	12.6%	5.7%
Neither Good nor Bad	11.5%	8.4%	14.6%	19.8%	16%
Good	51.2%	57.7%	52.3%	40.3%	57.9%
Very Good	21.3%	20.6%	21.3%	16.4%	18.8%
Nonresponse	5.7%	9.4%	4.2%	9.1%	5.3%
Total Number of Responses	3183	5098	3473	4068	4961

Table Three: Overall Economic Condition

Perceived Corruption in the National Government

	Wave One	Wave Two	Wave Three	Wave Four	Wave Five
Almost everyone is corrupt	0.3%	1.1%	3.2%	1.0%	0.3%
Most officials are corrupt	6.7%	4.5%	13.0%	6.7%	2.2%
Not a lot of officials are corrupt	28.7%	17.5%	33.4%	58.8%	56.8%
Hardly anyone is corrupt	9.2%	16.0%	16.3%	6.3%	12.1%
Nonresponse	55%	60.8%	34.8%	27.2%	28.6%
Total Number	3183	5098	3473	4068	4961

Table Four: Perceived National Corruption

Access to Healthcare

	Wave One	Wave Two	Wave Three	Wave Four	Wave Five
Very Difficult		1.3%	3.6%	13.4%	2.0%
Difficult		8.4%	15.8%	60.2%	13.8%
Easy		70.9%	68.6%	17.6%	62.2%
Very Easy		13.6%	9.2%	4.6%	16.0%
Nonresponse		5.7%	2.8%	4.2%	6.0%
Total Number		5098	3473	4068	4961

Table Five: Access to Healthcare

weakened significantly. In the first wave in 2002, more than 87% of respondents said they had "a great deal of trust" in the government, and only 6% reported having "quite a lot of trust" in the government. However, by the fifth wave in 2019, only 45% of respondents reported having "a great deal of trust," and 48% reported "quite a lot of trust" in the government. Throughout the waves, trust has remained uniformly high, but the percentage of citizens with a "great deal of trust" has declined almost by half, while the amount with "quite a lot of trust" has grown by more than 900%. This change suggests that citizens still trust the government, but their degree of trust has weakened significantly. Ultimately, this factor did not play a role in the regression, as the responses were coded as either trusting or distrusting the government. Nevertheless, the decrease in the strength of the trust and its causes remains an interesting area for future research.

The high levels of trust in the national government match the findings of previous studies examining trust in China.[42] Similar to other studies, the total number of citizens reporting distrust of the government remains small when compared with the broad-based trust in the government. Nevertheless, in the regression, the total quantity of respondents who answered with distrust of the government never numbered above 200 in any of the waves.

The regression results show that demographic variables had almost no significant impact on citizens' perceptions of the national government. While previous studies have concluded that demographic factors, such as age, gender, education, income, and subjective social status, can impact a person's perception of the government, I found no such relationship in my regression analysis.[43] The sole expectation of this is the first wave, where higher education levels had a somewhat significant adverse effect on trust in the government.

Economic Satisfaction:

The results show that perceived economic conditions significantly impact trust in the national government. In waves one, three, and four, the confidence levels were at 99%, and perceived poor economic conditions caused distrust in the national government. In wave two, the confidence level was somewhat lower at 95%, and in the fifth wave, the results were insignificant. No obvious explanation exists for the drop in significance in the second wave and then lack of significance in the fifth wave. Nevertheless, together, the waves reveal that citizens who have negative perceptions of the economy are much more likely to distrust the national government over time.

As previously mentioned, China has grounded its legitimacy through its claim to be able to provide economic and social stability for its citizens.[44] However, when

42 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.;Edelman. "Edelman Trust Barometer." 2024. https://www.edelman.com/sites/g/files/aatuss191/files/2024-01/2024%20Edelman%20Trust%20Barometer%20Global%20Report_0.pdf.; Li, Lianjiang. "Assessing Popular Support for the Communist Party of China." China: An International Journal 21, no. 2 (2023): 51–71. Project MUSE. https://doi.org/10.1353/chn.2023.a898341.

43 Heurlin, Christopher. "NGO Governance and Management in China." In NGO Governance and Management in China, edited by Reza Hasmath and Jennifer Hsu. Routledge, 2016.

44 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.; Zhao, Dingxin. "The Mandate of Heaven and Performance Legitimation in Historical and Contemporary China." American Behavioral Scientist 53, no. 3 (2009): 416–433. https://doi.org/10.1177/0002764209338800.; Zeng, Jinghan. "Introduc-

the government cannot provide economic stability, it has failed at one of the central tenets of its claim to rule. The Chinese people have made a contract with the government, where the people agree to obey the government as long as it can provide economic stability and growth. As Sun Yu of the Financial Times explains, recently, as the economy in China has worsened, the social contract has come under increasing scrutiny. "The Communist party used to allow its people abundant economic opportunity in exchange for heavy restrictions on their political freedom. Now, the so-called social contract is no longer clear," leaving its citizens unsure about the future.[45] Should the Chinese Communist Party prove unable to reverse the present economic downturn, it might mark the end of the political system and the CCP's rule.[46]

Nevertheless, previous scholarly literature and the regression results all support the present hypothesis that poor economic conditions lead to increased distrust of the CCP. The causal mechanism for the distrust is the implicit social contract where the government promises its citizens economic security in exchange for their support. So far, the CCP has remained in power, and when the economy begins to sour, individuals feel distrust towards the government because they have failed to uphold their end of the bargain – while the citizens have upheld theirs.

