Title
Essays on Finance, the Environment, and Philanthropy

Author
Roth Tran, Brigitte

Publication Date
2016

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on Finance, the Environment, and Philanthropy

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Economics

by

Brigitte Roth Tran

Committee in charge:

Professor Richard T. Carson, Chair
Professor Julie Cullen
Professor Joshua Graff Zivin
Professor Christopher Parsons
Professor Richard Somerville

2016

The Dissertation of Brigitte Roth Tran is approved and is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2016

DEDICATION

To my husband, Lou, for supporting me on many, many levels as I pursued my dream.

To my children, Arlo and Sybil, who made graduate school the ultimate growing experience and have enriched my life by broadening my perspective.

To my mother, Diana, and my sister, Isadora, who helped with the kids when it mattered most, so I could tackle my dissertation.

To my father, Howard, for his steadfast support and excellence as a sounding board.

To all in my community who supported me during this process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

A special thanks to my dissertation committee members: Richard Carson, Julie Cullen, Josh Graph Zivin, Chris Parons, and Richard Somerville. Additionally I would like to thank the participants at the following seminars and conferences: UCSD, Center for Marine Biodiversity and Conservation at Scripps Institution of Oceanography, Federal Reserve Board of Governors, Resources for the Future, Federal Deposit Insurance Company, Federal Housing and Finance Agency, Bureau of Labor Statistics, National Institute of Standards and Technology, Cal Poly San Luis Obispo and University of San Francisco. Additionally I would like to thank the following people for their comments and support on this research: Mark Jacobsen, Shanthi Manian, Erin Troland, Ivana Komunjer, Jelena Bradic, and Elizabeth Hastings.

A special thanks also to the apparel and sporting goods company that generously gave me access to its sales data.

Chapter 1 has material that is being prepared for submission for publication. The dissertation author is the principal author of chapter 1. Chapter 2 is based on a paper that has been submitted for publication. The dissertation author is the principal author of chapter 2. Chapter 3 has material that is being prepared for submission for publication. The dissertation author is the principal author of chapter 3.

Chapters 2 and 3 contain material based upon work supported by the National Science Foundation under Grant No. 0903551.

VITA

1999	Bachelor of Science, cum laude, California Institute of Technology
2010	Master of Arts, University of California, San Diego
2016	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

"Discounting Behavior and Environmental Decisions" (with Richard T. Carson), Journal of Neuroscience, Psychology, and Economics, 2009, Vol. 2, No. 2, 112-130.

ABSTRACT OF THE DISSERTATION

Essays on Finance, the Environment, and Philanthropy

by

Brigitte Roth Tran

Doctor of Philosophy in Economics

University of California, San Diego, 2016

Professor Richard T. Carson, Chair

This dissertation examines environmental impacts on economic activity and finance issues for private philanthropic foundations that might want to fund environmental efforts. The first chapter examines the effects of weather on retail activity using the lasso machine learning method to develop a flexible weather index. The second chapter presents a theoretical model showing that large investments in objectionable firms can be optimal for foundations when they yield opportunities to hedge missions. The third chapter examines tax return data for private foundations to explore intertemporal spending strategies and the tax on net investment income.

Chapter 1

Blame it on the Rain: Weather Shocks and Retail Sales

Weather plays a pervasive role in economic activities. In addition to influencing variables like crop yields, energy consumption, crime rates, and mortality, weather affects our moods, ability to travel, enjoyment of outdoor amenities, and purchasing decisions. Weather touches on nearly every aspect of our lives, and retail activity is no exception. Failure to properly attribute variation in retail sales to weather shocks results in suboptimal contracting, undue volatility in commission-based pay, poor demand forecasts, and misinterpretation of brand-specific and macroeconomic financial indicators.

Academic research has shown that weather influences sales of products ranging from mobile purchases to cars. Quantitative investigations into the retail-weather relationship began with the 1951 Steele *Journal of Marketing* paper which showed that cold temperatures, precipitation, and snow cover reduced sales in a department store in Des Moines, Iowa. Despite a promising start, subsequent research has remained sparse and narrowly focused on the marketing and behavioral aspects of the relationship. The popular press, the Federal Reserve Board, and company documents filed with the Securities and Exchange Commission have all cited unseasonable weather as causing disappointing retail sales, which in turn lead to lower GDP growth. However, serious assessments of the role of weather-related volatility in retail sales volatility are limited to Starr-McCluer's (2000) Federal Reserve Board working paper and a Lazo, Lawson, Larsen, and Waldman

(2011) paper in the *Bulletin of the American Meteorological Society*. These works both point to two key problems. First, to the extent that there are offsetting activities, spatial and temporal aggregation can mask a lot of interesting economic behavior. Second, simple weather statistics such as cooling degree days or precipitation may not adequately represent the nonlinear influence of weather shocks, particularly when those indicators are aggregated across time or considerable geographic space.

In this paper I use proprietary national daily store-level sales data for a major apparel and sporting goods brand combined with a new weather index method to estimate how retail sales respond to weather shocks. My in-store sales data span almost four years and about 100 stores across the country in both indoor and outdoor shopping centers. The granularity of this data allows me to overcome the first problem of aggregation and examine economic behaviors like substitution and adaptation on a variety of temporal and spatial scales. Daily zip code level sales data from the company's high traffic internet site allow further examination of substitution behavior to a different and increasingly important shopping channel.

I address the second problem of oversimplifying weather characteristics by using comprehensive weather data from individual stores' closest weather stations and applying a modern variable selection technique, the lasso, to create weather indices to explain sales data variability. Along the way I address a number of interesting questions, such as whether intertemporal effects of weather shocks differ for positive and negative shocks, how important regional heterogeneity is in consumer responses to weather shocks, and how much expected variability in same-store sales over time is due to weather shocks. I conclude with a discussion of how climate change might be expected to influence sales at the firm studied.

Highlights of results include the following. Weather effects on sales are large, nonlinear, and heterogeneous by region, season, and whether a store is in an indoor or outdoor mall. The worst five percent of weather days for shopping reduce same-day sales by 22 and 12 percent for stores in outdoor and indoor malls, respectively. Amplified on surrounding days, these losses are permanent, with very limited recovery over longer time scales. While some moderate weather shocks may shift sales between indoor and outdoor malls, I find no support for substitution between in-store shopping and e-commerce. Estimates suggest that weather shocks account for up to about one-third of monthly store-level sales variation and that weather causes about 5 to 35 percent of total variability in quarterly year-over-year same-store sales growth.[1]

My retailer may not be representative of all retail sales, but the methodology presented can be applied more broadly. This first paper is the first in the literature to implement a rigorous variable selection method to develop a flexible comprehensive weather index with which to address questions of intertemporal and other substitution effects. It is also the first to examine weather shock effects on shopping at indoor versus outdoor shopping centers, a distinction which yields large and significant effects, and the first to examine the relationship of weather effects on in-store and online sales.

The remainder of this paper is laid out as follows. Section 3.2 reviews the relevant literature. Section 1.2 lays out a theoretical model for purchasing behavior. Section 3.6 outlines key features of the data used. Section 1.4 explains the empirical strategy. Section 3.7 covers results. I end in section 3.8 with summary remarks and suggestions for future research.

1.1 Literature

As mentioned above, Steele (1951) is the first to examine weather effects on shopping behavior. This and subsequent research are largely marketing-oriented and

[1]The low value corresponds to the Southwest in the second quarter and the high value the Northeast in the first quarter.

performed on either very small temporal or geographic ranges or at large aggregated scales. Maunder (1973) uses weekly national sales and a national level weather index to show that effects of abnormal temperature and precipitation on retail sales vary by season. Starr-McCluer (2000) shows that national level monthly retail sales are affected by weather, but that these effects are offset by neighboring months, yielding no meaningful effects at the quarterly level. Parsons (2001) finds that precipitation and maximum temperature decrease daily foot traffic at one shopping center over about a half-year worth of data. Lazo, Lawson, Larsen, and Waldman (2011) examine state- and sector-level data to show that weather variability could cause interannual variation in economic activity on the order of 3.4% of gross domestic product. Bahng and Kincade (2012) examine temperature effects on women's business wear in South Korea and find that periods of higher variance in temperature increase sales of seasonal garments. Bertrand, Brusset, and Fortin (2015) study national level apparel sales in France by garment type with a focus on managers using weather information to increase firm profitability.

One strand of literature has focused on the psychological impact of weather on purchases. For example, Howarth and Hoffman (1984) show that anxiety and skepticism decrease with higher temperatures. Hannak and Riedewald (2012) find that weather has significant effects on aggregated sentiments of twitter posts. Coupled with research like that of Spies and Loesch (1997), which shows that better moods may increase purchases, these works present a psychological mechanism through which weather affects purchasing behavior. Levi and Galili (2008) present a psychological argument involving gambling, mood and weather whey they show that cloudy weather causes certain individuals to buy more stocks. Conlin, O'Donoghue, and Vogelsang (2007) show that people are more likely to return cold-weather items ordered from a catalog during cold weather, indicating the purchases were made in part due to projection bias (wherein they incorrectly projected future value of the good based on current conditions.) Busse, Pope, Pope, and Silva-

Risso (2014) similarly use micro data on daily car sales across the country to show that sunshine increases convertible purchases, a phenomenon they attribute to projection bias. In contrast to major purchases like cars, Li, Reinaker, Zhang, and Luo (2015) use an experimental setting to examine the effects of weather on purchases made on mobile phones by over 10 million users. They find that purchases increase with sunshine and decrease with precipitation. Furthermore, the relative effectiveness of different framing strategies differ with weather.

Given that I examine an apparel and sporting goods brand, the literature on how weather affects outdoor activity also applies. An example of work in this area, Smith (1993) shows that beach use, swimming, golf and tennis all respond to temperature, with some nonlinear effects. Graff Zivin and Neidell (2014) find that high temperatures decrease time allocated to outdoor leisure. Tucker and Gilliland (2007) review 37 studies on how weather and seasonality affect physical activity and find that 73 percent of the articles examined report significant impacts.

Finally, there is also a significant literature on climate change and economic activity that ties to this work. Dell, Jones, and Olken (2014) review literature linking weather to various economic elements, especially as they relate to climate change. They cite significant research in the areas of agricultural and labor productivity, energy consumption, health and mortality, conflict and political stability, and crime and aggression. Dell, Jones, and Olken (2014) does not cite any work on how weather affects retail activity in the context of climate change. Beatty, Shimshack, and Volpe (2015) examines purchases of bottled water before and after hurricanes, natural disasters which are likely to increase in frequency under climate change. They use store level data on this one specific non-durable good and find that storm warnings increase sales somewhat, with hurricane landfall resulting in sharp increases in sales. However, a large gap remains in terms of examining non-extreme weather events and consumer durables. The work presented in this paper addresses that gap by exploring how consumer demand is affected by weather extremes, which has implications for climate change. The work presented in this paper seeks to fill that gap by exploring how consumer demand is affected by weather extremes and what possible implications are for climate change.

1.2 Theoretical Model

To explain how weather shocks can cause rational shifts in time, venue and channel of consumer purchases as well as permanent changes in demand for products, I build on the Starr-McCluer (2000) approach which adds a weather component to the Ghez and Becker (1975) household-production model.

In this model, an individual produces n household "commodities" C_{it} in period t. These commodities are broadly defined and include activities like golfing, going to the beach, eating dinner, and recreational shopping. A commodity is produced through a combination of household labor h_{it} and purchased goods vector $\boldsymbol{q}_{it}$ as follows:

$$C_{it} = C_i(\boldsymbol{q}_{it}, h_{it}, \boldsymbol{\theta}_{it}) \tag{1.1}$$

The vector $\boldsymbol{\theta}_{it}$ is an addition introduced by Starr-McCluer (2000) that "represents factors that shift the productivity of goods or labor in the production of C_{it}." Weather is an important component of $\boldsymbol{\theta}_{it}$.

In my model, I allow utility in period t to depend on the production of household commodities in surrounding periods as well as the current one. Denote vectors of commodity production levels over time, as anticipated during period t, as C_i^t, defined as

$$C_i^t = \{C_{i0}, \ldots, C_{it-1}, C_{it}, E_t[C_{it+1}], \ldots, E_t[C_{iT}]\}. \tag{1.2}$$

Then individual utility in period t is

$$U_t = U(C_1^t, C_2^t, \ldots, C_n^t) \tag{1.3}$$

This allows individuals to prefer a mix of activities over doing the same thing day after day. It also allows enjoyment of an activity to change with anticipation of some future activity. For example, an individual may enjoy shopping for golf clubs when it is snowing if she anticipates a future summer golf vacation. An individual may appreciate a beach outing less if he has just spent the last five days on the beach than if he has been in the office all week.

The individual maximizes a discounted utility stream as follows:

$$\max_{h,q} U = \max_{h,q} \sum_t \delta^t U_t. \tag{1.4}$$

This is subject to the budget constraint

$$\sum_t \frac{1}{(1+r)^t} \left[\sum_i \sum_j p_{jt} q_{ijt} \right] = A_0 + \sum_t \frac{1}{(1+r)^t} w_t H_t, \tag{1.5}$$

where r is the interest rate, p_{jt} and q_{ijt} the price and quantity[2] for good j at time t (for commodity i), A_0 the initial assets, w_t the wage rate, and H_t the hours of paid work. The time budget constraint requires total time L_t to equal paid work time plus household production time as follows:

$$L_t = H_t + \sum_i h_{it} \tag{1.6}$$

Given some basic tractability assumptions, optimal levels of goods (q_{ijt}^*) and

[2]Note that q_{ijt} may be durable or non-durable. In the case of a durable good, if the individual already owns the item, it may not be necessary to purchase the good at time t in order to use it then. However, every use contributes to some depreciation of the good, thus in effect imposing a cost on use of the good at time t.

household labor (h_{it}^*) will be functions of current and expected future wage rates $\{w_t, \ldots, w_T\}$, prices $\{\boldsymbol{P}_t, \ldots, \boldsymbol{P}_T\}$, and household productivity factors $\{\boldsymbol{\theta}_t, \ldots, \boldsymbol{\theta}_T\}$ (including weather), as well as the interest and discount rates r and δ.

$$\begin{aligned} \boldsymbol{q}_{it}^* &= \boldsymbol{q}_{it}^*(w_t, \ldots, w_T; \boldsymbol{P}_t, \ldots, \boldsymbol{P}_T; \boldsymbol{\theta}_t, \ldots, \boldsymbol{\theta}_T; \delta; r) \\ h_{it}^* &= h_{it}^*(w_t, \ldots, w_T; \boldsymbol{P}_t, \ldots, \boldsymbol{P}_T; \boldsymbol{\theta}_t, \ldots, \boldsymbol{\theta}_T; \delta; r), \end{aligned} \tag{1.7}$$

In this model, a shock to current or expected future weather can affect sales of good $q_i j$ in the following ways:

1. **Intertemporal substitution:** A weather shock that shifts productivity of labor and goods allocated to commodity C_{it} relative to C_{it+h} may cause an individual to produce commodity C_i in a different period and purchase good $x \in \boldsymbol{q}_i$ at a different time.

2. **Channel substitution:** A weather shock can affect the relative productivity of different types of shopping commodities, leading to a change in venue of purchase for good x. For example, suppose that C_{it} is a mall shopping trip, while C_{lt} is online shopping, with $x \in \boldsymbol{q}_{it}$ and $x \in \boldsymbol{q}_{lt}$. Inclement weather, like heavy rain that makes roads dangerous to travel, can reduce the relative productivity of C_{it}, making online shopping more attractive and causing individuals who otherwise would have purchased x at the mall to make the transaction online. Similar substitution can occur between purchasing a good at an indoor mall versus a store in an outdoor shopping area.

3. **Sales gained from or lost to outside options:** A weather shock may result in substitution between commodities C_{it} and C_{lt}, where good x is an element of $\boldsymbol{q}_{it}$ but not $\boldsymbol{q}_{lt}$. This weather-induced substitution to or from an alternative activity may increase or decrease sales for good x without corresponding intertemporal or

channel substitutions.

1.2.1 Sandals example

Suppose an individual is considering the following commodities:

$$\begin{aligned}
C_{1t}(\boldsymbol{q}_{1t}, h_{1t}, \boldsymbol{\theta}_{1t}) &= \text{shopping in indoor malls for sandals} \\
C_{2t}(\boldsymbol{q}_{2t}, h_{2t}, \boldsymbol{\theta}_{2t}) &= \text{shopping at outdoor malls for sandals} \\
C_{3t}(\boldsymbol{q}_{3t}, h_{3t}, \boldsymbol{\theta}_{3t}) &= \text{wearing sandals at the beach} \\
C_{4t}(\boldsymbol{q}_{4t}, h_{4t}, \boldsymbol{\theta}_{4t}) &= \text{shopping online for sandals} \\
C_{5t}(\boldsymbol{q}_{5t}, h_{5t}, \boldsymbol{\theta}_{5t}) &= \text{alternative activity}
\end{aligned} \tag{1.8}$$

Suppose there is an unexpected downpour, reflecting a shift in $\boldsymbol{\theta}_{it}$. One can imagine that:

$$\begin{aligned}
\frac{\partial C_{1t}}{\partial \boldsymbol{\theta}_{1t}} &> 0 \quad \text{(shopping in indoor malls for sandals)} \\
\frac{\partial C_{2t}}{\partial \boldsymbol{\theta}_{2t}} &\gtreqless 0 \quad \text{(shopping at outdoor malls for sandals)} \\
\frac{\partial C_{3t}}{\partial \boldsymbol{\theta}_{3t}} &< 0 \quad \text{(wearing sandals at the beach)} \\
\frac{\partial C_{4t}}{\partial \boldsymbol{\theta}_{4t}} &= 0 \quad \text{(shopping online for sandals)} \\
\frac{\partial C_{5t}}{\partial \boldsymbol{\theta}_{5t}} &\gtreqless 0 \quad \text{(alternative activity)}
\end{aligned} \tag{1.9}$$

These relationships could result in an individual shopping for sandals sooner than originally planned, as he reschedules his beach outing to a later date. It could result in the individual shifting from buying sandals (still needed for next week's vacation) in the downtown shopping district to doing so at an indoor mall. Or he could simply buy them online. Finally, the consumer could abandon the sandal purchase entirely, as he really needs a good pair of rain boots instead.

1.2.2 Adaptation

The same weather occurrence on the same day will affect sales heterogeneously across regions due to differences in:

- **infrastructure** (e.g. storm drainage systems that minimize flooding, indoor malls, snow plows)
- **individual adaptability** (e.g. ability to drive in rain or snow, four-wheel drive, rain gear for hiking)
- **local weather norms** $\rightarrow$ rarer events likely to be more disruptive (e.g. warm and sunny day in January in Minneapolis vs. Los Angeles)

These three adaptation elements are likely to be highly correlated. Infrastructure and individuals are likely to be well-adapted to weather that is within regional norms.

I represent these adaptation elements with $\boldsymbol{\eta}_{it}$. To account for the fact that weather will impact household commodity production differentially, I update equation 1.1 as follows:

$$C_{it} = C_i(\boldsymbol{q}_{it}, h_{it}, \boldsymbol{\theta}_{it}, \boldsymbol{\eta}_{it}) \tag{1.10}$$

1.3 Data

This paper combines data on retail sales in stores and online with weather. I now discuss each of these data sources in turn.

1.3.1 In-Store Sales

I use a proprietary sales data set for a major national apparel and sporting goods brand. My data consist of daily observations of net sales per store for about 100 U.S. locations across the country. These in-store data span 3.8 years from March 2010 through

December 2013. I exclude stores open less than 100 days, which represent about 5 percent of the total sample. About 10 percent of stores in my final sample close at some point during the observed time period. The majority of these exit in 2013. About 15 percent of stores in the final sample open during the time frame of the data set. The majority of these enter by the end of 2011.

Based on satellite and street view images as well as conversations with employees at store locations, I have identified each store as located in either indoor or outdoor shopping centers. I will henceforth simple call these indoor and outdoor stores, where I define indoor stores to be those that customers enter from within a building, so they are not exposed to the elements between stores. I define stores at outdoor malls or urban shopping districts as outdoor. About 80 percent of my store sales observations are in outdoor stores.[3]

1.3.2 E-commerce

In addition to in-store sales, I use zip code level data on daily online sales for the same apparel and sporting goods brand as the in-store sales. These data span January 2010 through August 2013, for a 3.4 year overlap with in-store sales. In the normal course of business, the firm does not track the date an online order is placed. Instead it defines the transaction date as the date on which the order is filled, typically the first day of business after the order has been placed. This means that orders placed on Saturdays and Sundays will largely be marked with the date corresponding with the following Monday.

I make two key adjustments in combining online and in-store sales data. First, to minimize measurement error bias introduced by this shift in timing, I combine e-commerce data with in-store sales at a weekly level, matching Wednesday-Tuesday in-

[3]Outlet stores represent a significant fraction of the stores in my data set and are typically located in outdoor malls.

store data with Thursday-Wednesday online purchases. Second, to address the different geographic scales, I compare in-store and online sales at the Metropolitan Statistical Area (MSA) level. Thus I can examine aggregate e-commerce, outdoor store, and indoor store sales on a weekly basis at the MSA level.

1.3.3 Weather

I use daily maximum temperature, minimum temperature, precipitation, snowfall, and snow depth observations by weather station from the National Oceanographic and Atmospheric Administration (NOAA) National Climactic Data Center (NCDC). Weather stations vary greatly in data quality and completeness. I use data from weather stations at commercial airports and weather forecast offices, which use weather data in the normal course of operation and are generally considered reliable. I inverse-distance weight observations from all such weather stations within 70 miles and 400 meters elevation of each store. Specifically, the value for element $ELEM_{it}$ for store location i on day t is calculated using the observations for the element at the $j \in J$ individual weather stations assigned to store i as follows:

$$ELEM_{it} = \sum_{j} ELEM_{jt} \cdot \frac{\frac{1}{d_{ij}}}{\sum_j \frac{1}{d_{ij}}}, \tag{1.11}$$

where $ELEM_{jt}$ is the observation of the element at weather station j on day t and d_{ij} is the distance from store i to station j.

I exclude weather stations missing more than 5 percent of precipitation, maximum temperature, or minimum temperature observations during the 2010-2013 time-frame.[4] This yields up to seven stations per store.[5]

[4] This also minimizes bias due to weather station addition and attrition. Suppose that a weather station that is relatively cold enters halfway through the observed time-frame. When weather variables are averaged over the new set of stations, it will appear that the region has gotten colder, even if it has not.

[5] Three stores had to be dropped from the analysis because of inadequate weather observations.

When an element observation is missing for a particular station on a given day, I drop that element observation for any stores that include the station in their weighted averages. However, many stations in the data are missing a significant fraction of snowfall and snow depth observations, often but not always because lack of snow has been coded with blanks rather than zeros. To avoid removing all of these observations, I take the following steps. I examine national monthly snowfall maps to identify store locations that received zero snow during the 2010-2013 period.[6] I replace missing snowfall and snow depth observations with zeros where I can clearly identify that there has been no snow during the observed time period. I next identify remaining stations with poor snow data coverage and exclude these from the snowfall and snow depth inverse-distance weighted averages, redefining weightings for equation 1.11 as appropriate. I further replace with zeros any missing snow observations on days when minimum temperature exceeds 40°F.[7] Consider, for example, a store with three weather stations within range, one of which is missing about half of the snow data. I average the temperature and precipitation observations over all three weather stations. However, I average snowfall and snow depth over only the two stations with adequate data.

Table 1.1 shows summary statistics for about 140,000 daily weather observations at the stores in my data. Average temperature is the simple daily average between maximum and minimum temperatures. Note that weather observations vary within a significant range.

Weather Norms

I use NCDC weather station level observations of temperature, precipitation, snowfall and snow depth from the 1980-2009 time-frame to create weather normals. For

[6] I further confirmed this with news searches for reports of snow events.

[7] This minimum temperature-based replacement of missing snow observations with zeros affects less than 0.1% of final observations. The majority of these replacements are for observations in the summer.

every element at every station, I create a separate normal for each calendar day. I apply a linear Bartlett weighting kernel to smooth the 14 days before and after the day of interest t^* (for a total of 29 observations each of 30 years).[8] This kernel places the greatest weight on t^*, with the weights on neighboring days decreasing with distance from t^*. This smoothing incorporates rare events without introducing too much day-to-day noise in normals.

To create store-level normals, I apply the inverse-distance weighted average process described in section 1.3.3 to these station-level normals.

1.3.4 Geographic Distribution of Sales Data

I group the data using the climate regions defined in the U.S. National Climate Assessment (Melillo, Richmond, and Yohe (2014)), a collaboration by NOAA and the Executive Office of the President. Figure 1.1 shows a map of the regions included. The distribution of stores in my sample reflects the distribution of the general population, with clusters of stores concentrated in or near major urban areas.

Figure 1.2 shows how observed temperature distributions vary by region, while Figures 1.3 - 1.5 show how the distributions of observed non-zero precipitation, snowfall, and snow depth vary by climate region. Note that in Figures 1.3 - 1.5, the values in parentheses next to the region names indicate what fraction of observations have positive values. These figures show that I have a wide range of observed weather events as well as a wide range of climates. For example, Figure 1.3 shows that while the Northwest most closely resembles the Northeast in terms of frequency of precipitation, it looks a lot more like the Southwest in terms of the amounts of precipitation it experiences within a given day. Analyses that aggregate weather across large spatial scales will not be able to

[8] For stations with fewer than 30 years of data, the window was widened proportionately to maintain roughly the same number of final observations. A station with 20 years of data, will have a window 1.5 times as long (e.g. 21 days). Stations with fewer than 10 years of historical data were dropped.

capture the differences in how these norms affect local responses.