Corruption: Perceived Levels of in National Government:

The regression results show that high levels of perceived national corruption are extremely significant in terms of their impact on trust in the national government. Each wave had a significance level of 99% and the largest correlation coefficient in each of the individual waves. This suggests overwhelmingly that perceived levels of national corruption lead to distrust in the government.

Interestingly, the nonresponse numbers for the question about corruption were the highest of any question asked, and in wave two, more than 60% of participants did not respond. While nonresponse rates began extremely high, waves one and two had nonresponse rates of 55% and 60%, respectively, but by the fourth and fifth wave, the percentages of respondents declining the answer decreased approximately by half to 27% and 29%, respectively. The number of respondents answering that "almost everyone is corrupt" and "most officials are corrupt" remained constant at roughly 5-10% combined. Meanwhile, the number of respondents who answered that "not a lot of officials are corrupt" rose from 28% in wave one, about doubled to 59% and 57% in waves four and five, respectively. The decline in perceived corruption suggests that China's recent corruption campaigns have convinced at least a portion of its population that it has cracked down on corruption. Despite relative improvements in perceptions of corruption, in the fifth wave, which occurred in 2019, more than 28% of citizens still did not answer the controversial question, and four times as many did not answer the question as compared to those who did not answer

tion." In The Chinese Communist Party's Capacity to Rule, Critical Studies of the Asia-Pacific. London: Palgrave Macmillan, 2016. https://doi.org/10.1007/978-1-137-53368-5_1.

45 Yu, Sun. "The Breakdown of China's Social Contract." Financial Times, November 1, 2023. https://www.ft.com/content/d7b9db0a-9275-48ad-be3a-99afcc6bbf4c.

46 Taplin, Nathaniel. "China's Economic—and Social—Contract Is Fraying." Wall Street Journal, August 15, 2023. https://www.wsj.com/articles/chinas-economicand-socialcontract-is-fraying-70f03f44.

the question about trust in the national government. Although nonresponses to the question of trust in the government were reclassified as distrust in the government, nonresponses to the question of corruption and the rest of the independent variables were not.

Irrespective of the reclassification, the anomalous number of nonresponses suggests a high amount of corruption and widespread fear of sharing that opinion, even in an anonymous survey. Recent surveys have also shown that the Chinese public still believes the nation has high levels of corruption.[47] Studies have shown that the government has no easy way to assuage the public's belief in the widespread nature of government corruption.[48] Not acting on corruption will harm trust; however, if the government acts, it will only confirm and potentially exacerbate the citizens' preexisting beliefs about the widespread scale of corruption.[49] The issue of corruption provides a perplexing conundrum for the government with no path toward improving citizens' opinions of the government.

Corruption threatens to erode the informal social contract as it tilts the economic playing field against the average citizen. Meanwhile, a government system based to some degree on the flow of cash naturally creates an unstable social structure. Loyalty is not grounded in any ideology or belief system. Instead, loyalty is a factor of whoever can provide the most cash at a given time, causing unpredictable changes and fluctuations. Corruption erodes the trust between the state and its people. Of the three independent variables, corruption was the only significant variable in all waves, and its significance exceeded 99% in all waves. The overwhelming evidence from the regression, combined with the high magnitude of nonresponse rates and the existing literature, suggests that perceived levels of corruption play a significant role in causing distrust in the CCP.

Healthcare Access:

The regression results showed that lack of access to healthcare had some significance in causing distrust of the government. No question regarding access to healthcare was asked in the first wave, nor were the results significant in the second wave. In the following waves, access to healthcare was significant, with a confidence level of 95% in the third wave, rising to a confidence level of 95% in the fourth wave, and falling to a confidence level of 90% in the fifth wave. In the third wave, the correlation coefficient is negative; however, this is because of the reverse order of the questions. However, this was reversed in the fourth and fifth waves, leading to a positive value in the correlation coefficient.

The degree of significance with respect to healthcare might be lower than the other independent variables because, as some scholars have noted, the expectations for the government are incredibly low.[50] Economic growth and social stability are indirect and universal services. However, access

47 Willemyns, Alex. "Report: 'Little to No' Progress on Fighting Corruption in Asia." Radio Free Asia, January 30, 2024. https://www.rfa.org/english/news/china/corruption-asia-index-01302024104623.html.

48 Gong, Ting, and Wenyan Tu. "Fighting Corruption in China: Trajectory, Dynamics, and Impact." China Review 22, no. 2 (2022): 1–19. ProQuest.

49 Wang, Yuhua, and Bruce J. Dickson. "How Corruption Investigations Undermine Regime Support: Evidence from China." Political Science Research and Methods 10, no. 1 (2022): 33–48.

50 Ratigan, Kirsten. "Riding the Tiger of Performance Legitimacy? Chinese Villagers' Satisfaction with State Healthcare Provision." International Political Science Review 43, no. 2 (2022): 259–278. https://doi.org/10.1177/0192512120927482.

to healthcare is a direct and individualistic service. The difference in the nature of the services means healthcare is typically provided at a local level instead of at a national level. Therefore, even if citizens might be dissatisfied with access to healthcare at a local level, that does not mean they will blame the national government.[51]

Nevertheless, the regression shows that lack of healthcare access leads to distrust in the national government. Without healthcare, it is difficult for citizens to feel a sense of stability from which they can pursue their economic aspirations. While the regression revealed that lack of healthcare access has some degree of significance in impacting levels of trust in the national government, it remains relatively small in comparison to other independent variables, such as economic satisfaction or perceived levels of national corruption.

Conclusion

While overall satisfaction with the CCP remains high, support for the government is not unanimous.[52] Among those who distrust the government, the results from this study have shown that demographic features play little to no role in explaining distrust of the government. However, perceived poor economic conditions and corruption in the national government likely lead to increased distrust towards the national government. Previous scholars have argued that the Chinese government has based its legitimacy on its ability to provide economic and social stability. Performance legitimacy means that should the government provide these factors, citizens will trust the government. However, should they fail, that leads to distrust of the national government. This study supports the theory of performance legitimacy by showing that the lack of economic and social stability is a significant cause of distrust in the national government.

While the results of this study support the theory of performance legitimacy, other possible explanations for the results exist. As previously mentioned, scholars Holbig and Gilley argue that increased trust in the government does not come from real economic prosperity and perceived economic prosperity.[53] Often, real economic growth leads to perceived economic growth, but this is not always the case. Economic growth can slow or even turn negative, but the government benefits as long as citizens perceive the economic situation positively. Holbig and Gilley's theory makes it difficult to determine if citizens distrust the government because their economic situation has worsened or if they believe that their economic situation has worsened.

Critically, the data used in the regression did not ask for objective levels of the respective independent variables but instead the perceived level of independent variables. The study's results do not show

51 Lee, D.C., J. Wang, L. Shi, E. Jiang, J. Akinbami, and R. Hoffman. "Health Insurance Coverage and Access to Care in China." BMC Health Services Research 22 (2022): 140. https://doi.org/10.1186/s12913-022-07498-1.

52 Cunningham, Edward, Tony Saich, and Jesse Saich. "Understanding CCP Resilience: Surveying Chinese Public Opinion Through Time." Ash Center, 2020. https://rajawali.hks.harvard.edu/resources/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time/.; Edelman. "Edelman Trust Barometer." 2024. https://www.edelman.com/sites/g/files/aatuss191/files/2024-01/2024%20Edelman%20Trust%20Barometer%20Global%20Report_0.pdf.; Li, Lianjiang. "Assessing Popular Support for the Communist Party of China." China: An International Journal 21, no. 2 (2023): 51–71. Project MUSE. https://doi.org/10.1353/chn.2023.a898341.

53 Holbig, Heike, and Bruce Gilley. "Reclaiming Legitimacy in China." Politics & Policy 38, no. 3 (2010): 395+. Gale Academic OneFile. https://link.gale.com/apps/doc/A233048048/AONE?u=brun62796&sid=bookmark-AONE&xid=1ca64f53.

if poor economic conditions or high levels of corruption lead to distrust, but instead, they do show that the perceived levels of these factors lead to distrust. Unraveling the relationship between objective reality and perceived reality, while incredibly important, is not the focus of this paper and is better left to those such as Immanuel Kant. Despite the differences between objective reality and perceived reality, at the macro-level, the two have enough similarities to consider the separate realities interchangeable for the purpose of this paper.

This study shows that the theory of performance legitimacy indeed explains the high levels of trust in the Chinese Communist Party and explains the reason for distrust despite being uncommon. The regression showed that both economic satisfaction and corruption at the national level impact the level of trust in the government. The regression results align with the existing body of scholarly literature and other broad national-level surveys, which have found similar results. While performance legitimacy as a model has flaws, all models have flaws. Models, by definition, are not supposed to encapsulate all of the nuances of reality. Instead, they are supposed to simplify reality so this world can be studied and understood. Performance legitimacy provides an incredibly valuable and useful model for explaining the CCP's continued resources and why certain individuals distrust the government. Ultimately, performance legitimacy provides a simple explanation that those with good economic prospects and social stability tend to trust the government, while those with bad economic prospects and little social stability tend to distrust the government. While the future of China remains uncertain and perpetually in flux, hopefully, the theory of performance legitimacy can provide scholars, politicians, businesses, and the world at large an improved understanding of the sources of the Chinese Communist Party's legitimacy.