1.4 Empirical strategy

While weather variations are exogenous to other economic variables, weather and sales are both beholden to seasonality. For example, back-to-school and Christmas holiday shopping drive up sales. Failure to account for periods like these would bias regression results for weather variables. To avoid incorrectly attributing regular shopping seasonality to weather variations, I establish a baseline regression for my results with fixed effects, to which I later add weather variables. The baseline equation I estimate is as follows:

$$\begin{aligned}\ln(\text{Sales}_{it}) = \alpha + \boldsymbol{\alpha_i} + \boldsymbol{\beta_1} \cdot \text{year}_t \\ + \boldsymbol{\beta_2} \cdot t + \boldsymbol{\beta_3} \cdot t^2 + \boldsymbol{\beta}_{4i} \cdot \text{store}_i \cdot t + \boldsymbol{\beta}_{5i} \cdot \text{store}_i \cdot t^2 \\ + \boldsymbol{\beta_6} \cdot \text{month}_t + \boldsymbol{\beta}_{7i} \cdot \text{store}_i \cdot \text{month}_t \\ + \boldsymbol{\beta_8} \cdot \text{holiday}_t \\ + \boldsymbol{\beta}_9 \cdot \text{weekday}_t + \boldsymbol{\beta}_{10i} \cdot \text{store}_i \cdot \text{weekday}_t \\ + \boldsymbol{\beta}_{11} \cdot \text{store\_closure\_or\_opening}_{it} + \varepsilon_{it}\end{aligned} \tag{1.12}$$

My dependent variable is the natural log of sales at store i on day t. Thus each β coefficient represents the percentage change in sales due to a 1-unit increase in the regressor in question. I include store fixed effects ($\boldsymbol{\alpha_i}$) to accommodate the broad range of store level mean sales in my data set.

I don't want to conflate with warming and other climatic trends any changes that the brand may be experiencing in sales over time. I thus include a year fixed effect, a daily trend, and the square of the daily trend to accommodate this possibility. Furthermore, to allow for store-specific trends, I interact the daily trend and its quadratic with store indicators.

As discussed, shopping seasonality may be correlated with weather in a way that could bias results. I include fixed effects for each of the twelve months of the year and then a set of store fixed effects interacted with these twelve month fixed effects. This allows each store to experience its own seasons. If back-to-school sales occur in August in one region and in September in another, the month-store fixed effects will capture these seasonalities. If sales increase during the Christmas shopping season in December – when it is also colder than average – the December fixed effect will absorb the seasonal increase in sales. That way, I can see how relatively cold December days affect sales without attributing the fact that it is the holiday shopping season to the cold weather. This also means that I am measuring the effect of weather shocks on sales, not average weather.

I next include holiday fixed effects. In particular, I include a separate indicator for each federal holiday along with indicators for Easter, Cyber Monday, and Super Bowl Sunday.[9] While events like Black Friday increase sales, events like the Super Bowl decrease them. I include these indicators because they have significant explanatory power and occur during specific times of the year and are thus correlated with weather. I do not interact them with stores individually because for any given store each holiday is observed only four times at most.

I include weekday fixed effects, in the form of a separate indicator for each day of the week, because there are very strong shopping patterns by day of the week. I also interact these fixed effects with store fixed effects to allow for heterogeneous shopping patterns. While inclusion of these fixed effects increases the amount of sales variation explained by the model, I do not expect weekdays to be correlated with daily weather

[9] I use one indicator for combined sales on Thanksgiving and Black Friday. Some stores in the data set are open for Thanksgiving, while some are not.

variations.

The final non-weather regressor in my model is a set of indicators for store closure or opening. In particular, I include indicators for whether the observations are in the first or last week or 30 days of a store's presence in the data, if this does not coincide with the start or end of the entire data set. This allows stores to behave differently at these unusual times.

Having controlled for the fixed effects described above, I estimate the effects of weather shocks on sales through a variety of possible specifications. First, there is the simple linear specification in which the error in equation 1.12 includes a linear weather component as follows:

$$\varepsilon_{it} = \gamma \cdot w_{it} + \nu_{it}, \tag{1.13}$$

where w_{it} is a weather element like temperature, precipitation, or snowfall.

Second, I incorporate interactions between weather elements and include higher powers of weather variables to allow for non-linearities. I also incorporate interactions with seasons and regions. I use the lasso method to select among possible weather variables to develop a weather index, as described below in section 1.4.2.

Finally, I use non-parametric specifications, binning weather elements or index values into ranges, a method also used in papers like Busse, Pope, Pope, and Silva-Risso (2014) and Graff Zivin and Neidell (2014). Generically, I represent this as follows:

$$\varepsilon_t = f(\boldsymbol{\gamma}, \mathbf{weather}_{it}) + \nu_{it}. \tag{1.14}$$

1.4.1 Adaptation

Under adaptation, a weather shock is more disruptive in an area (or season) less accustomed to experiencing such an event. Whether it is an inch of snow, an inch of rain, or a hot day, under adaptation, the effect of that weather shock decreases as the normal amount of snow, rain, or heat increases. I follow the strategy discussed in Dell, Jones, and Olken (2014) and estimate the following equation to test for adaptation:

$$\ln(\text{Sales}_{it}) = \alpha + \boldsymbol{\alpha_i} + \boldsymbol{\beta X}_{it} + \phi_1 \cdot ELEM_{it} + \phi_2 \cdot \overline{ELEM}_{it} \cdot ELEM_{it} + \varepsilon_{it} \qquad (1.15)$$

Here $\overline{ELEM}_{it}$ is the normal value for a particular element on a particular calendar day. Adaptation is consistent with ϕ_1 and ϕ_2 having opposite signs. To allow for complex weather effects while examining questions of substitution across time, venues, and channels, I develop a weather index, described below.

1.4.2 Weather Index Creation

To create a one-dimensional weather index, I apply the lasso method which selects from a large set of interacted region- and season-specific weather variables and interactions. Following Hastie, Tibshirani, and Wainwright (2015), the lasso runs in two stages to select a highly predictive model. The first stage repeatedly minimizes the mean squared error subject to a series of constraints as follows:

$$\min_{\beta_0,\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_{ij})^2 \text{ subject to } \sum_{j=1}^{p} |\beta_j| \leq t. \qquad (1.16)$$

Here the constraint is that the l-1 norm on the coefficients β_j sum to less than some value t. At some sufficiently large t, this will yield the OLS estimates. As t gets sufficiently low, this constraint will yield zero coefficients on subsets of regressors. Thus the first stage repeatedly selects variables to exclude from the model for a series of t values.

In the second stage, I use 10-fold generalized cross-validation to identify which of the models developed in the first stage yields the lowest mean-squared error across subsets of data. I apply the lasso separately to the sets of indoor and outdoor stores, and then separately normalize each of these two distinct weather indexes.

Table 1.2 shows a list of weather variables included in the lasso. In particular, I use the base weather variables of minimum, average, and maximum temperature, precipitation, snowfall, and snow depth as well as the squares and cubes of those values. In addition I include all possible interactions of the base weather variables as well as their squares with one another. This yields 68 weather variables. I then interact each of these variables with indicators for season.[10] Finally, I interact each of the variables created thus far with indicators for climate region. This yields 2,854 variables in addition to the fixed effects. I use the lasso selection method to reduce this number of variables to avoid overfitting my model.

The lasso is designed to select which variables do the best job of predicting the dependent variable. However, as described above, I do not wish to attribute to weather seasonal effects that relate to cultural phenomena like Christmas and back-to-school shopping. Thus I want the lasso to include the non-weather fixed effects in equation 1.12. To force lasso to select among the weather variable effects only, I first estimate residuals for sales on all non-weather variables in equation 1.12. Then I estimate residuals for each weather variable of interest regressed on those same non-weather fixed effects from equation 1.12. By the Frisch-Waugh-Lovell theorem, regressing residuals on residuals yields the same coefficients as if I included all regressors. Using these residuals as inputs for the lasso, I run the standard two-stage lasso procedure outlined above to produce weather indexes for outdoor and indoor stores, each of which I standardize separately before combining them.

1.4.3 Substitution and Offsetting of Weather Effects

Because I do not have data on individual shoppers, I cannot test directly for substitution. However, I examine the extent to which contemporaneous gains and losses

[10] I define winter, spring, summer, and fall are defined as December - February, March - May, June - August, and September - November, respectively.

are offset at other times and places, which is consistent with substitution. First, I examine relationships between indoor and outdoor stores. I separately aggregate sales at the MSA level for indoor and outdoor stores, excluding MSAs that only have indoor or outdoor stores. I define the following indexes:

$$\begin{aligned} W_{own,imt} &= W_{outdoor,mt} \cdot \mathbb{1}[i = outdoor] + W_{indoor,mt} \cdot \mathbb{1}[i = indoor] \\ W_{other,imt} &= W_{outdoor,mt} \cdot \mathbb{1}[i \neq outdoor] + W_{indoor,mt} \cdot \mathbb{1}[i \neq indoor] \end{aligned} \tag{1.17}$$

I then estimate the following equation:

$$\ln(\text{Sales}_{it}) = \alpha + \boldsymbol{\alpha_i} + \boldsymbol{\beta X}_{it} + \gamma_1 \cdot W_{own,it} + \gamma_2 \cdot W_{other,it} + \varepsilon_{it} \tag{1.18}$$

Here $\boldsymbol{X}_{it}$ are the non-weather fixed effects in equation 1.12. Because the analysis is no longer performed at the store level, i represents an MSA-store type (indoor versus outdoor) combination. I also add an indicator for the number of stores in the MSA, allowing the sales to shift with entry and exit of stores in the area. The index that corresponds to the store type of location i is represented by $W_{own,it}$, with the index for the other store type represented by $W_{other,it}$. A negative γ_2 coefficient indicates that there is offsetting behavior consistent with substitution between venue types.

I also test for intertemporal effects of weather shocks. Here I follow the structure of equation 1.12 and add lags and leads of weather index values. In the simple weather specification of equation 1.13, negative coefficients on lags and leads are consistent with intertemporal substitution from the firm's perspective. When examining leads and lags of weather index quantile bins, offsetting and substitution is indicated by positive coefficients on low weather index value ranges and negative coefficients on high ones.

Finally, I look for evidence of substitution between in-store and online sales by regressing weekly MSA-level e-commerce sales on weekly in-store weather index values.

Negative coefficients would indicate substitution.

1.5 Results

I now report results of basic weather analyses, adaptation tests, the weather index developed with the lasso method, and various forms of substitution.

1.5.1 Basic weather results

I first confirm that weather shocks measured in their simplest forms have significant impacts on daily sales. The regressions in Table 1.3 control for year, month, day-of-week, store, store-month, store-day-of-week, and store-trend fixed effects, as described in equation 1.12. Serving as a baseline for comparison, column 1 shows that the equation 1.12 fixed effects alone yield an adjusted R^2 of 0.8556. The addition of each element in columns 2 - 7 yields a higher adjusted R^2 than the baseline, indicating that inclusion of weather variables improves explanatory power. These remaining columns show that on average, all other things being equal, colder temperatures and increased precipitation, snowfall, and snow depth all decrease sales. In particular, a decrease in maximum temperature of 10° Fahrenheit results in a 2.6 percent decrease in daily sales. An inch of precipitation decreases same-day sales by 4.6 percent, while snowfall and snow depth decrease sales by 12.9 and 2.6 percent per inch, respectively.

Figure 1.6 shows that the contemporaneous effects of temperature on sales are non-linear. Each bar in this plot shows the coefficient on observations in 5-degree bins (denoted in the graph by the bottom value of each 5-degree range) relative to a 70-75°F temperature range. Figure 1.6 shows that average temperatures below 40°F increasingly depress sales. In fact, a regression limited to observations with average temperatures below 40°F shows that each 10°F drop in temperature below 40°F results in an 11.3 percent sales reduction, which is four times the average effect shown in Table 1.3. In contrast, temperatures above 40°F do not appear to have significantly different effects except at extreme heat, where average temperatures exceeding 100° F decreases sales by about 15 percent.

Figure 1.7 shows that the average effects for daily average temperatures depicted in Figure 1.6 mask seasonal heterogeneities in responses. The left panel of Figure 1.7 shows sales responses to average temperature ranges relative to the 70-75°F range in the winter, while the right panel shows effects in summer. Note that temperatures in the 45-65°F range tend to have positive effects in the winter but negative effects in summer.

Figure 1.8 shows that precipitation affects sales very differently at indoor and outdoor stores. For example, relative to zero precipitation, 1.5-2.0” of rain will on average increase sales at indoor stores by about 13 percent and decrease sales at outdoor shopping centers by about 10 percent.

Using snowfall and snow depth as examples, Table 1.4 shows that weather effects are heterogeneous across regions. Here the Northeast is the base region. Column 1 shows that one inch of snowfall reduces daily sales by 10.7 percent in the Northeast. Stores in other regions experience even larger decreases in sales in response to snowfall. In the case of snow depth in column 2, where the Northeast is again the base region, sales appear to respond similarly in the Midwest and Northeast, the two regions with the lowest losses due to any given level of snow depth.

In summary, these results show that weather shock effects on sales are large, statistically significant, and non-linear, with meaningful interactions between elements and heterogeneity across regions, seasons, and indoor versus outdoor stores.

1.5.2 Adaptation

The results depicted in Figure 1.7 and Table 1.4 suggest that shopping patterns vary with current local norms. It would be unsurprising if intense rain is more disruptive in Southern California than in Florida where high levels of rain are much more common (see Figure 1.3.) I now test for adaptation, following the approach presented in Equation 1.15.

The regression results in Table 1.5 support the hypothesis of adaptation. The negative top coefficient on precipitation in column 3 shows that precipitation decreases sales. However, the second coefficient, on the interaction between current precipitation and normal precipitation is positive, indicating that as the normal precipitation level increases, the magnitude of the effect of precipitation shock decreases. In fact, in every column but column 2 the coefficients for the interactions with the normal terms and the coefficients on the weather elements themselves have opposite signs. Thus there is clear evidence of adaptation for maximum temperature, precipitation, snowfall, and snow depth. Column 2, which shows results for minimum temperature, is the only one which does not support adaptation. Here the two coefficients are both positive. However, the interaction term coefficient is only statistically significant at the 90 percent level.

These results further support the idea that local norms influence sales responses to weather and that aggregation of weather or sales variables may mask some interesting shopping behaviors. In the next section I show results of a weather index that is flexible enough to allow for all of these effects.

1.5.3 Weather Index Results

The previous two sections have revealed that sales responses to weather shocks are non-linear and heterogeneous by season, region, and store type. To handle these complexities, I might include squares and cubes of each weather variable along with interactions between them and interactions between all of these terms and season, region, and season-region pairs. I would then examine all of these variables separately for indoor and outdoor stores. This would yield over 2,854 weather regressors. Despite having about 140,000 observations to work with, including all of these variables would result in overfitting and make results difficult to interpret. I thus use the lasso method to develop a weather index that captures in one simple number how favorable local current weather conditions are for shopping.[11]

Figure 1.9 shows the distribution of weather index values at indoor and outdoor stores. For ease of interpretation, these weather index values have been standardized to have a mean of 0 and a standard deviation of 1. Figure 1.10 shows how the distribution of weather index values for outdoor stores varies by region. Of particular note, in the upper right and middle plots the Northeast and Midwest regions have much smaller fractions of observations right around 0 than other regions. This suggests these regions experience more variation in sales due to weather than the others.[12] In contrast, the Southwest has a relatively high fraction of weather index observations around 0, making its sales relatively unaffected by daily weather fluctuations.[13]

This regional heterogeneity of weather-induced sales variation is not too surprising given the ranges of weather observed in different regions. However, it does have some interesting economic implications. In particular, compensation and bonuses tied to short-run local store sales will experience more noise in regions like the Northeast and the Midwest that experience greater weather variability in terms of shopping. Thus any distortions in incentives that result from this noise will be more pronounced in these regions. Furthermore, same-store sales metrics will be subject to more weather-based volatility for firms with greater proportional exposure to these regions.

Figure 1.11 shows precipitation-based examples of how the weather index flexibly allows for nonlinear heterogeneous weather effects by region, store type, and season. The

[11] See section 1.4.2 for a description of this process.

[12] The Northwest and Midwest regions also experience the most extreme winter weather, including snow, as shown in Figures 1.4 and 1.5.

[13] Hawaii, not shown in Figure 1.10 has an even more narrow weather index distribution than the Southwest.

dark gray points represent indoor stores, and the light blue points outdoor stores. The left and right panels in each subplot include winter and summer observations, respectively. In each region (Southwest, Southeast, and Northeast), the winter and summer plots have different shapes and ranges. This confirms that the index is capturing nonlinear season- and region-specific responses of sales to weather.

Regression results in Table 1.6 confirm that the weather conditions captured by this index drive significant variation in sales. In particular, column 1 shows that a one standard deviation increase in the weather index corresponds to a 7.5 percent increase in daily sales. Column 2 shows that the effect is more pronounced on the negative end than the positive.[14] In particular, relative to the middle quintile, weather in the worst 5 percent of observations yields an average 20 percent drop in same-day sales across indoor and outdoor stores. Weather in the most favorable 5 percent of observations increases daily sales by 14.6 percent. Columns 3 and 4 show that the index effects are even more pronounced for outdoor stores, with one standard deviation change in weather index corresponding to 8.1 percent change in sales, and the worst 5 percent of weather shocks causing an average of 21.7 percent declines in sales. Finally, columns 5 and 6 show that weather shocks have a weaker effect on indoor store sales. A one standard deviation change in weather index corresponds to a 5.3 percent change in sales, while the worst 5 percent of weather shocks on average drop sales by only 12.4 percent. Thus outdoor stores appear to experience greater variation in daily sales due to weather than indoor stores. This makes sense as shopping in indoor malls protects individuals from the elements.[15]

[14]These quantiles have been determined separately for indoor and outdoor stores, though they are aggregated across the country. Thus an indoor store in the Northeast is as likely as an outdoor store to experience a bottom 5 percent weather shock, but more likely than an indoor store in the Southwest.

[15]Note that because the dependent variable is ln(sales), which is undefined if sales are zero, these results exclude weather events so extreme that they result in store closures.

1.5.4 Substitution and Offsetting of Weather Shock Effects

I now examine behaviors consistent with substitution. Because I have store-level data and cannot observe individual shoppers, I cannot clearly identify individual substitution behaviors. However, I will henceforth refer to the shifts in sales over time periods, venues, and channels as substitution, as per the perspective of the firm. I examine substitution between indoor and outdoor stores, from stores to e-commerce, and over time.

Indoor Versus Outdoor Stores

For an illustrative example, the upper panel of Figure 1.12 shows the weather index values over the first six months of 2011 at two Northeast region stores, one indoor and one outdoor, the two within an hour drive. The lower panel depicts the weather variables during the time frame. Representative of patterns found in other areas and times, Figure 1.12 reveals three regimes.

First there are periods of synchronized movement, where weather affects sales at indoor and outdoor stores similarly. This is likely driven by a change in underlying demand for products, a change in demand for (or cost of) shopping as an activity, or a combination of the two. In Figure 1.12, such periods are most prominent when weather indexes are very negative and coincide with events like major snowfall.

Second, there are periods when indoor and outdoor sales appear to be largely orthogonal to one another. This may be due to offsetting of counteracting factors.

The third regime is one of negative correlation between indoor and outdoor stores, when weather contemporaneously shifts sales between these venues. This evidence is consistent with the notion that some weather affects the type of shopping experience people enjoy most, but does not affect underlying demand for products or shopping itself. The shaded areas in Figure 1.12 correspond to periods when weather indexes for stores in indoor and outdoor shopping centers move in opposite directions, indicating that losses in one type of store are in part offset in the other type of store. This could result from different sets of people shopping at each venue type and behaving in offsetting manners or due to individuals deciding to switch shopping venues due to weather.

Negative correlations between indoor and outdoor weather indexes indicate that losses in one type of store are being recovered to some extent in the other type. These negative correlations are also suggestive of substitution between store types. Table 1.7 shows correlations between indoor and outdoor weather indexes by different subsets of data. The correlation shown for all data indicates that, on average, indoor and outdoor weather indexes are positively correlated. Positive precipitation also yields positive correlation. Consistent with Figure 1.12, snowfall and snow depth increase weather index correlations. Regions have heterogeneous index correlations, with the Midwest experiencing the highest level and the Northwest the lowest. Seasons also yield heterogeneous levels of correlation, with the milder fall and spring seasons yielding lower correlations. In the final group of correlations in Table 1.7, I show examples of broad periods of negative correlation between indoor and outdoor weather indexes, namely in the spring and fall in the Northwest. As suggested by Figure 1.12, it is reasonable to expect that there may be shorter periods of negative correlations that may not be captured in aggregated analyses like the one presented here in Table 1.7, where the intermittent strong positive correlations from severely negative weather events more than offset the periods of negative correlations.

Table 1.8 shows regression results that test for substitution between indoor and outdoor stores. These regressions are performed at the daily MSA-store type (indoor versus outdoor) level. Each column includes two sets of regressors. First, the "own" weather index is the local weather index that applies to the type of store for which sales are being observed, as defined in equation 1.17. Then the "other" weather index is for the other type of store in the same area. A negative coefficient on the "other" index is consistent with substitution between indoor and outdoor stores. Column 1 shows that while the own weather index is significant for sales, the other index is not. This indicates sales losses and gains are not, on average, being offset in other types of stores. Column 2 shows that there is no effect for either the indoor stores or outdoor stores separately. In column 3, I find that there is evidence of substitution during fall, when a standard deviation decrease in the other index yields a 1.5 percent increase in sales (significant at the 90% level). This is a fairly large effect relative to the 6.3 percent own weather index effect shown. In column 4, I find that positive precipitation, snowfall, or snow on the ground do not on average cause offsetting or substition.

Finally, column 5 of Table 1.8 shows that there are substitution effects in some regions but not others. In particular, in the Northeast, which is the excluded base region, a one-standard deviation shift in the other index yields 0.7 percent substitution. I interpret these results as follows. In the Northeast, a standard deviation shift in own weather index represents a 5.89 percent loss in same day sales, of which 12.7 percent will be recovered in the other type of store. In addition to indicating that weather shock-induced losses may be recovered across locations in some regions, the lack of results in columns 1 - 4 can be explained by the Midwest and Southeast regions. In particular, in the Southeast I find a positive coefficient on the other weather index that exceeds the negative baseline coefficient. This indicates that in the Southeast, any substitution effects between venues are outweighed by intensification or demand effects.[16]

Substitution between in-store and online shopping

I now examine whether weather that decreases in-store shopping contemporaneously increases online shopping. Two opposing forces that could apply here are

[16] A basic analysis by bins finds no evidence of substitution by store type, consistent with the first four columns of Table 1.8.

substitution and demand shifts. With local e-commerce sales as the dependent variable in a regression, substitution is characterized by negative coefficients on in-store weather indexes; weather that drives up sales in stores drives down sales online and vice versa. In contrast, demand shifts simultaneously increase or decrease sales in stores and online as weather shifts the underlying demand for products.

The regressions in Table 1.9 examine the relationship between weather favorable for in-store shopping and online sales. Column 1 provides a baseline, with all remaining columns yielding a higher adjusted R^2. Columns 2 and 3 show that the weather captured by the local average weekly indoor and outdoor in-store weather indexes have significant predictive power for e-commerce sales. For example, weather captured by an average 1 standard deviation increase in the indoor index results in a 2.6 percent increase in weekly local online sales.

Column 4 of Table 1.9 includes the weekly mean of indoor and outdoor weather indexes as the independent variable. Weather conditions that cause substitution between indoor and outdoor stores will yield mean indoor and outdoor index values close to zero, while weather shocks that drive both indexes in the same direction will yield very high absolute values. Again, the coefficient here is positive, consistent with shifts in demand for purchases outweighing substitution between sales channels.[17]

Figure 1.14 further breaks these results down into quantiles. Surprisingly, these plots show that none of the worst indoor, outdoor, or mean local weather index ranges have significant effects on e-commerce sales. Thus, there is no support for the notion that individuals substitute to online channels during especially bad weather (for in-store shopping.)

In contrast, the coefficients on the upper bin in Figure 1.13 are positive and

[17]Excluding the week that includes Thanksgiving, Black Friday, and Cyber Monday yields nearly identical results.

significant for the indoor, outdoor, and mean indeces. Thus weather favorable for in-store sales is also favorable for e-commerce. This is consistent with the idea that favorable weather is good because it causes underlying shifts in demand for products. It is also possible that these increases result from online purchases made in stores when unusually high demanded causes stores to sell out of popular products.

In summary, there is no evidence of weather-induced substitution between in-store and e-commerce sales. Losses in stores will not necessarily be recouped online. However, weather shocks that are unusually positive for in-store sales also increase e-commerce purchases, indicating that these types of shocks may be shifting underlying demand for the products sold in addition to making the household commodity of shopping more productive.

Intertemporal Substitution

To understand how meaningful weather effects are, one must look beyond contemporaneous effects to consider whether losses in one period are recovered at other times. Table 1.10 shows results from regressions of ln(sales) on one-day lags and leads of the weather index along with the Equation 1.12 fixed effects and contemporaneous weather. Intertemporal substitution would yield negative coefficients on leads and lags, as a negative weather event at another time increases current sales. All lead and lag coefficients in Table 1.10 are positive. Thus there does not appear to be substitution on a daily basis, nor should one expect losses today to be offset by tomorrow's or yesterday's sales. Sticky expectations about weather and product demand sensitivity to weather could be driving these results.

What if substitution occurs over longer time horizons? Figure 1.15 shows effects on sales of daily weather index leads and lags over the three weeks preceding and following a particular weather shock. The top panel shows day-by-day effects, while the bottom panel shows cumulative effects, telescoping outward from the day of the weather event. Although the top panel shows some individual daily substitution-like effects, when coefficients dip below zero more than a week before or after the weather shock, the cumulative effects shown in the lower panel are clearly positive when significantly different from zero. These results further indicate that intertemporal substitution on a daily level does not result in significant recovery of losses due to poor weather.