***

for acknowledgments and appendixes please visit www.tulanejournal.com

***

The Cost of Vaccine Exemption: A Natural Experiment with Vaccine Exemption and Pertussis Incidence

Sam Haines and Joshua Darrow
Washington and Lee University

Abstract

This study examines the effects of removing vaccine exemptions on the incidence of pertussis (whooping cough) in the United States, with a focus on policy changes in California (2015) and New York (2019). Using county-level data from five states (California, New York, Arizona, Minnesota, and Florida) between 2010 and 2022, we estimate the impact of religious and philosophical vaccine exemptions on pertussis incidence. Our findings suggest that both religious and philosophical exemptions are associated with notable increases in pertussis case incidence. We further identify that the impact of both religious and philosophical exemptions is substantially higher in counties experiencing very high transmission of the disease. Our results indicate that eliminating philosophical exemptions can be an effective public health strategy in controlling disease outbreaks, especially in areas with low vaccination coverage. Nonetheless, the overall effect of vaccine mandates on herd immunity does not appear to be limited to outbreaks. We recommend further research to explore the longitudinal impacts of vaccine policies and the role of exemption uptake rates in disease transmission.

Introduction

The COVID-19 era brought vaccine hesitancy and mandates back into the fold of hot-button American political issues. The literature is abundant with reasons that someone in the US might be hesitant to get a vaccine, including politics, personal beliefs, religious beliefs, and influence from family. As a result of this hesitation, and recognizing the communal benefits of herd immunity from widespread vaccination, state and local governments as well as employers have often required that citizens get certain vaccines in order to work, attend school, or travel.[1] Historically, these mandates have been backed by Supreme Court rulings, including Jacobson v. Massachusetts (1905).[2] Eradication of smallpox and polio, for example, required herd immunity, so the federal government was supported by the Supreme Court when it compelled citizens to get the vaccine. Similarly, schoolchildren in all 50 states are currently required to receive certain shots

1 Crouch and Dickes, "A Prediction Model of Childhood Immunization Rates"; Shaw et al., "Immunization Mandates, Vaccination Coverage, and Exemption Rates in the United States."

2 Jacobson v. Massachusetts, 197 US.

before enrolling and during their progression through the system, including those for measles, mumps, and rubella (MMR); tetanus, diphtheria, and pertussis (TDaP); hepatitis; and human papilloma virus (HPV) vaccines.

While there is precedent to require vaccinations regardless of personal beliefs, it is nonetheless uncommon. For many vaccinations "required" for young children (MMR, TDaP, DTaP, HPV, and hepatitis), most US states still have exemptions. All 50 US states offer a medical exemption for vaccinations that pose a threat to an individual's health, most have an exemption for religious belief, and some still have an exemption for personal belief—that is, some personal, non-religious philosophy that urges you ought not get the shot.

It is well known that subpopulations that choose to avoid vaccination are at much higher risk for disease outbreaks.[3] For example, a measles outbreak in 2018-2019 in New York largely affected the Orthodox Jewish population, a group that traditionally opts out of vaccination.[4] Other outbreaks, predominantly of measles, have occurred in communities with low vaccination rates in various parts of the US in the last several decades. Baseline levels of these diseases have also been increasing, as demonstrated by the recent increase in incidence of highly contagious diseases like measles and pertussis.[5] With major outbreaks of infectious disease, public officials are often pressured to reconsider the legitimacy of vaccine exemptions. Proponents of exemptions argue that government services like education should not be withheld simply because of an individual's personal medical decisions; opponents say that the maintenance of public health and the establishment of herd immunity outweigh the surrender of individual freedoms in this scenario.[6] These debates have resulted in actual policy changes in the last several decades, with some states eliminating both religious and personal exemptions in the last several years, citing growing concerns over public health. In 2015, California eliminated vaccination exemptions in response to a measles outbreak, thus requiring public school students to get a DTaP series before entering kindergarten. While the Supreme Court may have already ruled that eliminating vaccine exemptions is legitimate, the court of public opinion is still in deliberation. Thus, the benefit or consequence of disqualifying vaccination exemptions is one that is well worth investigating.

Literature Review

Extensive research has been conducted on the importance of childhood vaccinations. Feikin (2000), for example, identifies a relative risk factor between personal exemptions and incidences of pertussis and measles.[7] Other medical papers focus on the efficacy of childhood vaccinations like DTaP on infection later in life. The literature clearly stresses the importance of vaccination for personal health. There is also a substantial body of literature supporting the importance of high rates of vaccination with respect to herd immunity. Any novel vaccine or drug may be met with a certain

3 Feikin, "Individual and Community Risks of Measles and Pertussis Associated With Personal Exemptions to Immunization."

4 Jacobson et al., "Barriers and Enablers to Vaccination in the Ultra-Orthodox Jewish Population."

5 Constable, Blank, and Caplan, "Rising Rates of Vaccine Exemptions."

6 Colgrove and Samuel, "Freedom, Rights, and Vaccine Refusal"; Emhoff, Fugate, and Eyal, "Is There a Moral Right to Nonmedical Vaccine Exemption?"

7 Feikin, "Individual and Community Risks of Measles and Pertussis Associated With Personal Exemptions to Immunization."

amount of skepticism based on a fear or other environmental factors. Crouch and Dickes (2015) create a model predicting vaccination rates as a result of demographic information, and other studies demonstrate modeling based on personal and political beliefs.[8] Overall, vaccine hesitancy remains a key obstacle to public health, and national vaccine coverage has declined in recent years.[9]

Unfortunately, comparatively little research has been done concerning the efficacy, and thus the morality, of vaccine mandates. Emhoff (2021) approaches the moral question of vaccination exemptions, claiming that if the herd immunity threshold fails to be met, then there is no longer a right to refuse a vaccine and put others at risk.[10] However, other literature suggests that herd immunity is not sufficiently threatened by a philosophical exemption. Shaw et al. (2018) show in a longitudinal study of MMR and TDaP vaccinations that a state's decision to allow religious and philosophical exemptions decreases total vaccination coverage by a mere 2%, while other factors like parental education can increase vaccination rates by as much as 5%.[11] Given that a threshold of 83-95% vaccination is the general consensus for herd immunity, the marginal 2% decrease in coverage via non-medical exemptions seems inconsequential at first glance.[12] Thus, the literature here does not appear to favor the importance of vaccine mandates.

Various studies have expressed that non-medical exemptions can play a key role in determining whether a community is at risk of an outbreak. For example, A 2010 study of California public schools shows that clusters of non-medical exemption communities strongly correlated with clusters of pertussis cases at a statistically significant level.[13] Since California's laws regarding medical exemptions did not change until 2015, the study does not capture the distinct effect of a rollback on these exemptions. This timeline leaves a gap in the literature that we hope this study will fill, in addition to examining how the elimination of California's exemption policies compared to other states may weigh into the policy debate.

Research Question

Our research investigates the potential reduction in disease rate of removing vaccine exemptions, particularly pertaining to California's rollback of its personal belief and religious vaccine exemptions in 2015, as well as New York's removal of its religious exemption in 2019. If the arguments in favor of vaccine mandates have merit, then we would expect a state to see a substantial decrease in cases as a result of an exemption's elimination. Our null hypothesis is as follows: if the arguments allowing vaccine exemptions have merit, then we would expect no significant reduction in pertussis incidence as a result of a rollback. Thus, we explore the effects of a ban on religious and philosophical exemptions on county-level disease data for pertussis (whooping cough). We expect that data from California and New York, states which have

8 Crouch and Dickes, "A Prediction Model of Childhood Immunization Rates."

9 Veys-Takeuchi et al., "Determinants of COVID-19 Vaccine Hesitancy During the Pandemic"; Seither et al., "Coverage with Selected Vaccines and Exemption Rates Among Children in Kindergarten."

10 Emhoff, Fugate, and Eyal, "Is There a Moral Right to Nonmedical Vaccine Exemption?"

11 Shaw et al., "Immunization Mandates, Vaccination Coverage, and Exemption Rates in the United States."

12 Emhoff, Fugate, and Eyal, "Is There a Moral Right to Nonmedical Vaccine Exemption?"

13 Atwell et al., "Nonmedical Vaccine Exemptions and Pertussis in California, 2010."

recently banned these exemptions, will reflect a change in the disease incidence of pertussis compared to other states (Arizona, Minnesota, and Florida) that have maintained exemption policies.

Methods

Our research creates a model for incidence of pertussis as a function of the availability of vaccine exemption. Given the sensitivity of the spread of pertussis, we incorporated a number of control variables into the model. Crouch and Dickes studied the probability of childhood vaccination in preschool and took into consideration various race and demographic factors as well as income and education.[14] Following Crouch and Dickes, we incorporated control variables for race, income, and gender demographics. Since our data captures the actual spread (case count) of the disease rather than the probability of vaccination, we also incorporate factors that might influence the spread of pertussis, like population density. Finally, since our study focuses on vaccination requirements for children in public schools, we control for the proportion of the population that is enrolled in the county public school system. Thus, if more children are enrolled, we would expect a higher magnitude coefficient on the incidence of pertussis.

All fifty states in the US collect data on vaccine-preventable disease transmission, and most make these data public at the state level. Relatively few, however, publish county-level measures of vaccine-preventable disease incidence over time. As a result, we selected for analysis five states, each of which offered this county-level granularity for a sufficient time period. The states studied were California, New York, Arizona, Minnesota, and Florida. Though culturally very different, each of these states is predominantly rural but anchored by one or more major metropolitan areas. We also found that these states had sufficient variation in vaccination policy: New York and California were both states that eliminated exemptions over the observed time period, while Arizona, Florida, and Minnesota remain unchanged. Each state's exemption policy is discussed in more detail in Table 1. We take the counties in these five states to be a sufficient sample with which to conduct empirical analysis. We concede, though, that this choice of sample may leave our findings open to concerns regarding external validity, and indeed we agree that a fuller analysis including all fifty states has the potential to be more illustrative; we leave this task to future research.