It is possible that substitution occurs at one extreme of weather but not the other, and that these two effects offset each other when examining the means. Table 1.11 explores leads and lags of the weather index on in-store sales by bins. Here offsetting and substitution would yield positive coefficients on lower percentile ranges and negative coefficients in the upper ranges. However, I find the opposite is true. Column 3 shows that, in addition to the contemporaneous decrease of 17.4 percent, a day with the least favorable 5 percent of weather for shopping decreases sales the day before by 4.2 percent and the day after by 2.3 percent.[18] This suggests that weather is shifting underlying demand. In summary, on a daily level, there is no evidence of offsetting of weather shocks or behavior consistent with weather-induced intertemporal substitution.

Substitution-like behavior may occur at longer time scales. I now aggregate the sales and weather data to the weekly level to examine this possibility. Table 1.12 confirms that the daily weather index is still meaningful in predicting sales at a weekly level. Column 1 shows a one standard deviation increase in a store's average daily weather index for the week causes increases total sales for that week by 8.9 percent. Column 2 similarly shows that the lowest weather index value of the week has a meaningful effect when combined with the average for the week.[19] Finally, column 3 shows that weekend weather influences sales more than the rest of the week, which makes sense given that a

[18]The best 5 percent of days in terms of weather being favorable for shopping increase sales by 3.4 percent and 3.1 percent on the immediately preceding and subsequent days.

[19]Note that the standard deviation is for the daily weather index.

high fraction of sales occur on the weekend.[20]

Figure 1.16 examines weekly effects by number of days in the week in various daily weather index quantile bins. Only the extremes appear to have significant effects. For every day that weather is in the least favorable 5 percent of weather index values, the store will experience a 2.9 percent decrease in weekly sales.

Table 1.13 includes the same weekly regressors as Table 1.12, but with one or more lags for each regressor. The one-week lag term in column 1 is not statistically significant, but other terms are. There appears to be some offsetting and substitution-like behavior in some leads and lags (indicated by negative coefficients) as well as some amplification of weather effects (indicated by positive coefficients). In column 2, the one-week lag appears to amplify contemporaneous weather effects. However, column 3 results show that there is some substitution-like behavior on the one-week lag across weekends.

Figure 1.17 shows the weekly nonlinear results with three weeks worth of lags. For each percentile bin, the contemporaneous weekly weather effect is shown on the left, with the one-week lag immediately to the right, followed by the two- and three-week lags. Here some lags have coefficients of the opposite sign of the current weather effect coefficient, indicating intertemporal substitution. In particular, the left-most set of coefficients shows that each day with weather in the least favorable 5 percent one week reduces total sales that week by about 3 percent. However, it increases sales the next two weeks by no more than about 1 percent each week and then decreases sales again the following week. On the other extreme, in the top five percent of weather shocks, the first two weeks after a weather shock result in additional benefits to sales, while the third week finally yields some reversal or substitution.

I now examine sales aggregated at the monthly level. Table 1.14 shows results

[20]Results are similar for indoor and outdoor stores when examined separately.

from monthly regressions that control for the number of days per month and the number of weekend days per month.[21] Again, the weather captured by the index affects monthly sales in a meaningful way. Column 1 shows a 1 standard deviation increase in the mean daily index for the month yields an increase of 7.6 percent of monthly sales. Column 2 shows that the worst weather of the month has significant explanatory effects on monthly sales beyond the mean index.

Table 1.15 shows that there is no evidence of intertemporal substitution at the monthly level, when considering lagged mean daily weather index values. With no lag terms significant at the 95 percent level, Figure 1.18 shows the same result in a non-linear fashion. In particular, it shows that the number of days in the lowest 5 percent of the index largely drive the monthly effects. Every day of sales in the lowest 5 percent of weather index values decreases total monthly sales by 0.75 percent. On the other hand, every day in the 90th-95th percentile range increases total monthly sales by only 0.43 percent, though this is not statistically significant. Meanwhile, none of the lag term coefficients are statistically significant. Thus although weather still has meaningful monthly effects, there is no evidence that weather-induced sales gains and losses one month are offset in another month.

In summary, I find no evidence for daily- or monthly-level substitution. However, there are appears to be some minor substitution at the weekly level.

1.6 Conclusion

In this paper, I have used a proprietary apparel and sporting goods data set to examine weather effects on retail sales. The granularity of daily store and online zip code level observations combined with the breadth of over three-and-a-half years worth of

[21] Separate estimates were also calculated for indoor and outdoor stores only, with similar results. Separate results were also calculated for average daily sales over the course of a month, again with similar results.

observations at about 100 stores across the United States have allowed me to measure aspects of previously unexplored aspects of shopping behavior. I have found that weather shock effects on sales are nonlinear and heterogeneous across regions, seasons, and indoor versus outdoor shopping centers.

There are three key contributions of this paper to the literature. First, I have shown that weather effects on apparel and sporting goods retail sales are large and significant. The worst five percent of weather shocks reduce sales by an average 22 and 12 percent at stores in outdoor and indoor malls, respectively. The most favorable five percent of weather shocks increase sales by an average 16 and 10 percent in outdoor and indoor malls, respectively. More anomalous weather events induce larger responses in sales, suggesting local adaptation to weather patterns. Furthermore, I have shown that these effects are largely permanent, with no evidence of substitution at daily or monthly levels and only limited minor substitution at weekly levels.

The second significant contribution is that this is the first work to distinguish between weather effects on sales in indoor malls, outdoor malls, and e-commerce. I have shown that while there may be some shifting between indoor and outdoor stores, weather that decreases sales in stores does not increase online sales. In fact, weather shocks that increase store sales also increase local e-commerce sales, indicating that weather shocks may be driving underlying demand.

My third and probably most important contribution is a new methodology for measuring sales responses in a flexible but interpretable manner. Specifically, I utilize the lasso method to create an index based on a comprehensive set of local weather variables. This index allows for nonlinear heterogeneous responses to weather shocks. It predicts how favorable weather conditions are for shopping, net of seasonal, day of the week, holiday, brand, and store fixed effects. The same combination of elements may be favorable in one time and place and unfavorable in another. This index allows me to flexibly examine the broader question of how persistent losses due to “bad weather” are rather than biasing effects downward by treating weather as if it were not context-dependent.

This work has a number of important implications. Failure to properly attribute variation in retail sales to weather shocks can result in undue volatility in commission-based employee compensation. In the context of principle agent theory, examining outcomes net of seasonality and weather shocks makes those outcomes more informative of effort. Agents who know that their performance will be judged net of weather may experience higher morale, as they perceive the system to be more fair and more likely to reward their effort. Conversations with firm managers and retail sales employees suggest that sales staff do believe that weather is a major driver of sales. In fact, regional sales managers at the brand whose data I have examined here have complained of bad weather being responsible for disappointing sales despite the managers’ best efforts. Initial back-of-the-envelope estimates using bootstrapped historical weather data suggest that weather is responsible for up to one-third of seasonally-adjusted monthly same-store sales variation. If compensation is based on sales adjusted for weather shock effects, staff are likely to experience less volatile income, more direct rewards for effort, and greater incentives to increase effort in all types of weather.

Accounting for weather can also have significant impacts on demand forecasts, inventory management, and strategic planning. Suppose a brand experiences unusually favorable weather, which causes same-store sales to increase. Suppose further that the brand has been on the decline and would otherwise have seen a decline in same-store sales. If the brand accounts for weather, then it will recognize the true negative trend. In the long run, the brand may expect weather to be independently and identically distributed. If its demand forecasts are based on previous year changes, then it will use the negative net-of-weather same-store sales trend and estimate a relatively negative outlook for demand. It may decrease production levels and implement new strategies to change course. However, if the firm instead forecasts demand using same-store sales changes not net of weather, it may think the brand is gaining strength and overestimate future demand and continue with unsuccessful strategies, making problems worse. Consider an example in a shorter time frame. Suppose that ideal shopping weather leads stores in an area to sell out of particular products one weekend. Inventory managers who don't account for weather may not recognize that forecasts indicate unfavorable weather coming the next weekend. These inventory managers might make a large push at significant expense to restock those stores when the demand is unlikely to return in the short run.

Investors and policy analysts monitor indicators like same-store sales and aggregate national retail sales revenue growth. These metrics are taken as indicators of brand strength and macroeconomic health. As in the examples of forecasting, inventory management, and commission-based pay, failure to account for weather in these metrics will yield incorrect inferences. Upon seeing positive same-store sales numbers, an investor may incorrectly conclude that a firm is doing well when unusually favorable weather was responsible for the year over year growth. Similarly, an economist examining macro trends may misinterpret changes in aggregate retail activity if those changes are not net of weather shocks. Back-of-the-envelope estimates using bootstrapped historical weather data suggest that weather causes about 5-35 percent of year-over-year quarterly same-store sales variation. The relative influence of weather on sales growth variation depends on season and region, with, for example, second quarter Southwest same-store-sales growth relatively independent of weather and first quarter Northeast same-store-sales growth very sensitive to weather.

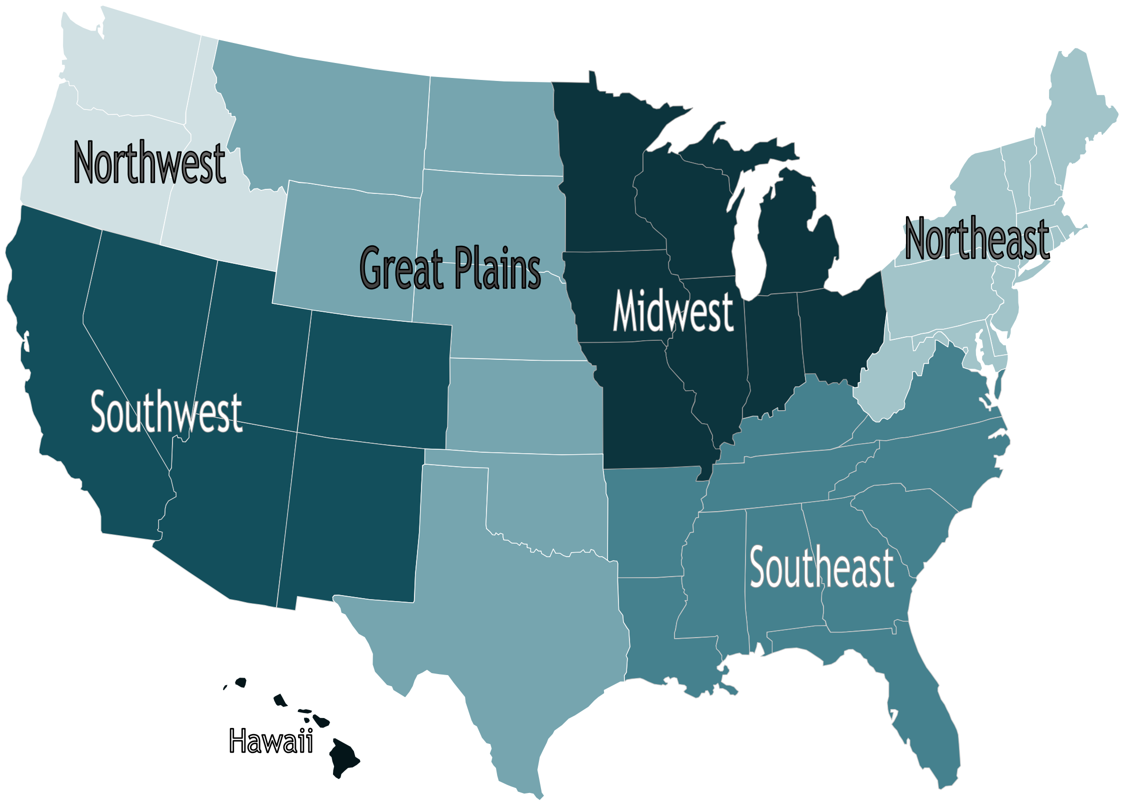
In terms of economic theory, these issues touch on contract theory, the principal agent theory, industrial organization, financial markets, macroeconomics, operations research, and forecasting. By showing how significant weather effects can be and how to measure them in a comprehensive manner, the work presented in this paper may encourage firm managers and policymakers to incorporate weather in their assessments of various financial metrics. This may become increasingly important as we experience more frequent extreme weather events under climate change. Current results suggest that shorter, milder winters predicted under climate change might benefit this brand. However, extreme weather events, including heat waves and intense snow and rain storms, which are expected to increase in frequency and/or intensity under climate change, will likely be bad for sales. One thing is clear, the effects of weather shocks on retail sales are already large enough that managers, investors, and policymakers should take notice now.

Chapter 1 has material that is being prepared for submission for publication. The dissertation author is the principal author of chapter 1.

1.7 Tables and Figures

Table 1.1. Weather summary statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
Temperature				
Max Temp (F)	70.27	17.92	-4.16	118.9
Avg Temp (F)	60.87	16.99	-11.11	105.0
Min Temp (F)	51.47	16.72	-23.09	94.5
Precipitation				
Precipitation (cm)	0.257	0.754	0	17.4
Precipitation indicator	0.399	0.490	0	1.0
Snowfall				
Snowfall (in)	0.041	0.411	0	17.5
Snowfall indicator	0.038	0.191	0	1.0
Snow Depth				
Snow Depth (in)	0.176	1.160	0	23.0
Snow Depth Indicator	0.049	0.216	0	1.0



Source: (Melillo, Richmond, and Yohe (2014), http://www.c2es.org/science-impacts/national-climate-assessment)

Figure 1.1. Climate regions defined in National Climate Assessment

Table 1.2. Variables included in lasso analysis

Variable	X^2 and X^3
TMAX - Maximum Temperature	Yes
TMIN - Minimum Temperature	Yes
TAVG - Average Temperature	Yes
PRCP - Precipitation	Yes
SNOW - Snowfall	Yes
SNWD - Snow Depth	Yes
Season indicators (winter, spring, summer, fall)	No
Climate region indicators	No

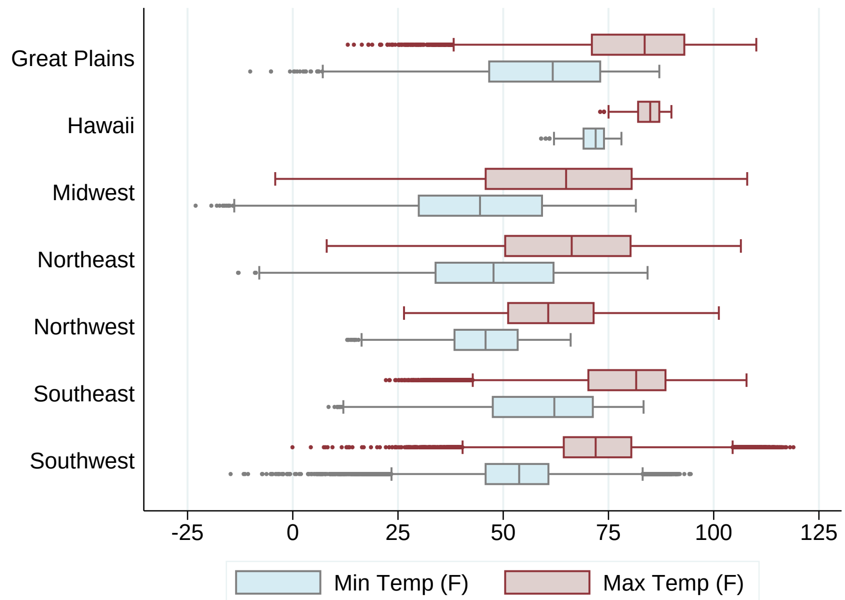


Figure 1.2. Temperature distributions by climate region

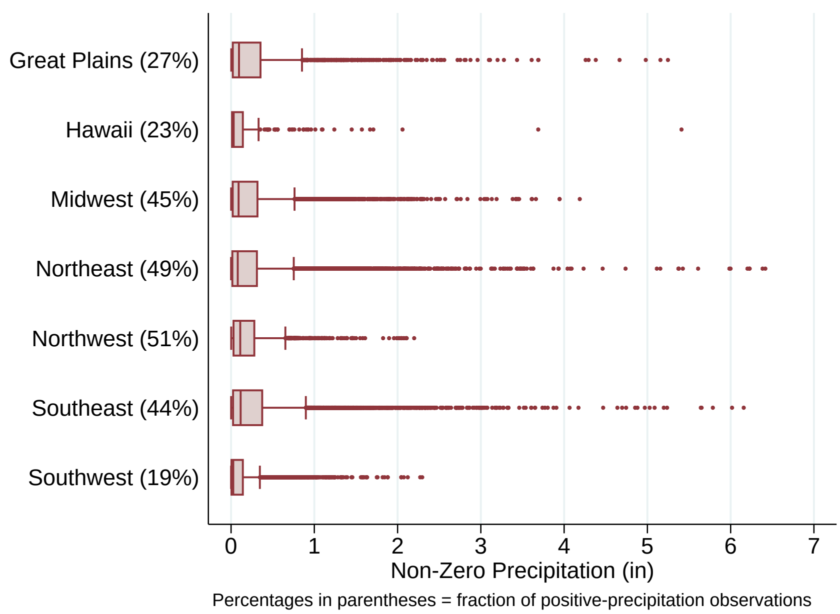


Figure 1.3. Distribution of precipitation by climate region

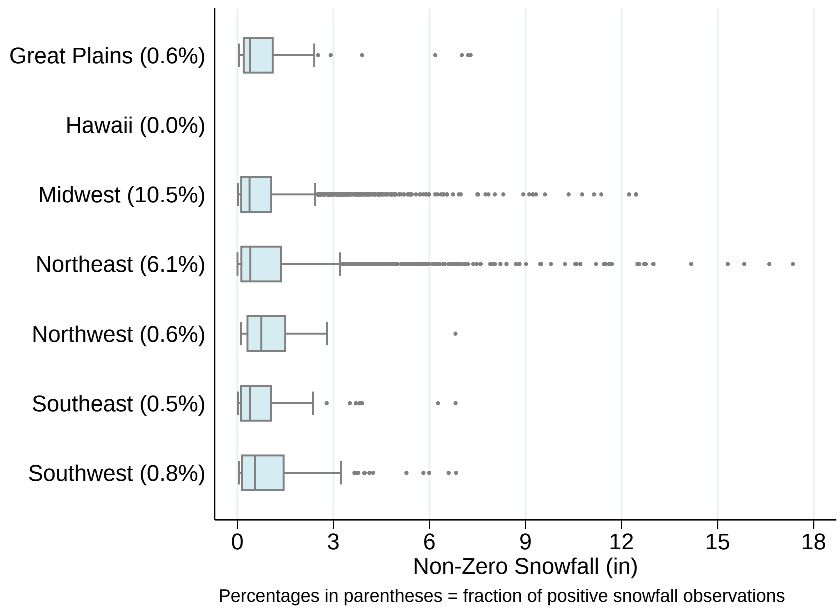


Figure 1.4. Distribution of snowfall by climate region

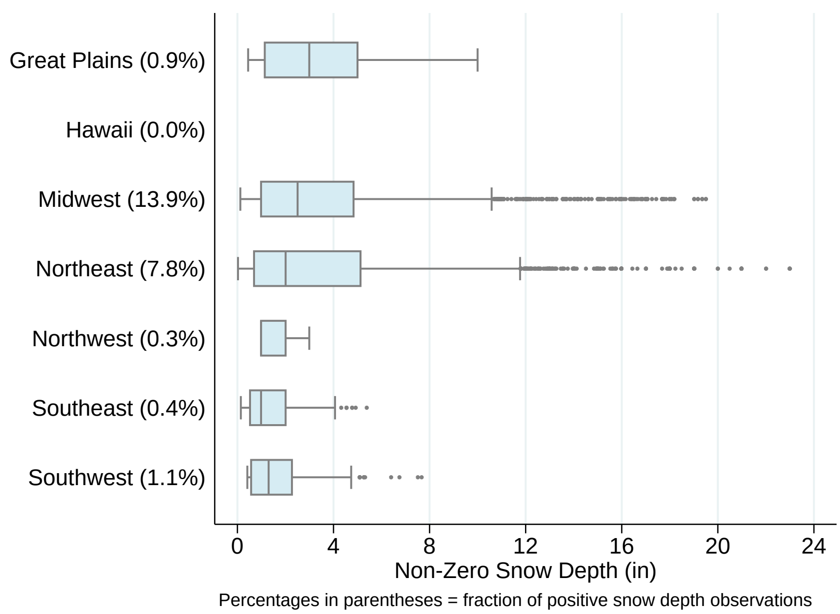


Figure 1.5. Distribution of snow depth by climate region

Table 1.3. Basic weather variable effects on ln(Sales)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Max Temp (F)		0.00260***					
Avg Temp (F)			0.00282***				
Min Temp (F)				0.00207***			
Precipitation (in)					-0.0455***		
Snowfall (in)						-0.129***	
Snow Depth (in)							-0.0262***
Observations	138148	137469	137380	137465	137747	138033	138008
Adjusted R^2	.8556	.8564	.8563	.856	.856	.8583	.8563

Notes: Results are clustered at MSA level. Regressions include year, month, day of week, holiday, store-trend, store-month, and store-day of week fixed effects. Controls also include indicators for store openings and closures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

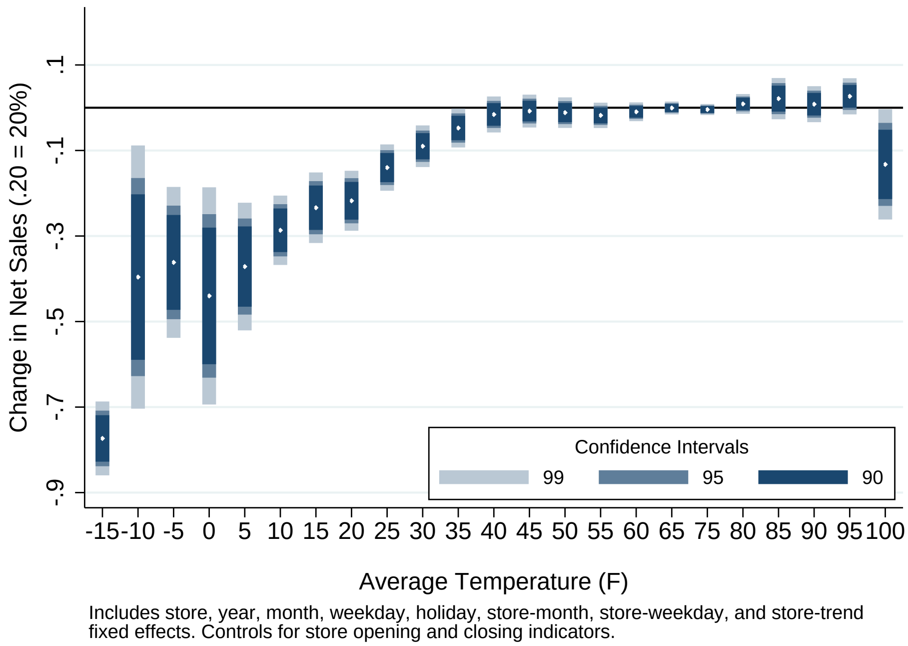
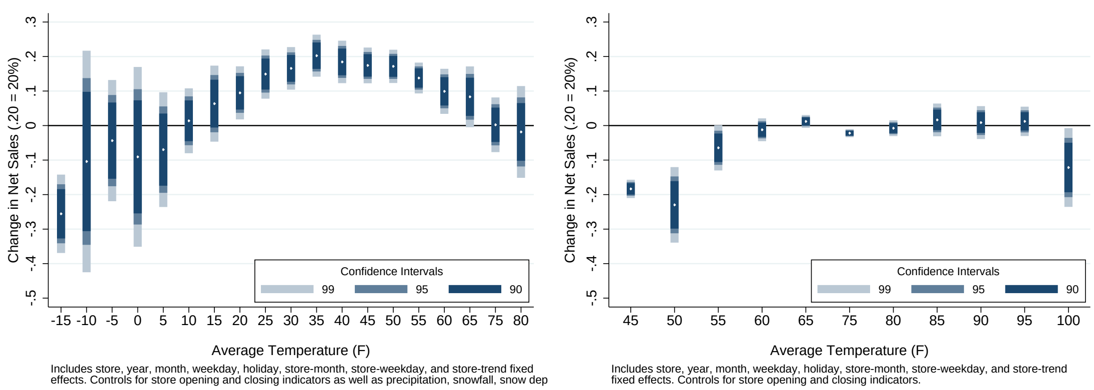
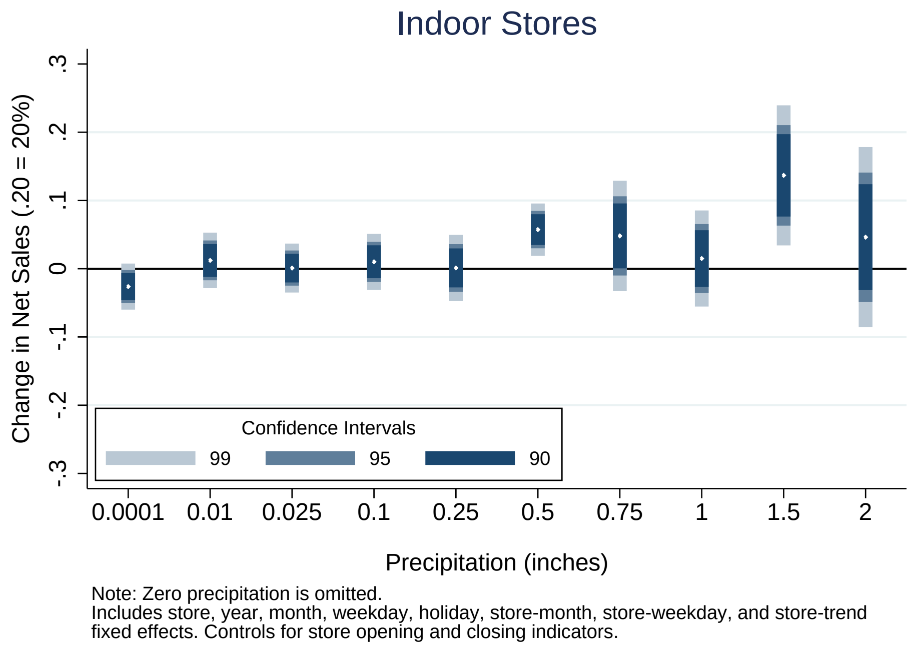