To construct our endogenous variable, we divided the pertussis case count in a given county and year by the total population in that county year. Population data was sourced from the US Census Bureau's annual population estimates. We then scaled this value to represent cases per 100 people in a county.

Our focal exogenous variables are both indicators: the first represents whether religious vaccine exemption is permitted in state s in year t, and the second represents whether philosophical exemption is permitted in state s in year t. These variables were constructed manually as follows. First, we referenced a report from the National Conference of State Legislatures (NCSL) that details whether a state offers a religious or philosophical vaccine exemption (or

14 Crouch and Dickes, "A Prediction Model of Childhood Immunization Rates."

State	Religious Exemption	Philosophical Exemption
California[16]	Parents must check a box indicating that their religion prohibits them from seeking medical advice. If this does not apply, they must get a signature from a medical professional. In either case, the parents must sign a form acknowledging the general risk associated with non-vaccination. This exemption was removed in 2015.	Parents must get a signature from a medical professional and sign a form acknowledging the general risk associated with non-vaccination. This exemption was removed in 2015.
New York[17]	Parents must write a statement explaining the basis of their request for exemption, describing their "religious principles," and indicating which immunizations they request exemption from. School district officials may request additional documentation, including letters from religious leaders, if the parent's statement is insufficient. This exemption was removed in 2019.	N/A
Arizona[18]	Parents must sign a form acknowledging the risks associated with the diseases each vaccine is designed to protect against.	
Minnesota[19]	Parents must sign a form acknowledging the general risk associated with non-vaccination. This form must be notarized.	
Florida[20]	Parents must contact their county health department to initiate the exemption process. Once initiated, parents must sign a form acknowledging the general risk associated with non-vaccination.	N/A

Table 1: Application conditions for vaccine exemptions by state

none of the above).[15] In some cases, NCSL gives detailed date information regarding the introduction or relaxation of vaccine mandates; in other cases, they simply report the contemporaneous requirements in each state. In either case, we went on to manually research vaccine mandate and exemption legislation by state to determine or verify the years for which a pertussis vaccine exemption may or may not have been available on religious or philosophical grounds. This research relied largely on news articles from the affected areas as well as analysis of the actual legislation in the state, where available. In line with NCSL's report, we assign a 1 to both the religious and philosophical exemption indicators when a philosophical exemption is available in a given state-year. We do so because a philosophical exemption allows a parent to avoid vaccination due to personal beliefs, and in the states in our sample where philosophical exemptions are permitted, parents need not provide a specific rationale for their opposition to the immunization.

The process for requesting a philosophical exemption is similar across the three states in our sample that currently offer these exemptions. New York and Florida, however, differed notably in their execution of religious exemptions during the period when such an exemption was available in both states. Table 1, below, gives a brief explanation of each state's application process for religious and philosophical exemptions. Though we are able to record the language conveyed to parents, we are unable to determine what portion of applications for these exemptions, if any, were denied. If a state made the approval process slow or uncertain, parents may choose, on the margin, not to opt for the exemption. For the sake of simplicity, we assume that each

15 "State Non-Medical Exemptions from School Immunization Requirements."

state approved the same portion of exemption applications.

Personal income is defined as the per capita personal income, measured in thousands of US dollars, and collected from the US Bureau of Economic Analysis. In order to generate the public-school enrollment data, we culled county level enrollment numbers from each state's Department of Education website, again dividing by the population estimate in the Census to establish the proportion of the county population enrolled in public school. For the purpose of our analysis, we take public school enrollment to mean enrollment in any of the thirteen grade levels, kindergarten through twelfth grade, in a public school in the county. We measure gender as the proportion of the county's population that identifies as male, sourced from the annual Census Bureau population estimates described above.

Similarly, we capture race as the proportion of the county's population that identifies as a non-Hispanic white individual using the same data source. Finally, we calculate population density by dividing a county's total population, sourced from the Census's annual estimates, by its land area, measured in square miles and sourced from the Census Bureau.

It should be noted that the New York State Department of Health combines New York County, Bronx County, Queens County, Kings County, and Richmond County into a single region ("NYC") for the purpose of case reporting. Thus, we do the same for all additional data (proportions regarding race, sex, income, education, etc.), taking a mean of means of each variable and weighting by each county's population. The proportion of the population in public school is collected by summing public school enrollment across all five counties and dividing by the area's total population. Population density is calculated by dividing the total population across all five counties by the land area of the entire region.
Our data analysis uses a weighted difference-in-differences model with two-way fixed effects, where the weights are set to the county's population. Equation (1) spec-

	No Exemption	Religious Exemption	Philosophical Exemption
Pertussis Incidence (cases per 100)	0.0046	0.0033	0.0129
	(0.0073)	(0.0060)	(0.0196)
Avg. Personal Income ($ thousands)	65.2	48.2	48.2
	(23.4)	(15.2)	(13.1)
Population Density (People per Sq. Mi.)	4.21	5.64	1.31
	(8.02)	(10.11)	(2.20)
Prop of Pop Male	0.496	0.488	0.497
	(0.010)	(0.012)	(0.008)
Prop of Pop White (Not Hispanic)	0.404	0.557	0.505
	(0.168)	(0.216)	(0.194)
Prop of Pop in Public School	0.147	0.133	0.161
	(0.030)	(0.019)	(0.024)
N	638	1393	1674

Table 2: Descriptive statistics

ifies the model:

$$P_{ct} = \beta_0 + \beta_1 X^e_{st} + \beta_2 I_{ct} + \beta_3 S_{ct} + \beta_4 M_{ct} + \beta_5 W_{ct} + \beta_6 D_{ct} + \gamma_c + \theta_t + \epsilon_{ct}$$

where P represents the incidence of pertussis, X^e indicates the presence of exemption e (either religious or philosophical), *I* represents the mean income, *S* represents the proportion of the population enrolled in public school, *M* represents the proportion of the population that is male, *W* represents the proportion of the population that is non-Hispanic white, and *D* represents the population density, each contained in county *c*, year *t*, and state *s.* Fixed effects by county and year are captured in γ_c and θ_t, respectively.

We use data from each county-year in Florida, California, Arizona, New York, and Minnesota from 2010 to 2022 as units of analysis in our regression. Arizona was our data availability constraint, since no state school enrollment data were publicly available from prior to 2010. Thus, 2010-2022 was the longest period of time for which all of the relevant data were publicly available.

Results

Descriptive statistics for the entire study population appear in Table 2. An observation refers to a county-year; 285 counties are represented across 13 years (2010-2022). No counties enter or drop from the study during this period. We find sufficient variation in both religious and philosophical exemption across county-years. Table 2 gives a tabulation of these exemptions. As described in Methods (above), no county-year could have a philosophical exemption but not a religious exemption; any religious rationale for abstaining from vaccines, we suppose, would qualify under the broader philosophical exemption umbrella.

Evaluating our model uncovers interesting results, displayed in Table 3. For each exemption, we first perform a simple bivariate regression (columns 1 and 4), then include the controls described above (columns 2 and 5), and finally add county and year-fixed effects to control for unob-

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Dependent Variable: Proportion of Population Infected with Pertussis (per 100)					
Exemption- religious	**0.0032*** **	**0.0033*** **	**0.0015**			
	(0.0008)	**(0.0009)**	**(0.0010)**			
Exemption- philosophical				**0.0091*** **	**0.0086*** **	**0.0034** **
				(0.0010)	**(0.0011)**	**(0.0014)**
Proportion of pop in public schools		0.0779***	-0.2057**		0.0235	-0.2088**
		(0.0145)	(0.0891)		(0.0150)	(0.0888)
People per square mile		0.0001*	0.0030		0.0000	0.0031
		(0.0000)	(0.0031)		(0.0000)	(0.0028)
Proportion male		0.0926***	0.1140		0.0139	0.0533
		(0.0260)	(0.1668)		(0.0239)	(0.1632)
Proportion white (not Hispanic)		0.0064***	0.0317		0.0068***	0.0204
		(0.0021)	(0.0303)		(0.0019)	(0.0306)
Mean personal income $K		0.0000*	-0.0001**		0.0000	-0.0001
		(0.0000)	(0.0000)		(0.0000)	(0.0001)
Constant	0.0046***	-0.0576***	-0.0293	0.0038***	-0.0108	0.0096
	(0.0006)	(0.0127)	(0.1023)	(0.0003)	(0.0114)	(0.1006)
Observations	3,705	3,705	3,705	3,705	3,705	3,705
R-squared	0.0125	0.0474	0.3130	0.1049	0.1149	0.3156
Year FEs	-	-	YES	-	-	YES
County FEs	-	-	YES	-	-	YES
Robust standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Table 3: Association between exemptions and pertussis infections

servable unit-level and temporal variation (columns 3 and 6).

Across all specifications, the sign on our focal indicator variable is as expected; an exemption is associated with increased pertussis incidence. The magnitude of this effect varies in its practical significance when evaluated against the mean pertussis incidence for each exemption group. Beginning with religious exemptions, we find modest evidence that religious exemptions increase pertussis incidence, all else equal. Comparing the estimate of 0.0033 in model 2, for instance, to the mean pertussis incidence of 0.0046 among county-years with no exemption, reveals a roughly 72% increase in the likelihood of infection when a religious exemption is introduced, all else equal. This effect loses statistical significance when fixed effects are included in the model, but the estimated coefficient of 0.0015 still represents more than a 30% increase in the risk of infection. This effect is even stronger with philosophical exemptions. Model 5 suggests a 187% (0.0086/0.0046) increase in pertussis infection risk when philosophical exemptions are instated, *ceteris paribus*. Even with county and year fixed effects (model 6), we find a consistently positive and statistically significant coefficient on the philosophical exemption indicator and an increase in the risk of nearly 75% (0.0034/0.0046).