Figure 1.6. Nonlinear effect of temperature on sales

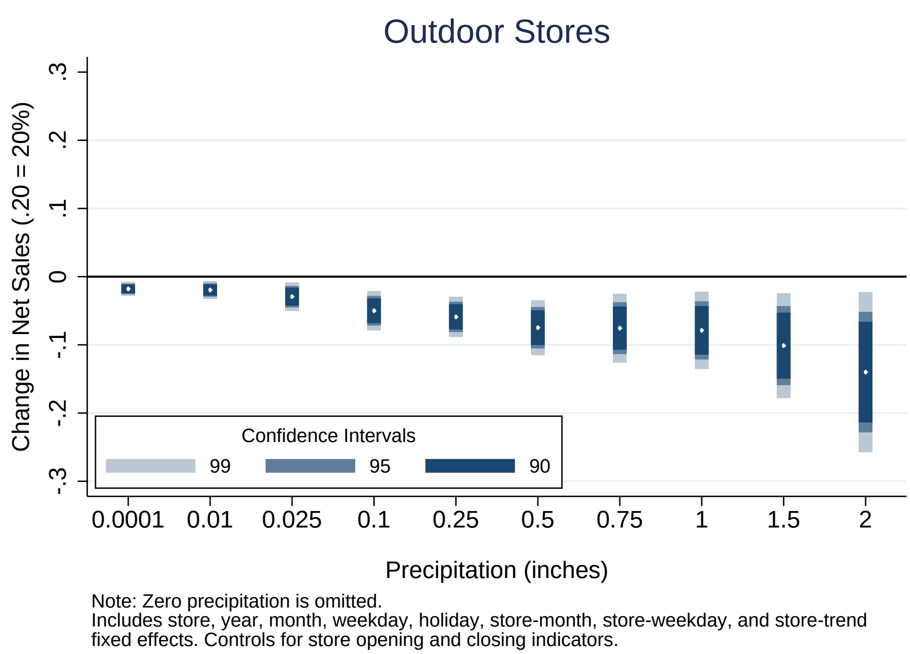


(a) Winter **(b)** Summer

Figure 1.7. Responses to average temperatures vary by season



(a)



(b)

Figure 1.8. Precipitation effects on sales by indoor and outdoor store locations

Table 1.4. Regionally heterogeneous effects of snowfall and snow depth on sales

	(1)	(2)
Snowfall (in)	-0.107***	
Great Plains × Snowfall (in)	-0.137***	
Midwest × Snowfall (in)	-0.050***	
Northwest × Snowfall (in)	-0.131***	
Southeast × Snowfall (in)	-0.106**	
Southwest × Snowfall (in)	-0.079***	
Snow Depth (in)		-0.024***
Great Plains × Snow Depth (in)		-0.036***
Midwest × Snow Depth (in)		-0.002
Northwest × Snow Depth (in)		-0.237***
Southeast × Snow Depth (in)		-0.105***
Southwest × Snow Depth (in)		-0.055***
Observations	138033	138008
Adjusted R^2	.8585	.8564

Notes: Northeast is base region (represented by coefficient on Snowfall). Results are clustered at MSA level. Regression includes year, month, day of week, holiday, store-trend, store-month, and store-day of week fixed effects. Controls also include indicators for store openings and closures.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.5. Adaptation regressions

	(1)	(2)	(3)	(4)	(5)
Max Temp (10F)	0.0357***				
Max Temp (10F) × Norm Max Temp (10F)	-0.0005**				
Min Temp (10F)		0.0149***			
Min Temp (10F) × Norm Min Temp (10F)		0.0006*			
Precipitation (in)			-0.130***		
Precipitation (in) × Norm Precip (in)			0.007***		
Snowfall (in)				-0.171***	
Snowfall (in) × Norm Snowfall (in)				0.013**	
Snow Depth (in)					-0.0343***
Snow Depth (in) × Norm Snow Depth (in)					0.0003**
Observations	124610	124606	124889	133891	134627
Adjusted R^2	.853	.8526	.8525	.8605	.8581

Notes: Results are clustered at MSA level. Regressions include year, month, day of week, holiday, store-trend, store-month, and store-day of week fixed effects. Controls also include indicators for store openings and closures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

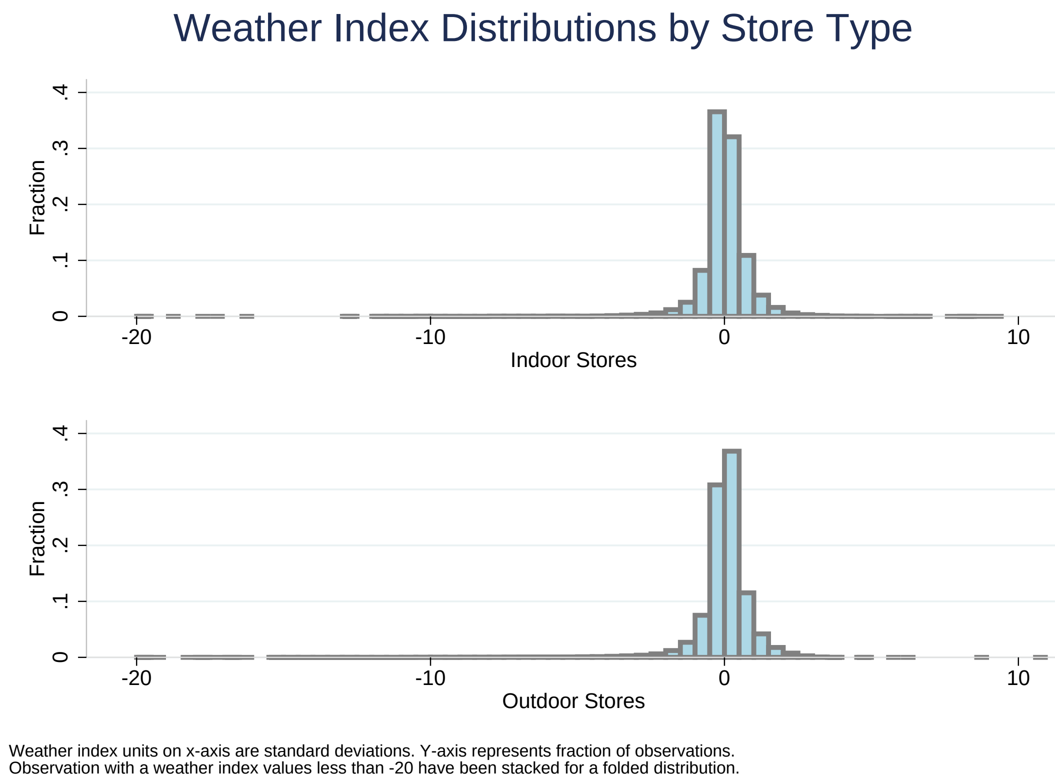
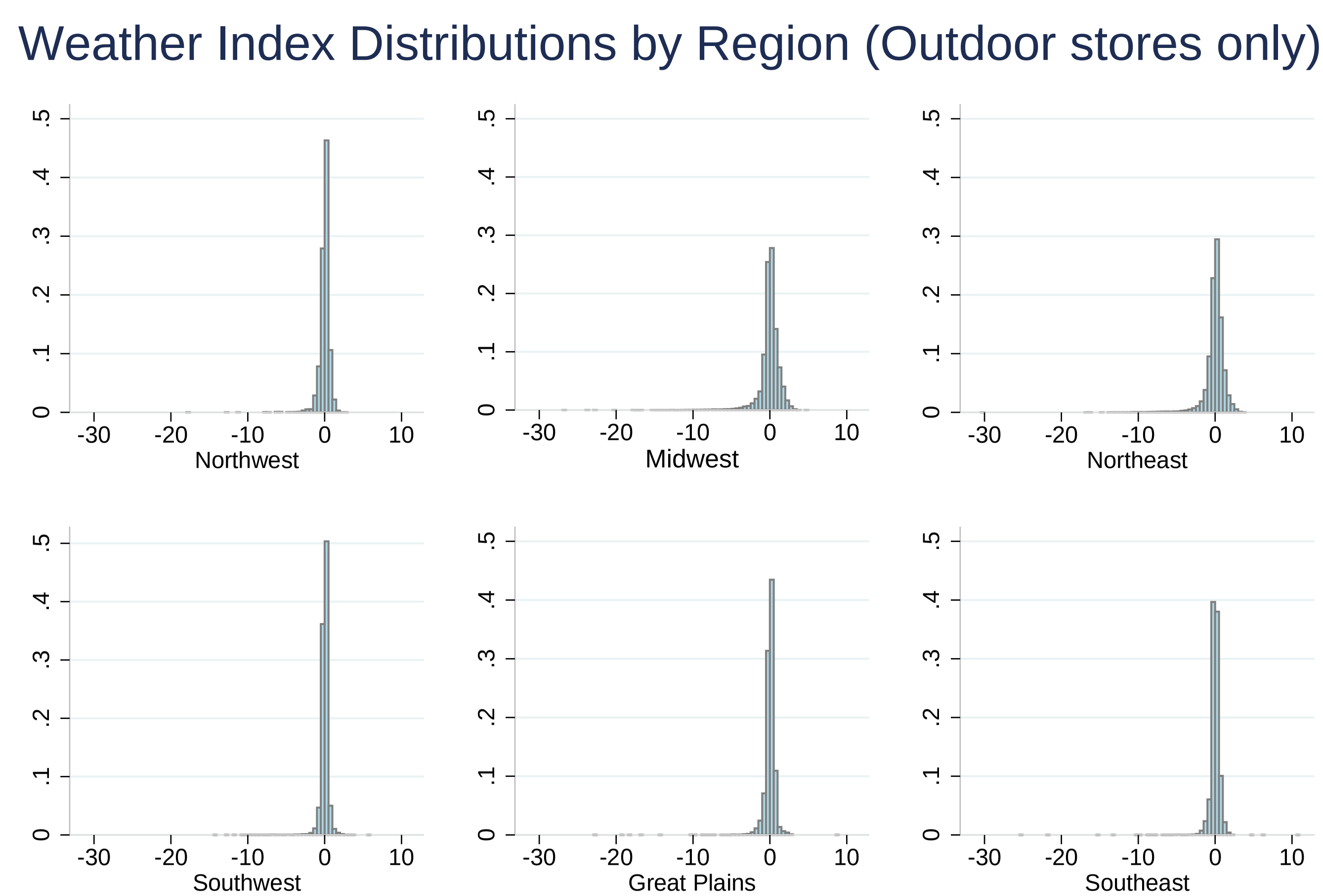


Figure 1.9. Distribution of weather indexes



Weather index units on x-axis are standard deviations. Y-axis represents fraction of observations.
Southeast excludes one observation with a weather index value of -60.14.

Figure 1.10. Distribution of weather indexes by region

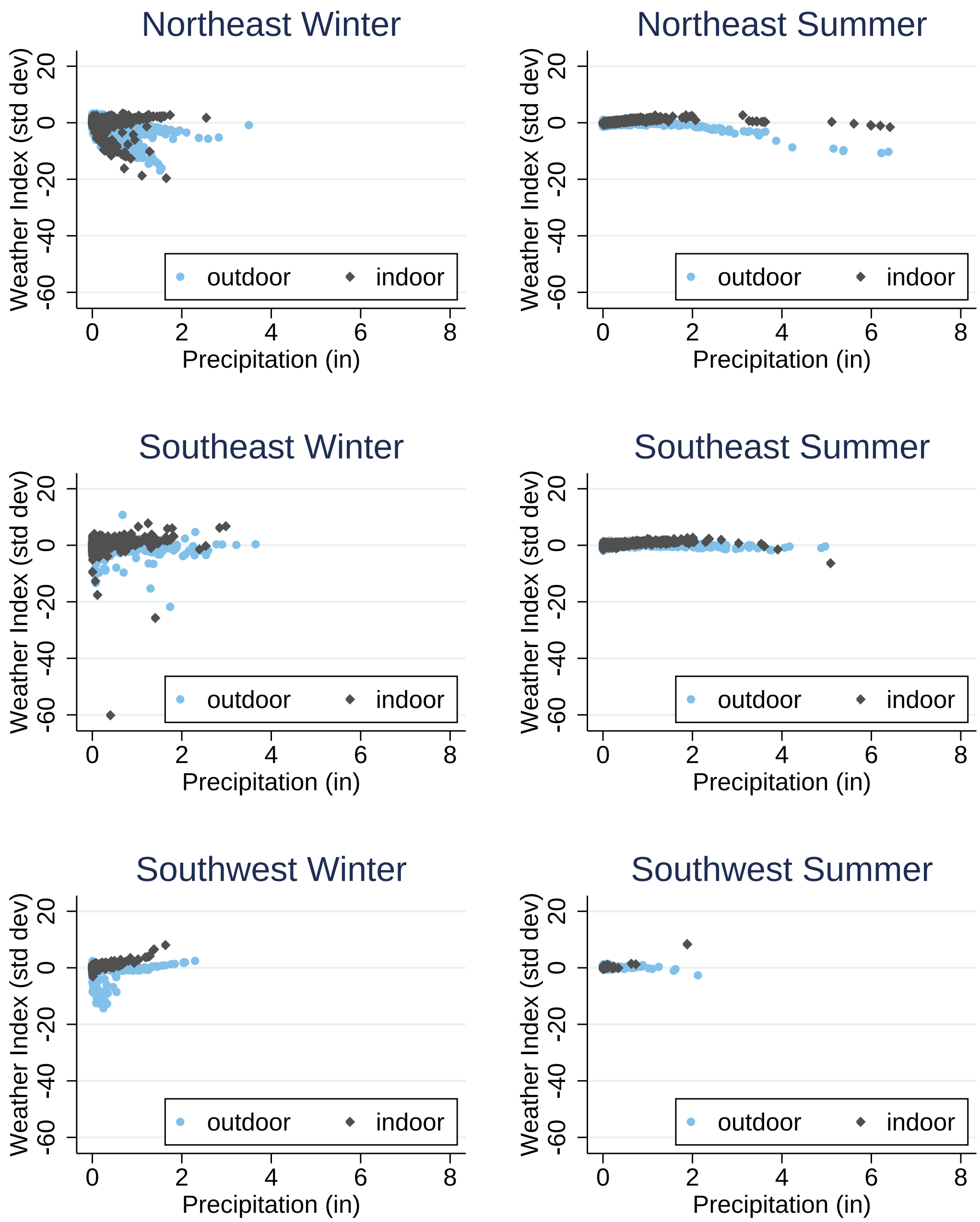


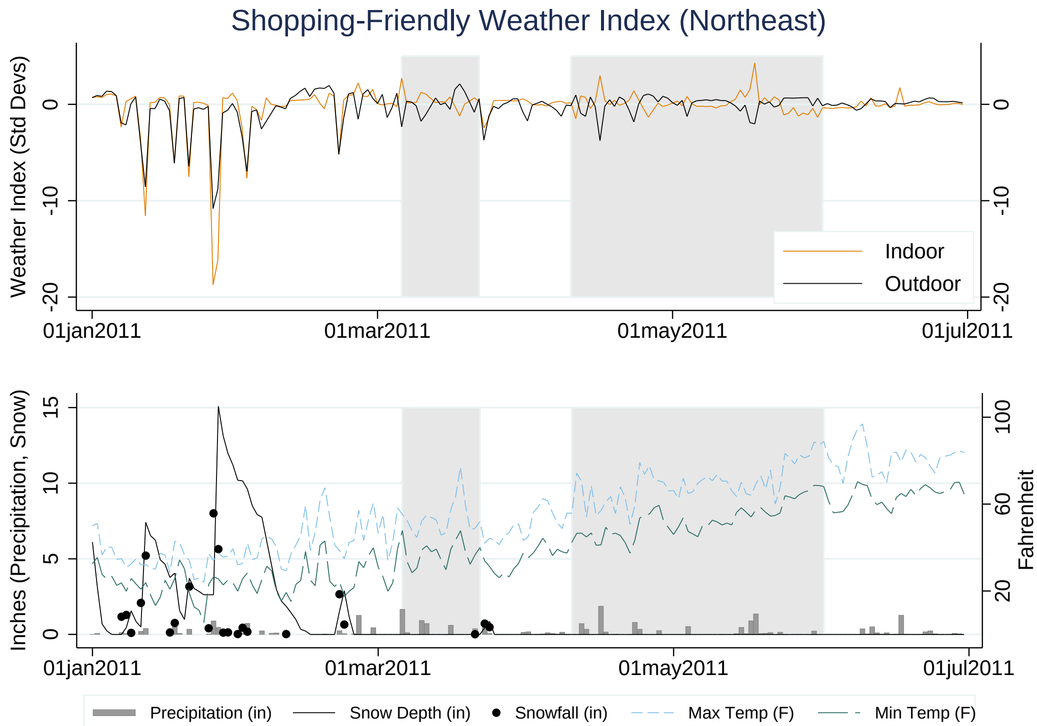
Figure 1.11. Weather index vs. precipitation by region, store type, and season

Table 1.6. Weather effects on daily sales based on weather index and quantiles of weather index

	All		Outdoor		Indoor	
	(1)	(2)	(3)	(4)	(5)	(6)
Weather Index	0.075***		0.081***		0.053***	
< 5th percentile		-0.199***		-0.217***		-0.124***
5th - 10th percentile		-0.065***		-0.069***		-0.048***
10th - 20th percentile		-0.038***		-0.041***		-0.025**
20th - 40th percentile		-0.019***		-0.020***		-0.017*
40th - 60th percentile		-		-		-
60th - 80th percentile		0.018***		0.018***		0.014
80th - 90th percentile		0.049***		0.048***		0.048**
90th - 95th percentile		0.088***		0.095***		0.057***
> 95th percentile		0.146***		0.158***		0.101***
Observations	136888	136888	108765	108765	28123	28123
Adjusted R^2	0.862	0.860	0.861	0.859	0.865	0.864

Notes: Results are clustered at MSA level. Regressions include year, month, day of week, holiday, store-trend, store-month, and store-day of week fixed effects. Controls also include indicators for store openings and closures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



Shaded areas indicate periods with consistent substitution between indoor and outdoor stores.

Figure 1.12. Indoor and outdoor weather indexes in a northeast metropolitan area

Table 1.7. Correlations between indoor and outdoor weather indexes

Subset	Correlation
All	0.518
Conditional on weather	
Positive precipitation	0.521
Positive snowfall	0.808
Positive snow depth	0.808
Regions	
Northeast	0.445
Northwest	0.130
Southeast	0.582
Southwest	0.548
Midwest	0.825
Great Plains	0.454
Seasons	
Winter	0.688
Spring	0.194
Summer	0.479
Fall	0.234
Northwest by season	
Winter	0.529
Spring	-0.194
Summer	0.769
Fall	-0.514

Correlations are between indoor and outdoor weather indexes by MSA. All correlations are significant at 99% level.

Table 1.8. Substitution between indoor and outdoor stores

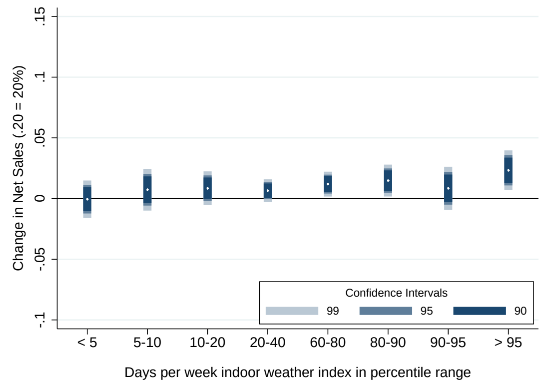
	(1)	(2)	(3)	(4)	(5)
Ln(Net Sales)					
Own weather index	0.0637***	0.0520***	0.0634***	0.0601***	0.0589***
Other weather index	-0.00220	0.00232	-0.00408	-0.00157	-0.00746**
Outdoor × Own weather index		0.0294***			
Outdoor × Other weather index		-0.0102			
winter × Own weather index			-0.00346		
summer × Own weather index			-0.00350		
fall × Own weather index			0.00847		
winter × Other weather index			0.00747		
summer × Other weather index			0.00266		
fall × Other weather index			-0.0146*		
it's raining=1 × Own weather index				0.0127	
it's snowing=1 × Own weather index				-0.00425	
it's snowed=1 × Own weather index				-0.0116	
it's raining=1 × Other weather index				0.00412	
it's snowing=1 × Other weather index				0.000644	
it's snowed=1 × Other weather index				-0.00420	
Great Plains × Own weather index					0.000490
Midwest × Own weather index					0.00670***
Northwest × Own weather index					0.0301***
Southeast × Own weather index					0.00460*
Southwest × Own weather index					-0.00174
Great Plains × Other weather index					0.000758
Midwest × Other weather index					0.00815**
Northwest × Other weather index					0.00311
Southeast × Other weather index					0.0117**
Southwest × Other weather index					0.00199
Observations	32036	32036	32036	32036	32036
Adjusted R^2	.943	.9431	.943	.943	.943

Notes: Observations are at MSA level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

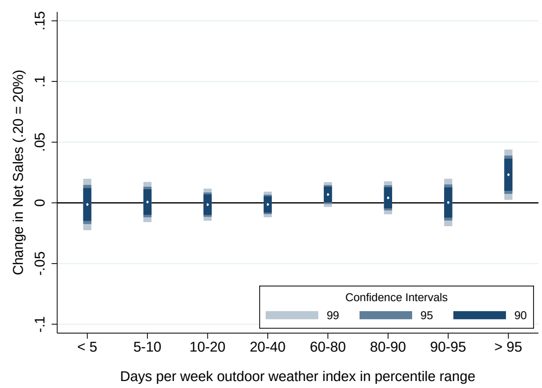
Table 1.9. Regressions of ln(e-commerce sales) on weather indexes for outdoor and indoor stores

	(1)	(2)	(3)	(4)	(5)
Indoor avg. weekly index		0.0261***			0.0260***
Outdoor avg. weekly index			0.0293***		0.0198*
Weekly mean indoor and outdoor weather index				0.0352***	
Indoor avg wkly index × Outdoor avg wkly index					0.0070**
Observations	2145	2145	2145	2145	2145
Adjusted R^2	.9637	.964	.9639	.964	.964

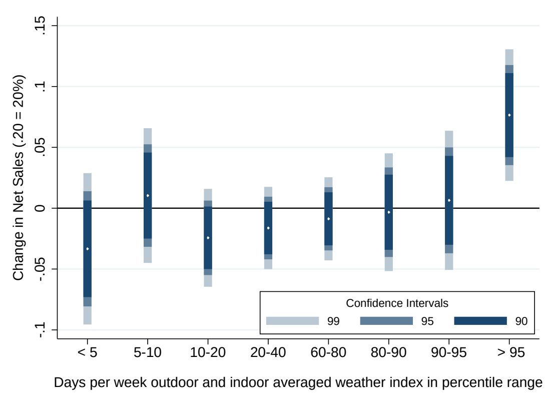
Notes: Observations are at MSA level. Regressions control for month, year, MSA, holiday, month-MSA, and MSA-trend fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



(a) Local MSA-level indoor index effects on e-commerce



(b) Local MSA-level outdoor index effects on e-commerce



(c) Local MSA-level mean indoor and outdoor index effects on e-commerce

Figure 1.13. Non-linear in-store weather index effects on e-commerce

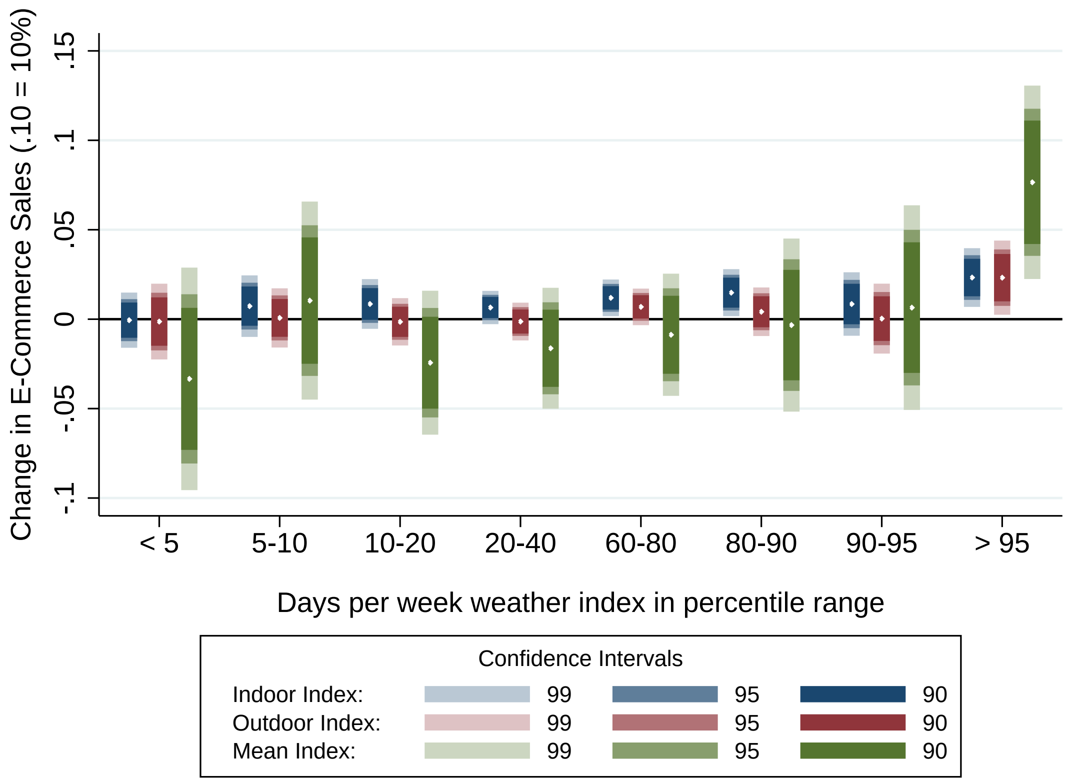


Figure 1.14. Indoor, outdoor, and mean in-store weather index effects on e-commerce