It is worth noting that, when considered separately from mean incidence levels, the magnitudes of these estimates are fairly low. The coefficient in model 6, for instance, suggests that *ceteris paribus* a county with a population of 300,000 which allows a philosophical exemption will have 1 more case of pertussis in a year than a county without any vaccine exemptions. We believe, though, that directly comparing conditions with exemptions to those without exemptions (i.e., by referencing non-exemption infection proportions) provides a much fuller picture.

As a robustness check, Table 4 (in the Appendix) gives regression results for equation 1 without accounting for population weights. Results remain largely unaffected by the presence of weights in the estimation, with the exception of model 3, which is the full estimation of the religious exemption's associated effect on pertussis incidence (including fixed effects). However, this coefficient's estimate isn't statistically significant in either specification. We therefore conclude that the model is, indeed, robust to the inclusion or exclusion of population weights.

Discussion

Regardless of model specification, our estimates suggest a substantial negative impact associated with vaccine exemption policies. We suspect, however, that the cost of vaccine exemptions may vary depending on the disease environment: specifically, exemptions may be more costly in county-years with relative spikes in pertussis compared to county-years with baseline incidence. To test this hypothesis, we augment our standard model by adding a simultaneous quantile regression at the 50th, 70th, 90th, and 99.9th percentile. The model specification remains otherwise unchanged, apart from county fixed effects, which are notably omitted from the simultaneous quantile regression due to a failure to achieve model convergence. Results are given in Table 5, found in the Appendix.

Those regression results offer relatively

strong support for our hypothesis when focusing on philosophical exemptions, but the support is somewhat weaker among religious exemptions. As shown in Table 5, the cost of philosophical exemptions is strongest in places with pertussis outbreaks. The largest estimates suggest that in areas with heavy outbreaks, an additional 2,000 cases per million result from the existence of a philosophical vaccine exemption. Religious exemptions account for virtually zero additional cases, even in pertussis outbreaks.

There are a number of issues that we encountered regarding the integrity of data and the timeframe in which this study was conducted. Perhaps the most obvious is the limited availability of public health data: we only scraped data from 5 states, only one of which had eliminated vaccine exemptions well before the pandemic. While states like New York and Maine implemented vaccine mandates in the last 5 years, all data regarding communicable diseases were influenced by the COVID-19 pandemic since 2020 and thus contribute poorly to our study. Other states without exemptions, like West Virginia, did not have disease data publicly available at the county level over a sufficient time period. We initially had concerns about a reverse causality problem; however, since policy changes often occurred in response to measles outbreaks, the incidence of pertussis was unlikely to be a key factor in a policy change. Since there seems to be little change in disease incidence from religious exemptions, that would not be as effective a tool to control the spread.

We also lack data on the utilization of religious and philosophical exemptions, so it is unclear how vaccine exemption uptake rates may impact the incidence of pertussis cases. It could be that, by happenstance, the states in our sample with only religious exemptions had very low rates of exemption utilization, depressing the estimated effect. Conversely, it could be that philosophical exemptions are very highly utilized across our sample, making our estimate a sort of upper bound on the true association between philosophical exemptions and pertussis incidence.

Conclusion

These results have several important findings and wide-ranging policy implications: first, exemption policies are the most costly for the most infected areas; and second, philosophical exemptions are magnitudes more dangerous than religious exemptions alone. Over the last ten years, it seems that policy changes to vaccine exemptions have often accompanied a preventable infectious disease outbreak: notably the Disneyland measles outbreak in 2015 and the 2019 Orthodox Jewish measles outbreak outside of New York City. Our regression, particularly the quantile regression noted in the appendix, supports these sorts of policy changes; it is perhaps most effective to more strictly enforce vaccinations in counties with a severe lack of coverage since those tend to fare the worst in the event of an outbreak.

Another important finding is that while religious exemptions tend to have little to no (sometimes even a negative) effect on the spread of pertussis in a county, personal and philosophical exemptions definitely have an impact on the disease's ability to spread. For states experiencing outbreaks, it can be very important to eliminate the exemption to prevent further spread and

ensure the safety of the community.

The main policy implication for these data is that vaccine mandates can be used as a responsive tool to outbreaks and particularly harmful spells of disease. This is largely how the policy has been implemented in the last ten years—only in the case of outbreaks of measles, pertussis, or other diseases has a state moved to eliminate vaccine exemptions.

We look forward to seeing what research continues to be done on vaccine exemptions. While we have confidence in the results we attained, there are a number of steps that could be taken to strengthen the significance of this study. Further research could include a more longitudinal assessment of the impact of these mandates, as the vaccines are administered in early childhood life but often have doses that last through adulthood. Researchers may also wish to capture a greater portion of the population, perhaps examining as many as all fifty US states. We would also hope to control for education and incorporate variables that show the actual effectiveness of a vaccine policy on people's propensity to get a vaccine exemption, rather than simple disease incidence. However, we have confidence that these results provide interesting insight into an issue that has vast policy implications today.

***

for acknowledgments and appendixes
please visit www.tulanejournal.com

***