Table 1.10. Intertemporal weather index effects

	(1)	(2)	(3)	(4)	(5)
Ln(Net Sales)					
Weather Index	0.0717***	0.0717***	0.0683***	0.0683***	0.0660***
L.Weather Index	0.00997***		0.0100***	0.00844***	0.00910***
F.Weather Index		0.0116***	0.0108***	0.0102***	0.0102***
Sunday=1 × L.Weather Index				0.0112***	0.00677*
Saturday=1 × F.Weather Index				0.00528**	0.00450
Saturday=1 × Weather Index					0.00432
Sunday=1 × Weather Index					0.0134***
Observations	135000	135000	133148	133148	133148
Adjusted R^2	0.863	0.865	0.866	0.866	0.866

Notes: Results are clustered at MSA level. Regressions include year, month, day of week, holiday, store-trend, store-month, and store-day of week fixed effects. Controls also include indicators for store openings and closures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.11. Nonlinear intertemporal weather index effects

	(1)	(2)	(3)
$<$ 5th percentile$_{t-1}$	-0.0201**		-0.0228***
5th - 10th percentile$_{t-1}$	-0.0098*		-0.0112**
10th - 20th percentile$_{t-1}$	-0.0020		-0.0024
20th - 40th percentile$_{t-1}$	-0.0016		-0.0020
40th - 60th percentile$_{t-1}$	-		-
60th - 80th percentile$_{t-1}$	0.0027		0.0025
80th - 90th percentile$_{t-1}$	0.0068*		0.0050
90th - 95th percentile$_{t-1}$	0.0151**		0.0142**
$>$ 95th percentile$_{t-1}$	0.0355***		0.0309***
$<$ 5th percentile$_{t+1}$		-0.0418***	-0.0421***
5th - 10th percentile$_{t+1}$		-0.0064	-0.0089*
10th - 20th percentile$_{t+1}$		-0.0091**	-0.0094**
20th - 40th percentile$_{t+1}$		-0.0072**	-0.0076**
40th - 60th percentile$_{t+1}$		-	-
60th - 80th percentile$_{t+1}$		-0.0000	0.0004
80th - 90th percentile$_{t+1}$		0.0143***	0.0132***
90th - 95th percentile$_{t+1}$		0.0299***	0.0280***
$>$ 95th percentile$_{t+1}$		0.0390***	0.0336***
Observations	135000	135000	133148
Adjusted R^2	0.861	0.863	0.864

Notes: Results are clustered at MSA level. Regressions include current period weather index quantile bins as well as year, month, day of week, holiday, store-trend, store-month, and store-day of week fixed effects. Controls also include indicators for store openings and closures.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.12. Weekly store level sales

	(1)	(2)	(3)
ln(Weekly Net Sales)			
Mean index value	0.0888***	0.1241***	0.0734***
Min index value		-0.0114***	
Max index value		-0.0192***	
Mean weekend index			0.0550***
Observations	19824	19824	19824
Adjusted R^2	.9156	.9157	.9157

Notes: Observations are at weekly store-level. Regressions control for month, year, store, holiday, month-store, and store-trend fixed effects. Excludes weeks without 7 observations of weather index.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.13. Weekly store level sales with multiple lags and leads

	(1)	(2)	(3)
ln(Weekly Net Sales)			
Mean index value	0.0886***	0.1270***	0.0753***
L7.Mean index value	0.0013	0.0195**	0.0183***
L14.Mean index value	-0.0086***		
L21.Mean index value	-0.0071**		
L28.Mean index value	0.0069**		
F7.Mean index value	0.0175***		
F14.Mean index value	-0.0160***		
F21.Mean index value	0.0214***		
F28.Mean index value	0.0046		
Min index value		-0.0120***	
Max index value		-0.0202***	
L7.Min index value		-0.0077**	
L7.Max index value		-0.0131*	
Mean weekend index			0.0467***
L7.Mean weekend index			-0.0679***
Observations	19057	19810	19810
Adjusted R^2	.9188	.9161	.9162

Notes: Observations are at weekly store-level. Regressions control for month, year, store, holiday, month-store, and store-trend fixed effects. Excludes weeks without 7 observations of weather index.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.14. Monthly store level sales

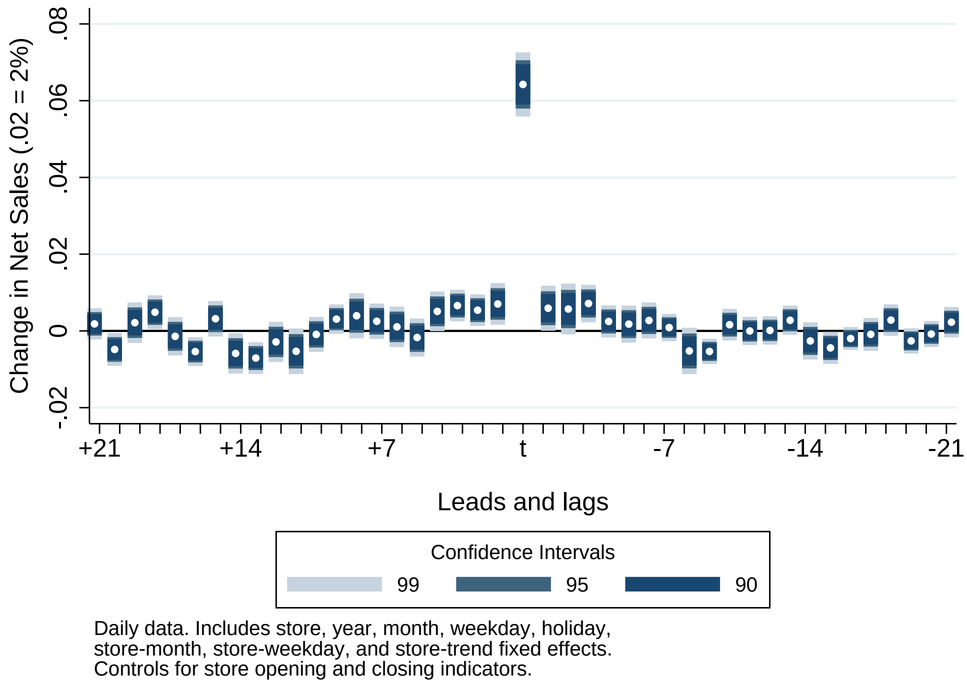
	(1)	(2)	(3)
Ln(Monthly Net Sales)			
Average index	0.0766***	0.0604***	0.0645***
Lowest index value		0.0048***	
Highest index value		-0.0051	
Mean weekend index			0.0477
Observations	4100	4100	4100
Adjusted R^2	.9471	.9473	.9471

Notes: Observations are monthly store-level. Regressions control for number of weekend days per month, days of sales observations per month, month, year, store, holiday, season-store, and store-trend fixed effects. Excludes store opening and closing months and months without at least 2 missing weather index observations or store sales.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

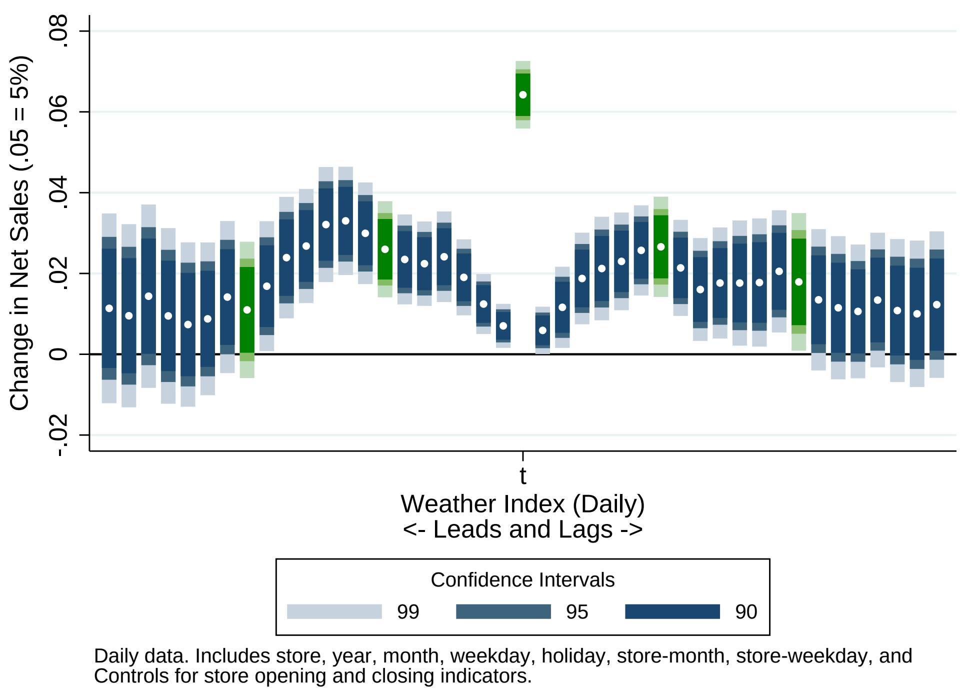
Table 1.15. Monthly store level sales with lags

	(1)	(2)	(3)
Ln(Monthly Net Sales)			
Average index	0.0625***	0.0440***	0.0418**
L.Average index	0.0015	-0.0116	-0.0200
Lowest index value		0.0060***	
Highest index value		-0.0018	
L.Lowest index value		0.0039*	
L.Highest index value		0.0022	
Mean weekend index			0.0745
L.Mean weekend index			0.0946**
Observations	3704	3704	3704
Adjusted R^2	.9464	.9466	.9464

Notes: Observations are monthly store-level. Regressions control for number of weekend days per month, days of sales observations per month, month, year, store, holiday, season-store, and store-trend fixed effects. Excludes store opening and closing months and months without at least 2 missing weather index observations or store sales.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



(a) Daily lead and lag weather index effects on sales



(b) Cumulative daily weather index effects on sales

Figure 1.15. Daily intertemporal weather effects over three weeks

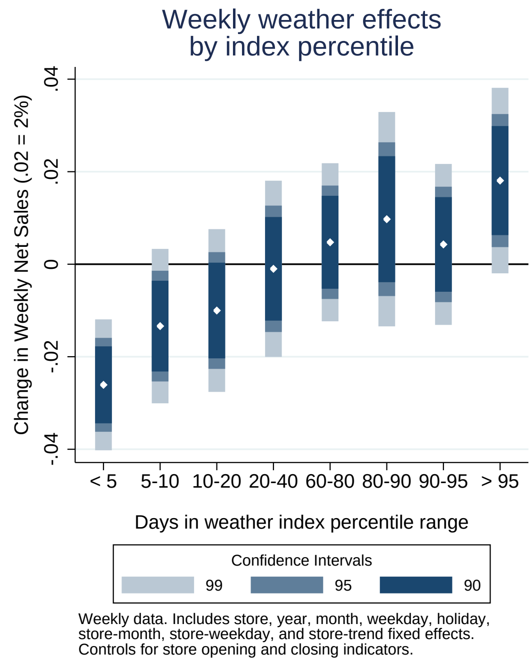


Figure 1.16. Weekly weather effects by percentile

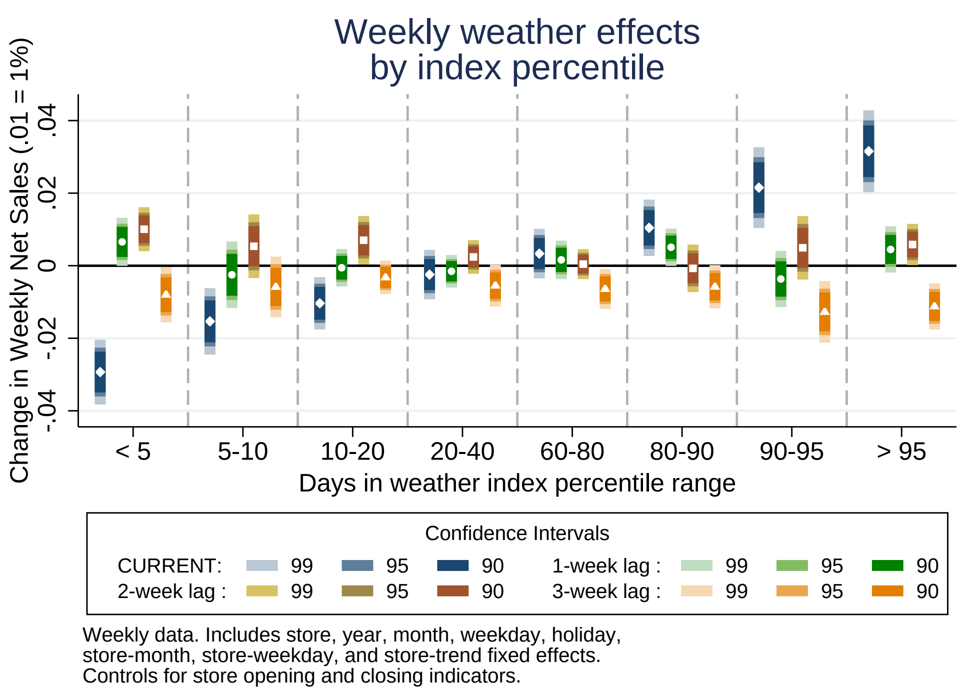


Figure 1.17. Non-linear intertemporal substitution by week

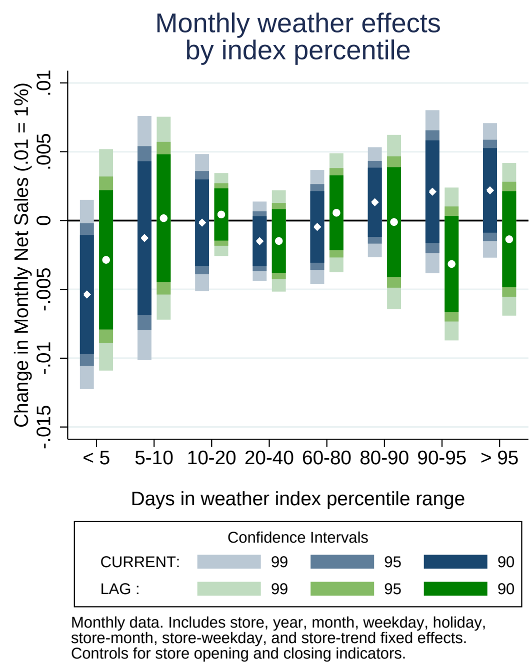


Figure 1.18. Monthly lags

Chapter 2

Divest, Disregard, or Double Down?

How should a philanthropic foundation's endowment invest in a firm whose activities run directly counter to the charitable missions the endowment funds? Any foundation seeking to combat, reduce, or eliminate something like pollution, tobacco consumption, gun violence, child labor, or nuclear power that can be tied to the activities of a firm must answer this question. Universities and colleges considering investments in sectors which relate to their research, policy, or teaching focuses may also face this question.[1] With ethicists, legal scholars, investors, journalists, governments, activists, universities, and foundation and endowment managers all weighing in, two main opposing strategies have emerged: socially responsible investing (SRI), which calls for divestment of objectionable assets, and the "firewall" approach of disregarding the objectionable nature of firms in investment decisions.[2] In this paper I present a new strategy, which I call "mission hedging," wherein the endowment doubles down on its investment in the firm it opposes. If increased objectionable activities coincide with both higher firm

[1]There are some differences in tax treatment of college and university foundations and endowments which are not important from the perspective of this paper.

[2]Shareholder activism is an alternative middle-ground approach in which investments are used to submit and vote on shareholder proposals that influence firms directly. Due to Securities and Exchange Commission rules, a foundation only has to own $2000 in market value of the firm's securities (continuously for one year) in order to submit a proposal to be voted on by all shareholders. (U.S. Security and Exchange Commission (1998)). Thus shareholder activism would be an additional benefit of investing in a firm but is not expected to motivate a large investment level.

Consider the example of People for the Ethical Treatment of Animals (PETA), which sought to submit activist shareholder proposals. It purchased just $2274 worth of shares in SeaWorld in April 2013 along with additional shares the next year after prices declined (Smith (2013), Sweatte (2014)).

returns and greater foundation needs for revenue (to counteract the firm's activities), then greater exposure to the firm aligns funds available with funds needed. This creates a hedge around the foundation's mission and maximizes its expected utility.

Classic arguments in the divestment debate are as follows. SRI proponents divest to advance missions directly through investing (Kramer and Cooch (2007)). Exclusion of objectionable firms from portfolios avoids implicit shows of support for the firms or benefiting from tainted profits, makes political statements, and potentially exerts downward price pressure on stocks to influence firm behavior.[3] Firewall foundations instead disregard interactions between missions and firms, keeping figurative firewalls between investing and charitable operations. They task money managers with maximizing risk-adjusted financial returns to yield the largest possible operating and grantmaking budgets with which to directly do good works. Firewall advocates believe any potential divestment benefits are outweighed by the costs, which include lower risk-adjusted returns from choosing investments based on non-financial factors, higher risks from a less diversified portfolio, and administrative burdens of implementation.[4]

Unfortunately, these approaches both ignore covariance between a firm's financial returns and activities related to the foundation's mission. The mission hedging strategy leverages this covariance by skewing investment toward the firm, making the foundation's mission outcome more certain much as traditional hedging makes financial outcomes more certain. Thus some endowments will want higher than normal exposure to stocks

[3]Casual conversation with non-economists suggests that some people mistakenly believe that spending money to buy shares in a company (even on the secondary market) is equivalent to giving the company that money. A more sophisticated view that stock purchases can lower a company's borrowing cost or reward executives through increasing stock option value, however, can lead to similar concerns. Evidence for this is discussed in section 2.3.1

[4]Other relevant concerns have historically included that foundation managers' fiduciary duties might prohibit them from engaging in divestment activities that could potentially lower returns. American legal scholars have argued that it is at the very least legally acceptable for trustees to engage in social investing and perhaps the prudent course (Solomon and Coe (1997a), Solomon and Coe (1997b), and McKeown (1997)). Recently students sued Harvard University over its failure to divest, claiming that fossil fuel investments violated fiduciary duties. The suit was dismissed by the court (Klein and Delwiche (2015).)

that reward the behaviors they are fighting. For example, a lung cancer-fighting foundation could benefit from investing even more heavily in tobacco than a standard portfolio would.

This paper relates to a number of strands in the literature. Public finance scholars incorporate covariance in social project valuations. For example, Hirshleifer (1966) argues that one must account for funds being valued more in some states than others when evaluating government projects under uncertainty. Minken (2008) shows that in public project cost benefit analyses the risk premium component of discount rates depends on the covariance of project returns with returns on all national assets. On the corporate and personal finance side, the Consumption Capital Asset Pricing Model (C-CAPM) builds on the basic CAPM to allow expected return to decrease with covariance between asset returns and marginal utility of consumption (Blanchard and Fisher (1989), p. 507-508). Others have also proposed hedging on non-purely financial dimensions. For example, Wolfers and Zitzewitz (2006) suggest that individuals use prediction markets to hedge personal risks. Particularly, an individual could bet on a candidate whose election might increase the individual's risk of job loss. However, no one has presented a formal model of this mechanism or the trade-offs involved, especially in the institutional context.

A large body of literature has focused on determining whether SRI is generally beneficial in terms of financial returns and influencing firm behavior.[5] Charged with a charitable mission instead of profit-maximization, a philanthropic endowment's objective function is inherently different from other investors', and thus yields different trade-offs in the SRI decision. This institutional perspective has not yet been formally modeled. The literature on divesting as an SRI strategy specifically for philanthropic endowments has primarily surveyed or advocated for (or against) the practice (see McKeown (1997), Wood and Hagerman (2010), Emerson (2003), and Kramer and Cooch (2007)). This

[5]See sections 2.3.1 and 2.3.3 for a discussion.

paper takes the first step toward developing a formal theory by presenting an endowment objective function and using the CAPM to decompose the trade-offs of hedging around idiosyncratic risk.

I start with a basic model of a foundation (or educational institution) that seeks to reduce a “bad” activity level and chooses how to allocate its endowment given assets’ random return distributions and known subjective levels of “evil.” After the investment is made, an exogenous shock simultaneously affects the bad activity level and asset returns. The foundation then spends its endowment to reduce the activity that it considers an economic bad.

I examine how a marginal shift in portfolio weights between two assets changes expected utility. I use the Capital Asset Pricing Model (CAPM) to decompose the effects of this shift into a set of trade-offs between expected returns, exposure to market-wide risk, and an idiosyncratic risk component. This final component (typically minimized through diversification) increases expected utility when endowment managers increase portfolio weights on firms whose returns correlate with activities the foundation seeks to reduce. In particular, a foundation with decreasing marginal utility or a proportional intervention technology can increase expected utility by skewing investment toward the firm, yielding second order stochastic dominance.

Despite the potential mission hedging benefits, foundations do not generally skew investments toward objectionable firms. To see why, I model the optimization problems implicitly solved by the SRI and firewall strategies. I allow portfolio evil levels to directly change fundraising, pre-intervention bad activity levels, and utility. I also assume that firm returns are independent of bad activity levels. I relax this assumption again to establish a comprehensive mission hedging model. I end by examining evidence for the comprehensive model trade-offs. I find little consensus on magnitudes, which depend on factors like endowment size, firm fundamentals, and foundation funding. I conclude that while mission hedging must be evaluated on a case-by-case basis, there will be foundations for whom the strategy is optimal.

To understand what is at stake, consider that more than 90,000 U.S. private foundations own over $640 billion in combined assets (Internal Revenue Service (2011)), while non-profit colleges and universities manage over $500 billion in assets (Berner (2015)). Foundations and universities with environmental focuses have recently joined the fossil fuel divestment movement, which now exceeds $50 billion in committed assets (Goldenberg (2014) and Wines (2014)).

The remainder of the paper is structured as follows. Section 2.1 lays out the basic mission hedging model, identifying trade-offs in the investment decision and demonstrating conditions under which skewing investments toward objectionable firms is optimal. Section 2.2 adds assumptions and elements consistent with the divestment and firewall strategies and presents a comprehensive mission hedging model. Section 2.3 explores the theoretical and empirical evidence related to the trade-offs. Section 2.4 discusses market reactions and potential feedbacks from mission hedging. Section 2.5 concludes with suggestions for potential tests of empirical implications of the model presented here and finally some summary remarks.

2.1 Mission Hedging Model

I consider a foundation established with a mission to reduce or eliminate some bad activity, the initial level of which is denoted by b_0.[6] This foundation begins with initial endowment a_0, which it invests as it chooses in a set of assets. The world then experiences a shock which simultaneously affects the bad activity level (now b_1) and asset returns, yielding endowment value a_1. After the state of the world has been revealed

[6] I will use the term foundation throughout the remainder of this analysis. However, as described in the introduction, the results apply to other entities like some educational institutions.

and the endowment has earned its returns, the foundation spends a_1 on its intervention to reduce b_1.

In investing its endowment, the foundation chooses an asset allocation consisting of a set of weights $\boldsymbol{\alpha} = \{\alpha_i\}$ for each possible asset i, where an asset may be an individual security or a collection of securities like a fund. It must always hold that

$$\sum_{i=1}^{n} \alpha_i = 1. \tag{2.1}$$

Each asset is characterized by the distribution of its random return r_i and a measure of evil e_i, a known static scalar that is subjectively determined by the foundation.[7] Thus an endowment's portfolio return (r_p) and evil level (e_p) are:[8]

$$r_p = \sum_{i=1}^{n} \alpha_i r_i \qquad \text{and} \qquad e_p = \sum_{i=1}^{n} \alpha_i e_i. \tag{2.2}$$

The foundation spends its assets on an intervention. I model this intervention technology as

$$y(a_1, b_1) = b_2. \tag{2.3}$$

where a_1 and b_1 are post-shock, pre-intervention endowment and bad activity levels, respectively, and b_2 is the final bad activity level. This intervention has the following

[7]The same asset allocation may be considered evil by one foundation and good by another. For example, a foundation opposed to abortion would consider an abortion pill producer objectionable, while a pro-reproductive rights foundation might favor it.

[8]For simplicity I have specified the portfolio level of evil to equal the weighted average of individual asset evil levels. However, it could take a different form so that $e_p = g(\boldsymbol{\alpha}, \boldsymbol{e})$, where $\boldsymbol{e} \equiv [e_1 ... e_n]'$. The functional form does not play a role in the basic model presented in this section. However, one can imagine manager and donor preferences (as included in section 2.2) that correspond to portfolio evil measure of, for example, $e_p = \max e_i$.

derivatives:

$$\begin{gathered}\frac{dy}{da_1} < 0, \quad \frac{d^2y}{da_1^2} > 0, \\ \frac{dy}{db_1} > 0, \quad \frac{d^2y}{db_1^2} \leq 0, \\ \text{and } \frac{d^2y}{da_1 db_1} \leq 0.\end{gathered} \tag{2.4}$$

The production function $y(\cdot)$ yields final bad activity level b_2. Thus a negative first derivative dy/da_1 means that a foundation's ability to decrease the bad activity level increases with final endowment value, though at a decreasing rate due to positive second derivative d^2y/da_1^2. A higher initial bad activity level increases the post-intervention bad activity level, as indicated by the positive first derivative of dy/db_1. However, this increase may have a constant scale (if $d^2y/db_1^2 = 0$ and $d^2y/da_1db_1 = 0$) or be proportional (if the second derivatives are strictly negative, meaning that a higher initial bad activity yields a higher absolute decrease in final bad activity.)

Figure 2.1 depicts two alternate intervention technologies fitting these characteristics. The bottom curve depicts an intervention with decreasing returns in a_1 given a lower initial bad activity level b_{1L}. The dashed middle curve shows the corresponding final bad activity level curve given a higher initial bad activity level of b_{1H}, if the intervention technology is proportional. In this example, the intervention yields a given fractional reduction in b_1, where this fraction increases at a declining rate with a_1. This corresponds to $\frac{d^2y}{db_1^2} < 0$ and $\frac{d^2y}{da_1db_1} < 0$. The dotted top curve depicts the final bad activity level given the higher initial bad activity level of b_{1H} and a non-proportional intervention technology. Here the absolute reduction in the bad activity level for a given a_1 is independent of the initial bad activity level. Thus the two non-proportional curves have the same slope at each a_1 with $\frac{d^2y}{db_1^2} = 0$ and $\frac{d^2y}{da_1db_1} = 0$.

In the end, the foundation seeks to maximize expected utility, which in this basic set-up simply decreases in the final bad activity level. (In section 2.2.1, the possibility that the portfolio's level of evil enters into the utility function directly is considered.) Specifically,

$$\max_{\alpha} E\left[U\left(b_2\right)\right] = \max_{\alpha} E\left\{U\left[y\left(a_1, b_1\right)\right]\right\}, \tag{2.5}$$

where

$$a_1 = a_0(1 + r_p) \tag{2.6}$$

are the assets available for the intervention and b_1 is the bad activity level realized after the shock that simultaneously affects r_p. Furthermore, $dU/db_2 < 0$ and $d^2U/db_2^2 \geq 0$. Optimality of mission hedging, or doubling down on an investment in a bad firm, follows given the additional assumption that the derivative products are as follows:

$$\begin{gathered} U_A \equiv \frac{dU}{da_1} = \frac{dU}{dy}\frac{dy}{da_1} > 0, \quad U_A^2 \equiv \frac{d^2U}{da_1^2} > 0, \\ U_B \equiv \frac{dU}{db_1} = \frac{dU}{dy}\frac{dy}{db_1} < 0, \quad U_B^2 \equiv \frac{d^2U}{db_1^2} < 0, \\ \text{and } U_{AB} \equiv \frac{d^2U}{da_1 db_1} > 0. \end{gathered} \tag{2.7}$$

These derivative conditions can be satisfied through decreasing marginal utility in the bad activity (which is equivalent to increasing marginal damages) and/or a proportional intervention technology. Pollution reduction and habitat loss prevention are examples of foundation missions that could yield decreasing marginal utility in the bad activity.[9] A marketing campaign that causes a certain fraction of smokers to quit is an example of a proportional intervention technology. The more money available for the marketing campaign, the greater is the fraction of smokers who are affected. The higher the number of smokers, the more people the campaign can help with the same amount of money. Decreasing marginal utility and proportional intervention technology both make the same dollar more valuable to the foundation on the margin when pre-intervention bad

[9]For intuition consider habitat loss, where the first portion of a species' habitat lost results in very low damages. But the marginal damages due to the loss of the last remaining bit of habitat are extremely high.

activity levels are higher (see Figure 2.2.)

A fully expanded version of this optimization problem is

$$\max_{\alpha} E\left\{U\left[y\left(a_0\left(1+\sum_{i=1}^{n}\alpha_i r_i\right), b_1\right)\right]\right\}. \tag{2.8}$$

Now consider a marginal shift between assets j and k, which is essentially a change in weights α_j and α_k:[10]

$$dEU = E\left\{\frac{dU}{dy}\frac{dy}{da_1}\left[a_0(d\alpha_j r_j + d\alpha_k r_k)\right] + \frac{dU}{dy}\frac{dy}{db_1}\left(\frac{db_1}{d\alpha_j} + \frac{db_1}{d\alpha_k}\right)\right\}. \tag{2.9}$$

For the time being assume that $\frac{db_1}{d\alpha_j} = \frac{db_1}{d\alpha_k} = 0$, so that the investment itself does not directly affect the pre-intervention bad activity level. (This assumption is relaxed in section 2.2.1.) Equation (2.9) can thus be simplified to the following:

$$dEU = dE\left\{\frac{dU}{dy}\frac{dy}{da_1}\left[a_0(d\alpha_j r_j + d\alpha_k r_k)\right]\right\}. \tag{2.10}$$

Substitute $d\alpha_j = -d\alpha_k$, which must hold in order to preserve equation (2.1). This yields the following:

$$\frac{dEU}{d\alpha_j} = E\left\{U_A\left[a_0(r_j - r_k)\right]\right\}. \tag{2.11}$$

Recall that $Cov(X,Y) = E[XY] - E[X]E[Y]$. Then:

$$\frac{dEU}{d\alpha_j} = a_0 Cov\left(U_A, r_j - r_k\right) + a_0 E\left[U_A\right] E\left[r_j - r_k\right] \tag{2.12}$$

The first term on the right-hand side of equation (2.12) shows that to the extent that asset j's returns have higher covariance with marginal utility, a shift from asset k

[10] This step and that in equation (2.10) follow the methodology in Shalit and Yitzhaki (1994).

to asset j increases expected utility. In other words, under the right circumstances, a foundation can benefit from skewing its investment toward assets that correlate with the foundation's mission-based need. The second term indicates that this increase may be augmented or offset by differences in expected returns, weighted by expected marginal utility of assets.

I now decompose these effects by applying the CAPM, according to which the expected return of a firm can be expressed as follows:

$$E[r_i] = r_f + \beta_i \left(E[r_m] - r_f\right). \tag{2.13}$$

where r_f is the risk-free rate, r_m is the market return, β_i is the beta coefficient such that $\beta_i = \mathrm{Cov}(r_i, r_m)/\sigma_m^2$, and σ_m^2 is the variance of market return.

The realization of firm i's returns are

$$r_i = r_f + \beta_i \left(r_m - r_f\right) + \varepsilon_i. \tag{2.14}$$

Thus according to the CAPM, an asset's realized return is the sum of the risk-free rate, a risk premium that accounts for the asset's exposure to market-wide risk, and an asset-specific or idiosyncratic risk realization ε_i.

Plugging equation (2.14) into equation (2.12) and without loss of generality assuming that $a_0 = 1$, the marginal effect of a shift from asset k to asset j is

$$\begin{aligned} \frac{dEU}{d\alpha_j} =& E[U_A] E\left[r_j - r_k\right] + \left(\beta_j - \beta_k\right) Cov\left(U_A, \left(r_m - r_f\right)\right) \\ & + Cov\left(U_A, \varepsilon_j - \varepsilon_k\right). \end{aligned} \tag{2.15}$$

The trade-off to expected utility between two assets is therefore a combination of three different effects. The first is expected returns, which may augment or offset the other two covariance effects. The second is covariance between foundation marginal utility and market risk. To the extent market returns are positively correlated with foundation marginal utility of assets, a higher β_i will increase expected utility. For example, a foundation focused on job training programs for the unemployed might prefer a low β_i if its needs are greater when the economy and markets are doing worse. On the other hand, a foundation focused on preserving open spaces from development might prefer a higher β_i if development accelerates in bull markets.

Covariance between foundation marginal utility and idiosyncratic risk is the third component. I call this the *idiosyncratic risk trade-off*. Consider, for example, shifts in smoking rates that affect a lung cancer fighting foundation and a tobacco company's returns but not necessarily the broader economy. This trade-off component lies at the heart of the mission hedging strategy to double down on investments in firms whose returns are correlated with bad activity levels. As depicted in figure 2.2b, marginal utility of assets a_1 increases with b_1 (as shown by the higher slope on the bottom curve for a given a_1) so that this covariance term will be positive if asset j's returns are more positively correlated with b_1 than asset k. All other things being equal, shifting to an asset whose idiosyncratic risk has a higher covariance with U_A yields second order stochastic dominance in expected utility. Although investors typically seek to eliminate this idiosyncratic risk through diversification, a foundation can benefit from taking more of it on when this risk is properly aligned with the foundation's mission-determined states of high marginal utility of assets.

In the optimal asset allocation, either these three trade-off components sum to zero or there is a corner solution. If one asset had an extremely high covariance between marginal utility of assets and idiosyncratic risk returns but is in other respects very similar to other assets, then it is conceivable that a foundation could maximize its expected utility by investing completely in that asset. What kind of asset might yield such a covariance?

A firm whose business activity is closely intertwined with the foundation's targeted bad activity may yield returns that covary positively with bad activity levels. And often these firms are considered to be evil incarnate. For example, foundations working on global warming, lung cancer, or animal welfare could, respectively, consider fossil fuel producers, tobacco companies, or meat producers to be objectionable. My result implies that such endowments should skew investment toward these seemingly reprehensible companies unless the expected return from such investments is so low relative to other opportunities that no investment in the objectionable firm is warranted.

It follows that foundations that do not account for covariance between idiosyncratic risk and marginal utility of assets will generally under-invest in high covariance assets. Because objectionable firms are more likely to have such covariance, firewall foundations will underinvest in these firms by disregarding the mission in the investment process. SRI foundations will tend to underinvest in these firms even more by avoiding them altogether.

2.2 Model Extensions

The basic model presented in the previous section demonstrates how mission hedging can make foundations more successful. However, foundations have not been implementing this strategy. I now explore why they choose to divest from or disregard in the investment process the nature of objectionable firms. In this section I formalize models consistent with the SRI and firewall strategies and then combine all the components presented into a comprehensive mission hedging model. In particular, I incorporate fundraising and pre-intervention bad activity level effects and expand the objective function to allow the level of portfolio evil to enter directly.

2.2.1 Socially Responsible Investing Model

An SRI foundation may fundamentally object to investing in reprehensible firms and therefore experience direct negative utility effects from doing so. Thus portfolio evil enters directly into the utility function as follows:

$$\max_{\alpha} E\left[U\left(b_2, e_p\right)\right]. \tag{2.16}$$

Furthermore, SRI proponents frequently argue that investing in objectionable firms helps those firms increase bad activity levels, or conversely that the divestment act directly lowers bad activity levels.[11] In other words, contrary to the assumption made in the basic model in section 2.1, $\frac{db_1}{d\alpha_i} \neq 0$. Finally, an SRI foundation may rely on fundraising and worry that resources could be negatively affected by objectionable investments. The SRI foundation solves the following expanded expected utility optimization problem:

$$\begin{aligned} \max_{\alpha} E\Bigg\{U\Bigg[y\Bigg(a_0\left(1+\sum_{i=1}^{n}\alpha_i r_i\right)+D\left(\sum_{i=1}^{n}\alpha_i e_i\right), \\ B\left(b_0, s_b, \sum_{i=1}^{n}\alpha_i e_i\right)\Bigg), \sum_{i=1}^{n}\alpha_i e_i\Bigg]\Bigg\}. \end{aligned} \tag{2.17}$$

This problem differs from the basic mission hedging model presented above in equation (2.8) in the following ways. First, $D(e_p)$, with $dD/de_p \leq 0$ where defined, represents donations received by the foundation as a function of portfolio evil level. Donors contribute less to a foundation with more objectionable investments if the first derivative is strictly negative. These donations are added to endowment assets in the inter-

[11]For example, student leaders of the Harvard fossil fuel divestiture campaign argued that they "do not expect divestment to have a financial impact on fossil fuel companies.... Divestment is a moral and political strategy.... Divestment calls on citizens to build a powerful climate movement and pressure elected representatives to enact meaningful legislation" (Maxmin and StudentNation (2013)).

vention technology function. Second, $B(b_0, s_b, e_p)$, where s_b is the stochastic component affecting the post-shock bad activity level, represents the viewpoint that investments in objectionable firms result in increases to bad activities (and divestments in decreases), with $B_e(e_p) \equiv dB/de_p > 0$. Finally, e_p enters directly into the utility function (as represented by the last term).

In addition, an SRI foundation behaves as if there is no relationship between any firm's returns and marginal utility of assets, or $Cov\left(U_A, \varepsilon_j - \varepsilon_k\right) = 0$. Solving this problem by the same steps as in the derivation of equation (2.15) yields the following trade-off equation:

$$\begin{aligned} \frac{dEU^{SRI}}{d\alpha_j} =& E\left[U_A\right] E\left[r_j - r_k\right] + \left(\beta_j - \beta_k\right) Cov\left(U_A, \left(r_m - r_f\right)\right) \\ & + \left(e_j - e_k\right)\left\{E\left[U_A D'(e_p)\right] + E\left[U_B B_e(e_p)\right] + E\left[U_e\right]\right\}, \end{aligned} \tag{2.18}$$

where $U_B \equiv {}^{dU}/_{dy} \cdot {}^{dy}/_{db_1} < 0$ and $U_e \equiv {}^{dU}/_{de} < 0$ (where defined). An SRI foundation optimizes with respect to expected returns, systematic risk premia, donation changes, effects on pre-intervention bad activity levels, and the direct disutility of a high portfolio level of evil. However, it does not generally consider the idiosyncratic risk trade-off.

Given all of these trade-offs, one can see how an SRI foundation might optimize by divesting completely of objectionable firms. In particular, if even very low portfolio evil levels trigger significant drops in donations or major dissatisfaction for foundation managers, then taking on additional evil in the portfolio might not be justified by other increases to expected utility. Evidence for these trade-offs is discussed in section 2.3.

2.2.2 Firewall Model

Firewall proponents typically argue that the direct effect of investing on bad activity levels is negligible. Their choices reveal a lack of concern about material impacts of their investment decisions on fundraising and insignificant disutility of evil investments. They tend to support the basic model from section 2.1 and the assumptions that $db_1/de_p = 0$ and $dD/de_p \approx 0$. However, in arguing that investments should be made in a values-vacuum, firewall advocates like SRI practitioners act as if $Cov\left(U_A, \varepsilon_j - \varepsilon_k\right) = 0$. In other words, firewall advocates optimize as if the trade-off faced by the foundation equals the following:

$$\frac{dEU^F}{d\alpha_j} = \left(\beta_j - \beta_k\right) Cov\left(U_A, \left(r_m - r_f\right)\right) + E\left[U_A\right] E\left[r_j - r_k\right]. \tag{2.19}$$

By ignoring the idiosyncratic risk trade-off, firewall foundations may make suboptimal portfolio allocations and miss opportunities to increase expected utility. In particular, they will underinvest in objectionable firms whose idiosyncratic risk returns covary strongly with marginal utility of assets.

2.2.3 Comprehensive Mission Hedging Model

Adding the components included in the SRI model in equation (2.17) to the basic mission hedging model while allowing for a non-zero idiosyncratic risk covariance term yields a comprehensive equation with trade-offs as follows:

$$\begin{aligned}
\frac{dEU}{d\alpha_j} = \; & E\left[U_A\right] E\left[r_j - r_k\right] && \text{[Expected Return]} \\
& + \left(\beta_j - \beta_k\right) Cov\left(U_A, \left(r_m - r_f\right)\right) && \text{[Market Risk]} \\
& + Cov\left(U_A, \varepsilon_j - \varepsilon_k\right) && \text{[Idiosyncratic Risk]} \\
& + (e_j - e_k) E\left[U_A D'(e_p)\right] && \text{[Fundraising]} \\
& + (e_j - e_k) E\left[U_B B_e(e_p)\right] && \text{[Direct Investment Effects]} \\
& + (e_j - e_k) E\left[U_e\right] && \text{[Direct Disutility]}
\end{aligned} \tag{2.20}$$

This is a generalized model, where equation (2.15) is a special case that assumes that $D'(e_p) = B_e(e_p) = U_e = 0$.

2.3 Evidence for Trade-Offs

I now explore the magnitude of the equation (2.20) trade-offs that endowment managers must weigh against the idiosyncratic risk trade-off in deciding whether to engage in mission hedging. My findings are as follows. Theoretical models show that investments can directly affect pre-intervention bad activity levels. But empirical evidence indicates these effects only come into play with very substantial investment levels, typically far beyond what mission hedging prescribes. The fundraising trade-off may be significant for some foundations but is immaterial to the majority that do not accept any donations. The expected return trade-off will often augment the idiosyncratic risk trade-off core to mission hedging, as many objectionable firms are found to have higher returns. However, this may be offset by the market risk trade-off. Finally, to address the direct disutility trade-off, I explore basic ethical arguments and find support for the ethical application of mission hedging. Evidence thus indicates that there will be foundations for whom the benefits of mission hedging are not completely offset by other trade-offs, though this should be evaluated on a case-by-case basis.

2.3.1 Investment Effect on Pre-Intervention Bad Activity Level

The basic mission hedging model presented in section 2.1 assumes investments do not directly affect pre-intervention bad activity levels. Because divestment proponents often disagree with this assumption, section 2.2 introduced the trade-off term dB/de_p to allow for investments to affect bad activity levels pre-intervention. Large positive dB/de_p may make divestment optimal.

There are two key ways in which investments could directly increase pre-intervention bad activity levels. First, an investment could be seen as an endorsement of a firm, thereby increasing the firm's goodwill. If this is a real concern, a foundation can explain that its investment is strategic and not a show of support for the firm.

Second, if the stock demand curve slopes downward, then mission hedging can raise the stock price by increasing the investor base. This could in turn cause the firm to increase bad activity through management incentives or lower cost of capital. In the context of imperfect information, Merton (1987) shows that an exogenous increase in a firm's investor base results in an increase in optimal investment level due to a lower cost of capital.[12] Rivoli (2003) cites diverging investor opinions and imperfect substitutes as additional theoretical market imperfections that could result in downward-sloping demand curves.

Heinkel, Kraus, and Zechner (2001) show through an equilibrium-based model that green investors shrinking a polluting firm's investor base may cause the firm to clean up. Here the mechanism is risk-sharing. It follows that if green investors choose instead to invest in polluting firms, this could increase pollution by those firms. However, in calibrating their model, Heinkel, Kraus, and Zechner (2001) find that green investors need to initially account for at least 20 percent of all investments in the polluting firm in order to cause it to invest in cleaner technology. If total market value is $30 trillion (the sum of the individual market capitalizations of all stocks in the Center for Research in Security Prices (CRSP) dataset), then this would correspond to $6 trillion of assets committed to a divestment movement. Therefore theoretical evidence suggests that investments could affect pre-intervention bad activity levels, but it remains unclear whether this effect is empirically meaningful.

To address this point, one must first determine whether the investment (or divestment) is likely to have a significant effect on share prices. Two main empirical approaches

[12]In the Merton (1987) model, each investor has information on only a subset of firms and is only willing to invest in firms on which it has information. Thus each firm has a different subset of investors for its base.

to answering this question are to examine specific socially-motivated divestment events and to estimate general elasticity of demand for stocks.

The mass boycott of firms doing business in South Africa during Apartheid is the most studied divestment event. However, even with this major event, there is no clear consensus on the divestment campaign's effectiveness. One key challenges in this line of research is isolating the effect of the divestment campaigns from related consumer boycotts and other pressures and news. Furthermore, investors may be more willing to divest from socially objectionable firms if those firms face fundamental risks that make them otherwise unattractive investments, which could lead to potential omitted variable bias in analyses.[13]

With the above caveats in mind, highlights of results on the South African divestment movement are as follows. Kumar, Lamb, and Wokutch (2002) survey research related to the divestment as a means to help end Apartheid in South Africa and report mixed results. In their own analysis, they find that firms remaining in South Africa during the boycott experienced increased institutional ownership and positive abnormal returns when Nelson Mandela called for an end to the boycott. However, Teoh, Welch, and Wazzan (1999) show that only the first in a series of voluntary pension fund announcements of divestment from South Africa (during Apartheid) had a significantly negative effect on relevant share prices. Thus even within one campaign, the benefits of specific divestment activities may vary significantly. In an event study of firms announcing exits from South Africa, Posnikoff (1997) finds a positive announcement effect. Meznar, Nigh,

[13] For example, a Lexis/Nexis news search on the recent fossil fuel divestment campaign reveals that the ramp-up in divestment announcements has coincided with unfavorable market conditions for the coal industry as dropping natural gas prices and tighter regulations have made coal prospects very uncertain (Macdonald-Smith (2014)). It is unclear how many organizations divesting from coal might not do so in the absence of the unfavorable outlook for the coal industry. A counterexample is the divestment campaign targeting Monsanto (see Food Democracy Now! at http://action.fooddemocracynow.org/sign/take_the_Monsanto_stock_plunge/), which has failed to gain traction while Monsanto performed well on the stock market.

and Kwok (1994, 1998) find that firms withdrawing from South Africa experienced negative abnormal returns. However, McWilliams and Siegel (1997) have raised concerns with Meznar, Nigh and Kwok's methodology and in a replication study show that the abnormal returns in the South African Apartheid case are small and insignificant.

Research on other divestment events is limited. However, Ding, Parwada, and Shen (2014) examine the boycott of firms doing business in South Sudan and find that quarters with increased divestment campaign news stories coincided with decreases in stock prices and institutional stock ownership and were followed by quarters with higher returns. However, inferences are limited by the fact that stock divestment news may coincide with news about consumer boycotts and other items relating to firm fundamentals.

On the question of elasticity of demand, Loderer, Cooney, and Van Drunen (1991) estimate price elasticity of demand for stock by examining primary stock offerings of already publicly traded firms. They find announcements of offerings increasing the number of shares yield negative price effects. However, they are unable to clearly attribute these changes to mechanisms like liquidity and heterogeneous beliefs that underpin theoretical explanations (like Merton (1987)) for downward sloping demand.[14]

Another approach examines the impact of inclusion in an index, which increases firm investor bases. Petajisto (2009) finds that S&P500 inclusion can result in price impacts of up to 3 percent. Similarly, Capelle-Blancard and Monjon (2012) find evidence for significant positive abnormal returns coinciding with SRI index inclusion. However, these abnormal returns may result from information conveyed by SRI index inclusion rather than the resulting investor base increase. In summary, there is some empirical support for downward sloping demand. However, it is not clear how steep that slope is or

[14]Note that one issue with the Loderer, Cooney, and Van Drunen (1991) results may be their use of relative price changes rather than abnormal returns.

how much of a price effect a particular investment might have.

In order for an investment (or divestment) to directly cause an increase (or decrease) in bad activity levels, any resulting stock price increase must be followed by a corresponding change in firm behavior. While it is well understood that stock returns predict firm investment levels, it is not clear to what extent this relationship is causal. Morck, Shleifer, Vishny, Shapiro, and Poterba (1990) show that incremental explanatory power of stock returns is small when fundamentals are accounted for, finding only limited support for the idea that stock price changes drive firm investments. Similarly, Blanchard, Rhee, and Summers (1993) find that share price changes have a limited impact on investment unless they are matched by corresponding changes to fundamentals. However, there may be some heterogeneity of firm investment sensitivity to share prices. For example, Chen, Goldstein, and Jiang (2007) and Baker, Stein, and Wurgler (2003) find that investment-to-price sensitivity is strongly positively correlated with the amount of private information in price and the level of firm equity dependence, respectively. These types of results may provide some guidance as to whether particular firms are likely to respond meaningfully to given stock price changes.

The magnitude of dB/de_p will increase with endowment size. Larger foundations are both more likely to negatively affect bad activities with their investments in objectionable firms and to be able to have a positive effect through screening. However, with a median asset size of about \$20 million (and a mean of about \$70 million), most of the 59 foundation signatories to the Divest-Invest Philanthropy fossil fuel divestment campaign (for whom asset data were available)[15] are unlikely to have much individual impact on multi-billion dollar firms.

In summary, there is limited evidence that investments made on the secondary

[15] **Divest-Invest Philanthropy.** “Signatories.” http://divestinvest.org/philanthropy/signatories/#/signatories (accessed June 10, 2015). **ProPublica.** “Nonprofit Explorer.” https://projects.propublica.org/nonprofits/ (accessed June 10, 2015).

market may increase pre-intervention bad activity levels. In order to change firm behaviors, divestment movements must generally occur on a massive scale beyond that prescribed by mission hedging. However, foundations considering doubling down on objectionable investments may mitigate concerns by ensuring that their investments (in aggregate with other foundations in the same field) are small relative to the market capitalizations and trading volumes of the narrowly targeted firms. Or endowments can, where possible, opt for alternative investments that through their correlations provide some hedge on the mission but are not in a position to influence bad activity levels. For example, an anti-tobacco foundation could invest in a tobacco-related medical device company. Or a foundation could potentially invest in a factor-mimicking portfolio of stocks highly correlated with the bad firm's idiosyncratic return component. Choices like these will likely reduce the expected utility benefit from mission hedging.

2.3.2 Foundation Fundraising

I now consider whether portfolio evil level e_p might reduce assets by decreasing donations due to public outrage. I find that the majority of private foundations do not accept donations and therefore are not affected by the fundraising trade-off. Table 1 shows results of an examination of IRS 990-PF foundation tax returns. 64 percent of all foundations (holding more than half of foundation assets) accepted zero donations in 2011. Furthermore, only 30.7 percent of foundations (holding about one-third of assets) received contributions, gifts, grants, etc., in excess of 5 percent of their total expenses and disbursements that year. I designate these as "Significant Donations" foundations, providing a proxy for foundations potentially at operational risk through mission hedging. However, this proxy likely overestimates how many foundations risk losing important donations. Many of the these foundations likely rely on donors (including founders, other foundations, governments, and other institutions) whose funding decisions will not be affected by endowment investments.[16]

The fact that most foundations are not currently divested from objectionable stocks[17] represents additional evidence that fundraising is not a broad concern for foundations when it comes to their investments. Firewall foundations have maintained their investment approach even after encountering fierce public criticism. For example, in 2007 the Los Angeles Times printed a dramatic 12-article investigative series criticizing the Bill and Melinda Gates Foundation for investing in firms working counter to its mission (Piller, Sanders, and Dixon (2007)). After the series ran, the Gates Foundation did not divest from the firms in question (Piller (2007)). Like most private foundations, the Gates Foundation does not rely on fundraising for continued operations. Furthermore, it is hard to imagine grantees rejecting funding from the foundation over this issue.

A foundation concerned about public backlash against investments can explain that its mission hedging strategy will improve the expected mission outcomes that it and its supports care about. If this explanation fails, the foundation may invest in alternative assets whose returns correlate with the bad activity level, as discussed in section 2.3.1, or use derivatives like stock options to benefit from the exposure to a firm's idiosyncratic risk without actually owning shares. These investments could be described as a form of insurance and thus be more palatable for donors. In summary, while it may pose public relations challenges, the majority of foundations will not put necessary donations at risk through mission hedging.

---

[16] This analysis does not include universities and colleges, which may legitimately be concerned about the impact of investments on attracting alumni donations. But divestment from particular firms or sectors may work to increase or decrease aggregate donations.

[17] In a 2012 survey, the US SIF Forum for Sustainable Investing found only 95 U.S. foundations that applied environmental, social, or corporate governance criteria in their investments (US SIF The Forum for Sustainable and Responsible Investment (2014)).

2.3.3 Expected Returns and Market Risk

I now explore whether the expected return and market risk trade-offs will offset or augment the idiosyncratic risk covariance trade-off for objectionable firms. In comparing a firm's expected returns to an alternative investment, the previously discussed possibility of a downward-sloping demand curve for stocks may augment the benefits of mission hedging. To the extent that the questionable firm is objectionable to some SRI investors, and therefore, has a smaller investor base, lower share price, and higher return, a mission hedging strategy will increase the foundation's ability to directly lower bad activity levels through both the hedging aspect and higher expected returns. However, this may be offset by higher systemic risk.

Hong and Kacperczyk (2009) find evidence for this in their study of "sin stocks," where they show that tobacco, alcohol, and gambling stocks are held less by norm-constrained institutions (like pension funds) and that they have higher expected returns than comparable stocks. Other studies like Fabozzi, Ma, and Oliphant (2008) also find significant excess annual returns for sin stocks. This could be due to downward-sloping demand curves or unusually high risk profiles. Renneboog, Ter Horst, and Zhang (2008) provide a good overview of some of the key empirical performance studies and find that "the existing studies hint but do not unequivocally demonstrate that SRI investors are willing to accept suboptimal financial performance to pursue social or ethical objectives." Capelle-Blancard and Monjon (2012), on the other hand, examine the trends in SRI literature and assert that there is consensus that SRI funds perform similarly to "conventional" peers and benchmark indexes. However, they do not provide clear empirical evidence for their conclusions.

In summary, available evidence indicates that the expected return trade-off component is likely to be negligible or to augment mission hedging benefits. However, foundations need to consider this on a case-by-case basis. If expected returns are low enough on an objectionable firm, this could outweigh other beneficial trade-offs enough to make divestment optimal. For example, if the recent challenges faced by the coal industry (as discussed briefly in section 2.3.1) are expected to continue, then expectations of inferior stock returns might make divestment optimal regardless of idiosyncratic risk trade-off benefits. On the flip side, to the extent that a foundation expects to earn superior returns on an objectionable firm, it should consider the additional market risk exposure the stock carries and whether this means that the demand curve is downward-sloping enough for the foundation's investment to meaningfully increase the firm's bad activities on the margin.

2.3.4 Direct Disutility of Evil Investments

To help evaluate the direct disutility trade-off component, I consider the question of how mission hedging might be viewed from an ethical perspective. Irvine (1987) provides an excellent overview of the basic ethical principles relevant to SRI. He argues that the "Evil-Company Principle" (that it is wrong to invest in an evil company) and the "Tainted-Profits Principle" (that it is wrong to benefit from the wrongdoing of others) are so inherently flawed that they are not reasonable grounds for screening. In recognition of the fact that some foundations (or their managers) may nonetheless be motivated by these principles, I allow for them through the inclusion of the last trade-off element in equation (2.20).[18]

Irvine (1987) then presents the "Enablement Principle" (that it is wrong to enable others to do wrong) as a valid reason to conclude that SRI is morally superior. Irvine further considers revisions like "act-utilitarianism" under which the Enablement Principle says: "It is wrong for me to do something that enables others to do wrong, unless my

[18] Note that even if there are no ethical concerns with an investment, there may be psychological or other sources of direct disutility associated with it.

failure to do the thing in question will have even worse consequences." He argues that the "Small-Purchase Objection," that an individual investor's small investment is acceptable because it won't have a significant impact, fails under the "Universalizability Principle," which states that an act is problematic if it causes problems when repeated by everyone.

To the extent that one agrees with Irvine (1987), the morality of mission hedging relies on the act-utilitarianism revision of the Enablement Principle. Although the foundation invests in what it sees as morally problematic firms, by doing so the foundation is uniquely positioned to (on average) do more good than harm by this action. However, there is still a concern about universalizability. What if all foundations engage in mission hedging?

Mission hedging is not a simple endorsement for all investors to concentrate their investments in firms they see as evil. Rather, mission hedging applies in a focused manner. A given foundation's mission may not lend itself to mission hedging for lack of firms whose returns correlate well enough with the foundation's targeted activities. Thus, for example, while the theory I have presented here indicates an anti-lung cancer foundation might benefit at least in part from skewing its investments toward tobacco firms, the mission hedging strategy would not in any way apply to its potential investments in fossil fuel firms. Even in cases where there are firms whose activities clearly conflict with foundation missions and whose returns correlate with bad activity levels, the aggregate assets held by all foundations targeting that particular set of firms under mission hedging are unlikely to be large enough to cause problems. To the extent this is a concern, foundations with common goals can work collaboratively to determine if their aggregated investments could potentially cause problems. Not being subject to anti-trust regulations, foundations may and often do cooperate.

2.4 Market Pricing of Foundation Mission Hedging

In the model presented above, the market does not react to the endowment's asset allocation and its (potential) effects on firms. A number of factors determine the validity of this assumption that a foundation's greater expected reduction of bad activity through mission hedging won't affect the share prices of a firm whose profitability is somehow intertwined with that bad activity. For example, a foundation seeking to fund carbon sequestration research might focus its mission hedging investments in the energy sector. Because greenhouse gases are a byproduct of the firms' activities and successful carbon sequestration does not reduce revenues or increase costs for the firms, the market should not shift the stock price in response to the foundation's decision to engage in mission hedging.[19]

On the other hand, a foundation with a proportional intervention seeking to fund a smoking cessation campaign that affects a fraction of smokers will on average hurt the bottom line of tobacco companies more if it implements mission hedging than otherwise. In this case the tobacco firm's share price should decrease ex ante (before the shock) under mission hedging.[20] In addition, a pre-intervention positive shock to (or increase in) smoking should result in a smaller increase in share price because the market knows the foundation, having skewed its investment toward tobacco, and thus having also had abnormal positive returns to its endowment, now has more funds with which to convince smokers to quit. Similarly, the negative price shock accompanying a negative smoking shock should be smaller because the mission hedging foundation now has fewer funds

[19]In fact, one could argue that the greenhouse gas emitting firms could benefit from successful carbon sequestration projects as consumers concerned about their carbon footprints might increase their consumption upon knowing that their emissions can be sequestered.

[20]The possibility has been raised that the foundation could benefit from shorting a firm that it will affect negatively through its intervention, causing the price to fall. This relies on an assumption that the market does not know about or believe in the effectiveness of the foundation's intervention. Here I am operating instead under the assumption that the market already knows about the foundation's activities and believes the foundation is simply investing in the alternative investment.

available for its campaign than it would have otherwise. In this sort of situation, the foundation might be wise to announce its strategy before making its purchases so that it can benefit from the lower ex ante price.[21]

Finally, if a foundation whose intervention technology is not proportional and will affect the bottom line of the firm[22] engages in mission hedging, the firm's share price should not be affected ex ante because the expected post-intervention bad activity level is not affected. However, the market reaction will dampen return shocks under mission hedging versus divestment because the post-intervention bad activity levels will be less positive and negative. This dampening of return shocks will result in a smaller idiosyncratic risk trade-off term and less benefit to mission hedging.

2.5 Conclusion

Both sides of the debate on whether foundations should disregard the objectionable nature of firms or divest from them have missed an important issue. Investing heavily in objectionable firms may increase foundations' expected utilities by aligning availability of additional funds with need for those funds.

In this paper I have formalized the endowment investment problems whose solutions are consistent with the divestment, firewall, and mission hedging approaches, articulated the trade-offs faced by foundations deciding how much to invest in objectionable firms, identified a key idiosyncratic risk trade-off that has been absent from the debate, and examined the theoretical and empirical evidence for these trade-offs. I have shown that foundations can increase expected utility by skewing investment toward firms

[21]In this scenario and the one below, it is also possible that the foundation might encounter a moral hazard problem. Staff might be induced to decrease efforts to fight activities of an objectionable firm when they know that the foundation endowment, which funds their salaries, has significant exposure to that firm's returns. This would dampen any mission hedging share price effects.

[22]This is probably the least likely scenario. One example might be a foundation seeking to help individuals reduce their fast food consumption by giving them fresh unprocessed food to consume instead. The objectionable firms in this case might be fast food companies.

whose returns correlate with the activities the foundations seek to reduce or eliminate.

Although I have not included a formal analysis of a foundation with more than one mission, the intuition for this extension is as follows. The first order condition for the *ex-post* allocation of funds between missions requires that the marginal utility of the money spent on program areas be equal. Making substitutions following on that fact yields the same basic analysis as the single-mission case, except that the idiosyncratic risk trade-off magnitude is probably smaller because of the split focus.

The model presented here yields a number of empirical implications that could be tested given currently unavailable data on missions, SRI participation, program area budgets, and investment allocations of foundations. For example, the model suggests that as a foundation's missions change, so should its investment allocations. Furthermore, a foundation with multiple missions will realign its budget and investment allocations as shocks shift the relative marginal utility of funds between missions. Consider, for example, a foundation seeking to reduce smoking and obesity. If smoking declines but obesity increases, the model predicts the foundation will increase the fraction of its budget allocated to obesity. If the foundation decides to dedicate itself entirely to fighting obesity, then the model indicates it would be better served by shifting mission hedging investments away from tobacco and toward fast food and sugary drink producers.

Future work on measuring trade-off magnitudes could help answer how beneficial SRI is to foundations. A difference-in-difference strategy that examines the effect of changes in SRI activities on donations could estimate the magnitude of the donation trade-off. Aggregating foundation assets by mission area would help determine whether there are any program areas in which total foundation assets represent a large fraction of industry market capitalization, leading to possible concerns that mission hedging investments could affect pre-intervention bad activity levels. Analyses of how firm fundamentals and returns affect divestment campaign participation would shed light on the role of the expected return trade-off in the divestment decision (and a possible source of bias in divestment event studies like those on South Africa during Apartheid.) Given a set of well-defined missions and assumptions about specific utility functions one could estimate covariance between marginal utility of foundation assets and idiosyncratic firm risks, the trade-off central to mission hedging. Finally, one could measure the perceived balance of trade-offs by examining relationship between foundation missions targeting activities of objectionable firms and investment in those firms. Mission hedging activity (versus divestment) would be indicated by greater-than-market weights on those firms.

Divestment is making headlines with college students pressuring administrators to divest their endowments primarily of fossil fuels, but also of guns and Israeli stocks. Major universities have announced their decisions to divest (or not) from fossil fuels. Private foundations have also joined the movement, including the high profile Rockefeller Brothers Fund (Schwartz (2014)) as well as many smaller funds, both with and without environmentally-oriented missions.[23] Just as in voting, where one individual's decision not to vote is unlikely to determine the outcome in a major election, one foundation's divestment is unlikely to change a company's behavior. The power in moves like that of the Rockefeller Brothers Fund lies more in the ability to encourage others to follow suit than in directly changing fossil fuel firm behaviors through selling shares of stock. The potential success of propelling a broad divestment movement must be weighed against all trade-offs, including the mission hedging benefit of making more funds available when they are needed the most.

While major divestment movements have the potential to bring about change, my results show that firms who are seen as bad actors may provide good opportunities for hedging foundation-specific risks. Endowment decision-makers need to ask whether

[23] See **Divest-Invest Philanthropy.** "Signatories." http://divestinvest.org/philanthropy/signatories/#/signatories (accessed June 10, 2015.)

divesting, disregarding values in investing, or doubling down on objectionable stocks will yield the best social outcomes given not only their values but their unique missions, talents, and possible correlations between firm financial returns and foundation spending needs.

Chapter 2 is based on a paper that has been submitted for publication. The dissertation author is the principal author of chapter 2.

2.6 Tables and Figures

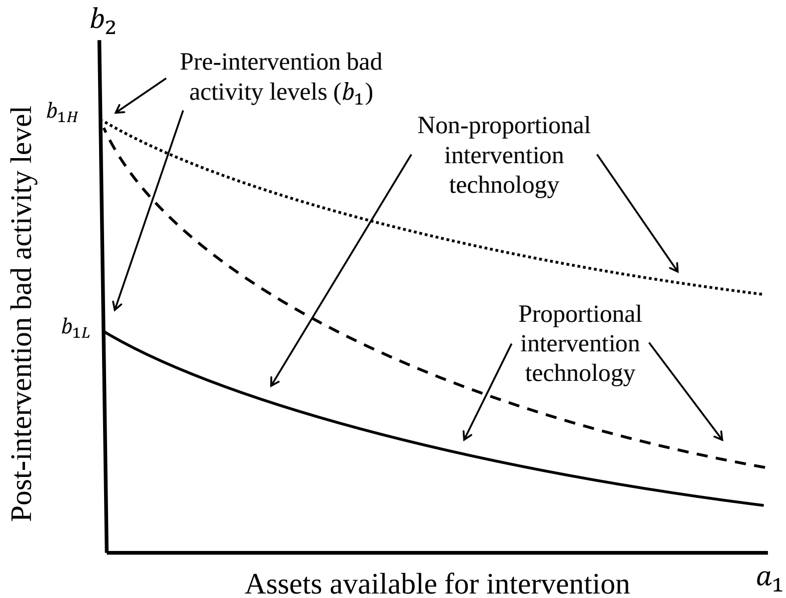
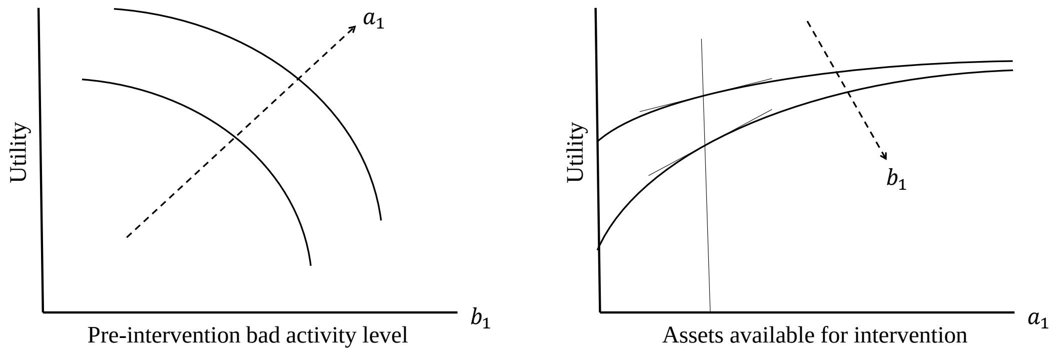


Figure 2.1. Proportional and non-proportional intervention technologies



(a) Utility of b_1 **(b)** Utility of a_1

Figure 2.2. Foundation utility function

Table 2.1. Foundation fundraising

	All Foundations		Zero Donations		Significant Donations	
Asset Range ($mil)	Number	Assets ($mil)	Percent of Total	Assets ($mil)	Percent of Total	Assets ($mil)
0.0-0.1	15,785	540	48.6	315	48.2	208
0.1-1.0	34,704	14,819	68.0	10,192	26.6	3,789
1.0-10.0	24,708	77,559	69.0	52,394	25.3	20,565
10.0-34,000	6,471	525,751	60.2	269,540	30.1	184,108
0-34,000	81,669	618,670	64.0	332,441	30.7	208,671

“Zero Donations” foundations are those who reported zero Contributions, Gifts, Grants, Etc. Received on their 2011 990-PF tax returns. “Significant Donations” foundations have received Contributions, Gifts, Grants, Etc., representing at least 5 percent of their total expenses and disbursements (including grants) in 2011. Foundations that started or ended the year with no assets or made no grant payments over the year were excluded, as were those that were terminated or in a 60-month termination. These summary numbers have been adjusted to reflect the stratification weights applicable to each observation.

These 2011 data are based on 990-PF microfile data from the IRS accessed at http://www.irs.gov/uac/SOI-Tax-Stats-Private-Foundations-Harmonized-Microdata-Files-ASCII on 2/20/2015.

Chapter 3

Private foundation spending responses to the tax on net investment income

Private foundations like the Bill and Melinda Gates Foundation, the Ford Foundation, and the Rockefeller Foundation, play a significant role in funding charitable and socially important activities in the United States. Foundation missions may involve short- and long-term goals, address issues like the environment, education, and the arts. To encourage and support donors' charitable activities, the US government generally exempts private foundations from income and capital gains taxes. However, to discourage wealth hoarding, private foundations are subject to special taxes designed to encourage spending. In this paper, I explore the relationship between intertemporal foundation spending and the notched tax on net investment income, with some discussion of the minimum spending requirements foundations face.

I build on existing literature by developing a model that characterizes the foundation's optimization problem subject to taxes. This model elucidates the complexity of the problems that foundations face in trying to optimize their intertemporal spending. I next characterize and compare several common intertemporal foundation spending policies. These strategies are relatively simple responses to the complex optimization problem. Using a hypothetical setting, I examine the implications of each spending policy on spending levels and endowment value.

I next use foundation tax filing data to explore the effect of the notched tax on net investment income. Following earlier works on notches, I use histograms to look for bunching around two tax thresholds. Consistent with previous results, I find that the threshold relating to minmum spending requirements appears to bind. Exploring the trade-off between higher benefits and costs of achieving the low tax rate in high return years, I build on earlier work by showing that the notched investment tax threshold exhibits more bunching in years with maximum returns than years with minimum returns. This suggests a more important role for tax savings than for endowment spending power preservation in high return years. Descriptive regressions show similar relationships.

For my final contribution I combine market return data with aggregate and foundation-level data on tax rates, bunching at the net investment income tax threshold, and initial foundation asset allocations to corporate stocks. I show that higher annual market returns and lower annual volatility decrease the fraction of foundations qualifying for the lower tax rate and increase bunching at the threshold. This provides new evidence that higher returns make it more difficult to achieve the low tax rate but also increase contemporaneous (but not necessarily total long-term) spending by foundations in order to achieve the low tax rate. Accounting for intial foundation level asset allocations to corporate stocks appears to amplify these basic effects.

This paper is laid out as follows. In section 3.1 I provide a brief summary of the historical background of the taxes analyzed herein. Section 3.2 reviews the relevant literature. Section 3.3 details the current tax rules, which I incorporate in an optimization model in section 3.4. In section 3.5 I categorize and compare potential intertemporal spending policies. Section 3.6 describes the data I use for the empirical analysis in section 3.7. Finally, I end with concluding thoughts in section 3.8.

3.1 Historical Background

Before 1969, it was possible for donors to establish private foundations, receive the immediate tax benefits, and then let the funds sit for years on end, accumulating income, without actually spending anything on charitable activities. In response to concerns that foundations were not sufficiently spending funds on charitable causes to justify the initial tax breaks, Congress enacted the Tax Reform Act of 1969. This act required private foundations[1] to spend at least 5 percent of their endowments per year. This is often called the minimum payout requirement.

To pay for implementation of the new payout regulations, the Tax Reform Act of 1969 also included a 4 percent tax on net investment income. In 1978 this tax rate was lowered to 2 percent. Even at this lower rate, the tax on net investment income generated much more revenue than was needed to cover the cost of regulating foundations. In 1984, In a purported attempt to encourage foundations to spend more than the required 5 percent minimum, this tax was restructured into a notched tax. Under the 1984 rule, still in effect today, a foundation can qualify for a lower tax rate of 1 percent of net investment income if its spending exceeds a threshold determined by past and current spending rates, asset levels, and net investment income. I call this the *investment tax threshold*.

To encourage spending beyond the minimum payout requirement, the investment tax threshold rewards foundations when they increase spending over previous years. However, increased spending one year by design also raises the investment tax threshold for future years. Foundations have argued that this tax therefore discourages one-time spending increases in response to humanitarian crises. Responding to the complexity

[1] The Internal Revenue Service defines two categories of private foundations: "operating" and "non-operating" foundations. Operating foundations are typically associated with hospitals, schools, and other entities for which they directly fund charitable activities. Non-operating foundations are typically primarily grantmakers that give funds to other organizations that do the actual charitable work. In this paper, I will say "foundations" to refer to non-operating foundations, which are the subset of foundations that the tax laws I discuss apply to.

and potential for perverse incentives, opponents to the notched tax have attempted to turn the net investment income tax back into a flat tax. In particular, the 2001 Clinton budget proposal aimed for a fixed 1.25% excise tax rate (Abelson (2000)). However, this change was not implemented. In 2009 Senators Schumer, Stabenow, and Levin proposed legislation that included a new fixed excise tax rate of 1.32% (Spector (2009),Wilhelm and Williams (2009)). The bill never made it out of committee[2] Further attempts have been made to change the tax, including in the February, 2016 President's Budget blueprint for Fiscal Year 2017 (Office of Management and Budget (2016)). However, the tax on net investment income remains a complex notched tax which I describe in detail in section 3.3.2.

3.2 Relevant Literature

There has been some work examining the foundation tax rules I explore in this paper. However, given the fact that foundations managed about $600 billion in assets and paid approximately $400 million in excise taxes in 2010, one would expect to see more literature on the effects of this tax. The relevant work dates back to Labovitz (1974), which provides a good overview of the Tax Reform Act of 1969 and very simple estimates of impact of rules. Wagner (1977) provides a good history of the rules and discusses the weak survivability conditions of private foundations relative to other organizations such as firms or other types of charities.

Sansing and Yetman (2006) examine the minimum payout requirement and the notched net investment income tax. They examine the ratio of actual spending levels relative to minimum payout requirements, consider how this ratio shifts with foundation characteristics, and categorize how constrained foundations are by the minimum payout

[2] See "S.676 - A bill to amend the Internal Revenue Code of 1986 to modify the tax rate for excise tax on investment income of private foundations" at https://www.congress.gov/bill/111th-congress/senate-bill/676.

requirement. Sansing and Yetman (2006) also examine annual net investment income tax rate determinants and characteristics of foundations that always select the lower or higher tax rate. They show that foundations that qualify for the lower tax rate in every period are large and professionally managed. Higher growth in assets corresponds to higher likelihood of paying the 2% rate, just barely missing the 1% rate, and always having the same tax rate (either 1% or 2%). This could be due to the fact that the growth rate they use is not net of donations or spending.[3]

Yoder and McAllister (2012) build on Sansing and Yetman (2006) and analyze the relationship between foundation-year investment tax rates and a variety of predictive variables which they also interact with contextual variables. Yoder, Addy, and McAllister (2011) examines the accounting methods and strategies implemented by foundations who are just barely above or below the investment tax threshold. They show that foundations that barely qualify for the lower tax rate are more likely to employ accounting strategies called "set-asides" that allow them to adjust their accounting spending levels after the year end.

Other relevant literature on private foundation spending includes Deep and Frumkin (2001), which uses survey-based data to examine trends in spending policies of private foundations, finding that they are largely targeting paying out the minimum amount required. Boris, Renz, Barve, Hager, and Hobor (2006) examine how foundation characteristics influence foundation expenses and compensation of staff and trustees. Without considering any behavioral responses, the Myers and Roberts (2008) report commissioned by the Council of Michigan Foundations, which sought to simplify the excise tax[4], estimates that a fixed excise tax rate of 1.32% would yield revenues equiv-

[3]Examining whether foundation tax rate patterns as a function of investment returns might yield interesting insights.

[4]http://ejewishphilanthropy.com/foundations-urge-congress-to-simplify-tax-law/

alent to the current notched tax.[5] Bignami (2013) shows that many foundations use the 5 percent payout as a minimum or floor and argues that this payout level makes it difficult to maintain purchasing power in perpetuity. Roth Tran (2015) examines how foundations can leverage investment strategies regarding objectionable assets to align funds for spending with states with higher needs.

3.3 Current foundation tax rules

Non-operating private foundations are subject to two taxes intended to incentivize spending: the tax on undistributed income (from failing to meet the minimum payout requirement) and the tax on net investment income. These taxes are calculated on the Internal Revenue Service (IRS) 990-PF tax form that every foundation must file annually. I explain the mechanics of each of these taxes in turn below.

3.3.1 Minimum payout requirement

The tax on undistributed income is designed to cause foundations to spend on average (at least) five percent of their assets annually. Using IRS 990-PF terminology, first define the *net value of noncharitable-use assets* (*NA*) as the sum of the average monthly fair market value of securities, the average monthly cash balance, plus the end of year fair market value of all other assets. The *distributable amount* (*DA*) equals 5 percent of *NA* less the tax on investment income.[6] The distributable amount is what must be spent on charitable purposes in order to satisfy the minimum payout requirement. The foundation spending that qualifies toward the minimum payout requirement is called its *qualifying distribution* (*QD*).[7]

---

[5] This is the same rate proposed in the 2009 legislation put forth by Senators Schumer, Stabenow, and Levin.

[6] Foundations subject to an income tax also subtract that from the 5 percent of *DA*.

[7] Qualifying distributions include "amounts paid (including administrative expenses) to accomplish charitable, etc., purposes." Typical major categories are expenses, grants, and net program-related investments. Note that tax on net investment income does not count toward qualifying distributions.

Foundations have a grace period of one year to make qualifying distributions to cover their distributable amounts. In other words, funds equivalent to the distributable amount generated in year t (DA_t) must be spent by the end of year $t+1$. Foundations can apply current year qualifying distributions QD_t to satisfy the current year minimum payout requirement DA_t. *Undistributed income* (UI) is defined as the distributable amount for which no qualifying distribution has yet been made. If the foundation fails to spend enough in qualifying distributions to cover this undistributed income within the one-year grace period, it will be subject to a 30% tax on the remaining undistributed income until it is paid. Thus I will also refer to the undistributed income as the *minimum payout threshold*. Conversely, when a foundation spends more than the current distributable amount plus undistributed income, it can carry this *excess distribution* (ED) forward to apply it to distributable amounts for the next five years.[8] This complicated setup allows foundations to optimize and smooth spending intertemporally.

3.3.2 Notched tax on net investment income

One can think of the tax on net investment income as having a default rate of 2 percent with the option for a lower 1 percent tax rate. In order to qualify for the lower rate, a foundation's spending must exceed what I call the investment income tax threshold. First define *adjusted qualifying distributions* (AQD_t) as QD minus the tax on net investment income in years that a foundation qualifies for the lower tax rate. Then define the *distribution ratio* (DR) as the ratio of the current adjusted qualifying distribution to the current net value of non-charitable use assets (i.e. $DR_t = AQD_t/NA_t$).

[8] For an illustrative example, consider a new foundation established at the very start of 2014. Suppose that over the course of 2014, its net value of noncharitable-use assets is \$10 million. Then its distributable amount for 2014 is \$500,000. If the foundation spent \$240,000 in qualifying distributions in 2014, then it will need to spend at least \$260,000 in 2015 in order to avoid paying a 30 percent penalty tax on undistributed income. Suppose in 2015 that the foundation's net value of noncharitable-use assets is \$11 million and that it spends \$1 million in qualifying distributions. Its 2015 distributable amount is \$550,000. When we combine this with the 2015 qualifying distributions and 2014 undistributed income, the foundation now has excess distributions of \$190,000 to carry over to 2016.

A foundation qualifies for the lower tax rate if the current ratio of (unadjusted) qualifying distributions to net value of non-charitable use assets is at least as great as the average of the previous five years' distribution ratios plus 1 percent of net investment income of the current year (R_t). More formally, the excise tax rate (τ_t) is defined as:

$$\tau_t = \begin{cases} 1\% & \text{if } QD_t/NA_t \geq .01 \cdot R_t + \frac{1}{5}\sum_{i=1}^{5} DR_{t-i} \\ 2\% & \text{otherwise} \end{cases} \tag{3.1}$$

Thus the investment tax threshold (ITT) equals

$$ITT_t = NA_t \cdot \left[.01R_t + \frac{1}{5}\sum_{i=1}^{5} DR_{t-i} \right] \tag{3.2}$$

A foundation whose qualifying distributions exceed this investment tax threshold pays the lower 1 percent tax rate on net investment income.

There are three important issues to note concerning this tax. First, while one compares the qualifying distributions of the current year with the *previous* year's net value of noncharitable-use assets to calculate the current minimum payout threshold, *current* net value of noncharitable-use assets and *current* net investment income determine the current investment tax threshold. A foundation with any risky assets cannot not know these values or the threshold until the end of the year in which the threshold must be met to qualify for the lower tax rate.

The second major issue is the all-or-nothing nature of the investment tax incentive. The cost of missing the investment tax threshold by even the smallest amount can be very dramatic. Missing the threshold by \$1 would cost \$10,000 in additional taxes for a foundation with net investment income of \$1 million. In contrast, missing the minimum payout threshold – which is known with certainty at the start of the fiscal year – by \$1 would only cost any foundation 30 cents in additional taxes.

The final important feature to note about the net investment income tax is that when a foundation qualifies for the lower investment tax rate, the foundation subtracts the investment income tax for the year from the qualifying distribution for calculations of future year investment tax threshold. This foundation's future investment income tax threshold will be lower than if it had missed the cutoff. This adds to the potential cost of missing the investment tax threshold by increasing the risk of missing the threshold in future years as well.

3.4 Model

I assume that foundations have missions which they use their endowment funds to work on. These missions generally extend over some number of years and often into perpetuity. Thus one could model the foundation's objective function as the following:

$$\max_{QD_t} \sum_{t=1}^{T} \beta^{t-1} U(QD_t) \tag{3.3}$$

where QD_t is the level of spending in year t and β is the discount factor. This objective function is subject to a budget constraint that accounts for taxes and investment returns. To incorporate these constraints in a formally stated problem that abstracts from the investment allocation decision, consider the following sequence of events:

1. Pay taxes for previous year: T_{t-1}
2. Earn net investment income: $R_t = r_t(a_{t-1} - T_{t-1})$
3. Choose and spend payout: QD_t

   Note the foundation already knows

   a) Minimum payout threshold: $UI_{t-1} - ED_{t-1}$

   b) Historical part of investment tax threshold: $\frac{1}{5}\sum_{i=1}^{5} \frac{AQD_{t-i}}{NA_t - i}$

4. At year end, calculate:

   (a) Year-end assets $a_t = a_{t-1} + R_t - QD_t - T_{t-1}$

   (b) Net Average Use Assets $NA_t = (a_t + a_{t-1})/2$

   (c) Distributable Amount $DA_t = .05NA_t$

   (d) Tax rate for current year $\tau_t = .02 - .01 \cdot \mathbb{1}\left\{QD_t \geq NA_t \cdot \left[\frac{1}{5}\sum_{i=1}^{5}\frac{AQD_{t-i}}{NA_t - i} + .01R_t\right]\right\}$

   (e) Taxes owed for current year $T_t = \tau_t R_t$

   (f) Adjusted Qualifying Distributions $AQD_t = QD_t + (\tau_t - .02)R_t$

   (g) Undistributed Income $UI_t = \max(0, QD_t - DA_t)$

   (h) Excess Distributions[9] $ED_t = \max(0, DA_t - QD_t)$

The foundation's basic optimization problem can thus be stated as follows:

$$
\begin{aligned}
\max_{QD_t} \quad & \sum_{t=1}^{T} \beta^{t-1} U(QD_t) \\
\text{s.t.} \quad & a_t = (1 + r_t)(a_{t-1} - T_{t-1}) - QD_t \\
& a_t \geq 0 \\
& a_T \geq A + T_T \\
& QD_t \geq UI_{t-1} - ED_{t-1} \\
& R_t = r_t(a_{t-1} - T_{t-1}) \\
& T_t = R_t \cdot \left[.02 - .01 \cdot \mathbb{1}\left\{QD_t \geq NA_t \cdot \left[\frac{1}{5}\sum_{i=1}^{5}\frac{AQD_{t-i}}{NA_t - i} + .01R_t\right]\right\}\right]
\end{aligned}
\tag{3.4}
$$

[9]Note that I have simplified the excess distributions here to abstract from the ability to apply excess distributions for a five year period.

where r_t, R_t, and NA_t are uncertain and exogenous.

The budget constraints make the model I have developed here complex. However, I have excluded random shocks that would affect returns and utility of payout over time. These elements would further complicate the problem and are potentially very important to foundation spending strategies. Under standard assumptions of increasing utility and decreasing marginal utility, a foundation would generally wish to smooth its spending over time and to maximize its total spending. Depending on its particular mission, a foundation might face changes in marginal utility of payout over time. For example, a foundation that funds flood recovery efforts would see stochastic needs with a potential for increasing needs over time driven by climate change.

When considering whether to meet the minimum payout and investment tax thresholds, every foundation must trade off the cost of failing to meet a threshold with the cost of reducing future spending power in order to meet that threshold. In the next section I discuss some common practical intertemporal spending strategies employed by foundations.

3.5 Intertemporal spending policies

Given the complexity of the foundation spending optimization problem, foundations often respond by following relatively simple spending policies. I now outline the primary examples.

a) **Spend enough to meet minimum payout threshold.** A foundation may opt to spend its undistributed income, which is the minimum amount required to avoid the 30 percent penalty tax. This foundation spends money in the year that it is legally required to avoid the 30 percent penalty, not sooner or later. This policy preserves endowment spending power over time while disregarding the costs and benefits of qualifying for a lower tax on investment income. In year t, the foundation following this strategy would set its spending level as follows:

$$QD_t = DA_{t-1} - ED_t \tag{3.5}$$

b) **Minimize net investment income taxes.** This foundation targets the low investment tax threshold every year that it has positive net investment income. It exceeds the minimum payout threshold to avoid paying the undistributed income penalty. This foundation maximizes the share of its expenditures that it allocates to its own charitable purposes rather than to taxes. Under this policy, the foundation seeks to spend:

$$QD_t = \max[TT_t, DA_t - ED_t] + \varepsilon_t \tag{3.6}$$

Assuming the foundation also wants to avoid the 30 percent penalty, it will seek to meet the tax threshold TT_t defined as follows:

$$TT_t = \begin{cases} .01 \cdot R_t + \frac{1}{5}\sum_{i=1}^{5} \frac{AQD_{t-i}}{NA_t - i} NA_t & \text{if } R_t > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

The challenge with this policy is that given risky assets the foundation does not know until the end of period t what, precisely, NA_t and R_t will be. I include the error term ε_t to reflect this uncertainty. The foundation must manage this uncertainty in trying to aim for the threshold that minimizes spending subject to qualifying for the lower tax rate.

c) **Constant percent of assets.** A foundation may choose to spend fraction s of current assets in any given year. Setting $s = .05$ would roughly mimic the minimum qualifying distribution requirement over time. Assuming foundations still avoid the 30 percent penalty, this policy is represented mathematically as:

$$QD_t = \max\left[s \cdot a_{t-k}, DA_{t-1} - ED_t\right] \tag{3.8}$$

Here k allows foundations to base their spending on asset levels for current or previous years. Foundations may also choose to base their spending on assets at the end of a fiscal year or on the net of noncharitable-use assets as calculated for IRS form 990PF.

d) **Moving average.** Many universities, which like foundations often have significant endowments to manage, spend a percent of a moving average of assets every year. While implementation of this is simpler for universities, which are not constrained by minimum payout requirements, foundations may also try to employ moving average methods to smooth their spending. Mathematically, this would would be equivalent to:

$$QD_t = \max\left[s \cdot \frac{1}{Y}\sum_{y=1}^{Y} a_{t-y}, DA_{t-1} - ED_t\right] \tag{3.9}$$

e) **Ad hoc spending.** Foundations may choose to fund worthwhile projects as needs arise. Spending is driven by exogenous request shocks, but subject to the constraint of avoiding the 30 percent penalty tax. For example, a foundation's spending levels might be modeled as follows:

$$QD_t - DA_{t-1} \sim N(\mu, \sigma^2) \quad \text{s.t. } QD_t \geq DA_{t-1} - ED_t \tag{3.10}$$

3.5.1 Spending policy simulations

Table 3.1 shows results of simulated spending paths for foundations following each of the above spending policies. Each hypothetical foundation starts with an initial endowment of \$1 million, spends 5 percent of that in the first year. Each foundation then follows one of the specific policies outlined above for every subsequent year. The first set of simulated outcomes are for a constant known rate of return of 8 percent.[10] The second set is based on actual monthly S&P 500 index returns for the years 1992-2011, reflecting the time period covered by the data examined in this paper.

Column 1 of Table 3.1 shows that in the absence of uncertainty, the foundation that minimizes investment income taxes also maximizes its spending, or qualifying distributions.[11] However, given the stream of S&P 500 returns in this time period, the moving average asset policy maximizes qualifying distributions under the uncertain return scenarios. This is in part because there were three consecutive years of negative returns (2000 - 2002), during which the foundation still had to spend significant sums to avoid the 30 percent penalty on undistributed income even though there were no benefits from a lower tax on investment income. Column 2 shows that the strategy of spending the minimum required to avoid the qualifying distribution penalty yielded the highest final asset value. This appears to be a winning strategy for foundations interested in preserving purchasing power. The last column suggests that spending a constant percent of the moving average of assets yields the most stable growth in qualifying distributions given market uncertainty.

3.6 Data

Foundation financial data were produced by the Internal Revenue Service, which provides detailed 990-PF tax return data for a stratified sample representing about 10 percent of all foundations in a given year (Internal Revenue Service (2011)). Large asset classes are sampled at a rate of 100 percent. Every foundation must complete the 990-PF form each year and must make it publicly available upon request. The data include most

[10]The mean annual return level I observe in the foundation data is 8.3%.
[11]This would not be true if $s > .05$.

of the key financial fields, though some (like negative net investment income amounts) must be imputed based on other fields. The dataset contains no information on foundation grant recipients or grantmaking areas.

I exclude observations with assets under $10,000 at the start or end of the year or with zero net noncharitable-use assets for the year. I also exclude observations marked as being in termination proceedings, designated as "initial" or "final", or for which the initial or final observation for the foundation after the start or before the end of the 1992-2010 dataset range. Foundations cannot qualify for the low tax rate in the first year and are not likely to be constrained by the tax in the final year. I also excluded foundations with fewer than five years worth of observations as these foundations are similarly likely to still be in initial and/or final stages of their existence. In Table 3.2 I present summary statistics for the data I use.

3.7 Empirical results

3.7.1 Does the investment income tax incentive affect spending levels?

I use two approaches to answer the question of whether the investment income tax incentive affects spending levels. First, consistent with Sansing and Yetman (2006) and Yoder and McAllister (2012), I examine whether foundation spending bunches around the investment tax threshold. The shaded bars in Figure 3.1 show distinct and dramatic bunching immediately above the minimum payout threshold. In contrast, the empty outlined bars in the same figure show much more noisy and less pronounced bunching around the investment tax threshold.

Figure 3.1 shows that the minimum payout threshold is clearly binding. If the investment tax threshold tends to be close to the minimum payout threshold, the noisy bunching around the investment tax threshold could be driven entirely by the minimum payout threshold. Figure 3.2 shows a scatter plot of the two thresholds along with a dashed line where the two thresholds would be equivalent. This figure indicates that the two thresholds tend to track close together. In Figure 3.3 I show that even if I only consider the subset of observations in which the investment income tax threshold is at least twice as large as the minimum payout threshold, that there tends to be bunching around the threshold. This result suggests that the investment tax threshold may influence spending, but that there is significant error in achieving the threshold.

Higher net investment income increases both the benefit of the lower tax rate and the threshold for (and thus cost of) achieving that tax rate. I explore whether there is more bunching when returns are unusually high or low. Figure 3.4 shows the distribution of years when foundations experience their highest and lowest returns (percentage-wise). As in Figure 3.3, I limit the data to the subset of observations for which the investment tax threshold is at least twice as large as the minimum payout threshold. Figure 3.5 examines qualifying distributions relative to the investment tax threshold for the subsets of observations in which foundations experienced their highest and lowest investment return years. Consistent with the lower threshold, this plot suggests that foundations are more likely to achieve the lower tax rate in low return years. However, there is more bunching at the investment tax threshold in high return years, suggesting that the higher benefit in those years is creating a larger incentive to target the threshold.[12]

In Table 3.3 I show results from regressions of log of qualifying distributions on the logs of minimum payout and investment tax thresholds as well as net investment income. These results show that foundation spending increases with the investment tax threshold, conditional on the minimum payout threshold and net investment income. However, as past spending levels affect current thresholds, I am not able to use these regressions to identify causal effects.

[12] An examination of only the capital gains portion of net investment income yields similar results.

In Tables 3.4 and 3.5 I regress an indicator for whether the foundation qualified for the low tax rate on a variety of investment related regressors. I show that investment returns and capital gains are highly predictive of the foundation net investment income tax rate. Table 3.4 shows regression results without year fixed effects, while Table 3.5 regressions include year fixed effects. Major charitable giving shocks like catastrophes that spur spending demand inclusion of year fixed effects. However, the impact of investment returns is of primary interest in this regression and is highly correlated by year. Including year fixed effects may wash out too much of the exogenous variation in investment income. Note that in the model in column 2, inclusion of year fixed effects increases significance of regressors. One possible interpretation of this result is that when one controls for the market-wide variation in returns that year fixed effects capture, the remaining variation in returns and capital gains (which can be strategically timed) have an important strategic relationship with the net investment income tax incentive. For example, a foundation that plans to realize significant capital gains one year will also strategically plan to spend enough to qualify for the lower tax rate.

In terms of the individual coefficients, the negative coefficient on the indicator for whether the investment income tax threshold is greater than the payout threshold is consistent with previous results showing that the payout threshold is much more of a constraint. A positive coefficient on investment return indicates that when returns are higher, increased benefits from the lower tax rate outweigh higher opportunity costs of achieving the low rate. The positive coefficient on the highest return year and negative coefficient on the lowest return year are consistent with the bunching patterns shown in Figure 3.5. The lag return and capital gains percent terms do not appear to have significant coefficients, suggesting that foundations are responding quickly to contemporaneous incentives, without a lag response.

3.7.2 How do market returns affect investment income tax rates?

Foundations may strategically shift their asset allocations, the timing of capital gain realizations, or spending in early periods. These sorts of actions introduce endogeneity to analyses of whether the investment income tax threshold influences spending. However, stock market returns are exogenous and unpredictable. I now examine how market-level returns affect foundation investment income tax levels and bunching around the investment tax threshold.

Figure 3.6 shows four scatter plots with annual and monthly S&P 500 index returns and volatility on the x-axes and the fraction of foundations qualifying for lower excise tax rates on the y-axes. Figure 3.7 shows the wide range in asset allocations that foundations have to corporate stocks, yielding differential exposure to market risks. Figure 3.8 shows that while the majority of foundations in the data have fiscal years ending in December, there is some variation across fiscal year ends as well. Figure 3.6 depicts December year-end observations with larger and darker points.

Figure 3.6 indicates that higher annual returns correspond to a decrease in foundations qualifying for the lower rate (upper left plot), while higher volatility over the entire fiscal year or just the last month of it corresponds to more foundations qualifying for the lower rate (upper and lower right plots.) These descriptive results contrast with Table 3.4 and 3.5 results, which suggest that higher investment returns correspond to more foundations foundations qualifying for the lower tax rate. Given the uncertainty in the threshold during a year, one might also expect more foundations to qualify for the lower investment tax rate in years when the market is more stable and foundations are better able to plan accurately for the investment tax threshold.

Table 3.6 I regress the fraction of foundations qualifying for the low investment tax rate in a given month and year on the S&P500 returns and volatility. For these regressions, I aggregate observations across foundations at the month-year level and exclude foundations with non-varying investment tax rates. The results in the first four columns mirror the relationships depicted in the Figure 3.6 scatter plots. In column 5 I regress the fraction of foundations qualifying for the low investment tax rate on all four of the market metrics. These results indicate that lower annual returns and higher annual volatility increase the fraction of foundations qualifying for the lower tax rate. However, the sign on the volatility of the last month of the fiscal year reverses relative to the relationship depicted in Figure 3.6. This indicates that conditional on the volatility of the rest of the year, increased volatility in the final month decreases the fraction of foundations qualifying for the low investment tax rate. This result suggests that some foundations trying to achieve the low investment tax rate fail to do so if there is too much market movement in the final month of the fiscal year.

In Table 3.7 I examine how market returns affect the fraction of foundations whose spending bunches around the investment tax cutoff. This bunching indicates that the investment tax threshold is binding. Each column shows results for different bunching definitions. For example, column 1 shows results when bunching is defined as spending levels being within plus or minus 15 percent of the investment tax threshold. Column 6 shows results when the bunching indicator equals 1 if the spending level is greater than or equal to the investment tax threshold, but does not exceed it by more than 5 percent. For these regressions, I use the same underlying data as for Table 3.6, with the additional exclusion of observations for which the investment tax threshold is less than twice as large as the minimum payout threshold.[13] To see what these bunching ranges capture, the bars in Figures 3.9 and 3.10 each represent 1 percent of the investment tax threshold, providing details for the histograms originally presented in Figures 3.3 and 3.5.[14]

[13]This ensures that bunching measures constraint by the investment tax threshold and not the minimum payout threshold.

[14]Note that the bars in Figures 3.3 and 3.5 each represent 5 percent of the investment tax threshold.

The negative coefficients on Annual S&P500 volatility in Table 3.7 indicate that high annual volatility in market returns decreases bunching around the investment tax threshold. Volatility in investment returns increases the difficulty of achieving the threshold accurately and may discourage foundations from attempting the feat. Statistically significant for the majority of bunching definitions, the consistently positive coefficient on Annual S&P500 return shows that higher returns increase bunching. This indicates that higher market returns – which increase the benefit of the lower tax rate – cause foundations to increase spending in order to qualify for the low tax rate. This evidence clearly shows that the net investment income tax does incentivize foundation spending increases in some periods, though not necessarily total spending over time.

At a first glance, the results in Tables 3.6 and 3.7 may appear contradictory. Higher annual returns and lower annual volatility simultaneously decrease the fraction of foundations qualifying for the lower tax rate and increase bunching at the investment tax threshold. Taken in combination, these results show that higher annual returns make it more difficult and more valuable to achieve the low tax rate.

Table 3.8 examines the effect of market returns on individual foundations qualifying for the lower tax rate or bunching at the investment tax threshold, allowing for interaction with foundation-level asset allocations to corporate stocks in the first year observed. Columns, examining low tax rates without asset allocation interactions, show results similar to those in Table 3.6. However, column 2 indicates that the effect of annual volatility is strongly tied to asset allocations, while annual returns are not. Accounting for asset allocations also suggests that foundations with a high exposure to stocks are less likely to qualify for the low tax rate in fiscal years when they experience high returns in the final month. Column 3 results are consistent with those in Table 3.7, except that returns do not appear to be significant. Column 4 shows that the negative impact of market volatility on bunching increases with allocation to stocks, a result consistent with the notion that the uncertainty of the investment income tax makes it difficult to respond to.

In combination, the results presented in this section show that market returns have a significant effect on foundations abilities and/or incentives to respond to the low investment income tax rate incentive. Future analyses that would be worth performing include developing an instrument for foundation investment income by interacting starting asset allocation to corporate stocks with annual stock market returns.

3.8 Conclusion

In this paper I have made several contributions to the literature examining foundation spending with regard to the minimum payout and net investment income taxes faced by foundations in the United States. First, I developed a model of the foundation's intertemporal spending problem subject to these taxes. This simplified model demonstrates the complexity of the foundations' constrained optimization problems. I then characterized in a tractable way spending strategies that foundations can employ in response to the complex problem. Using a hypothetical setting with and without uncertainty in investment returns, I have explored how these strategies fare in trading off increased spending with increased endowment value preservation. I find that minimizing spending without regard to the investment tax rate appears to preserve endowment spending power the most. Spending a constant percent of a moving average of asset values appears to minimize variance in year-to-year spending growth, while a strategy that involves alternating the targeting of the lower tax rate and the minimum payout rate (subject to always meeting the minimum payout threshold) balances spending with endowment preservation.

This paper also makes several empirical contributions. Using foundation tax return filing data, I have examined how binding the tax constraints are. Consistent with existing literature, I have shown that the minimum payout tax threshold appears to be binding. I have also shown that this payout threshold correlates strongly with the investment tax threshold. I build on existing literature by examining separately the subset of observations for which the investment tax threshold is at least double the payout threshold. Using this subset, I have shown that foundation spending also bunches around the investment tax threshold. Furthermore, spending bunches around the investment tax threshold more in years when foundations experience their highest returns than in years when they experience their lowest returns.

My final contribution to the literature is to examine foundation tax levels and incentives in combination with exogenous contemporaneous market returns. Here I show that foundations are more likely to be constrained by the investment tax threshold and less likely to achieve it under higher annual market returns and lower annual volatility. This contributes new evidence to the literature by showing that higher investment returns make it more harder to achieve the low tax rate but also increase spending by foundations in order to achieve the low tax rate. Note that these spending increases may be short-term and offset in other periods over the long run. Accounting for intial foundation level asset allocations to corporate stocks appears to amplify these basic effects.

Future work that could be promising includes an examination of intertemporal effects of returns. Adding lags may show that contemporaneous spending shifts spurred by market returns and volatility are offset in other periods. It would also be interesting to examine the results in this paper separately for foundations that do and do not receive significant donations and by other categories that might affect foundation responses to market forces.

The taxes that private foundations face in the United States are complicated. While the minimum payout tax does appear to be binding, the extent to which the notched net investment income tax incentivizes increased total spending (versus shifting of spending between high and low return periods or random qualification for the lower tax rate), remains unclear. The results in this paper suggest that foundations are more likely to target the lower tax rate when returns are higher but are more likely to qualify for it when returns are lower (by construction). This suggests that the notched net investment income tax incentivizes foundations to increase spending when markets are performing well. Are boom times the years when we want foundations to increase their charitable giving?

Chapter 3 has material that is being prepared for submission for publication. The dissertation author is the principal author of chapter 3.

3.9 Tables and Figures

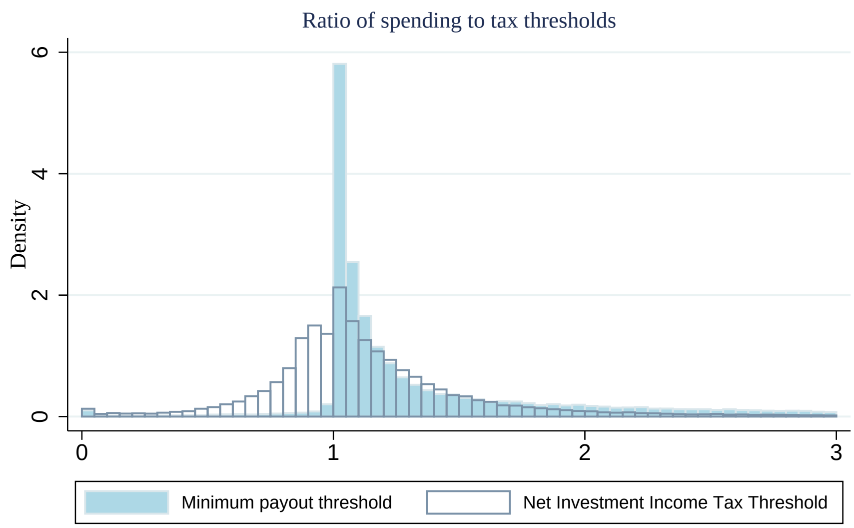
Table 3.1. Comparison of simulated outcomes for different spending policies

Policy	Qualifying dist's ($)	Taxes paid on inv income ($)	Taxes as pct of spending (%)	Final assets ($)	Years with low tax rate[1]	Std dev of YoY change in qual dist's (%)
Constant 8% annual rate of return						
Minimum payout threshold	1,288,960	41,388	3.11	1,819,555	0	0.8
Investment tax threshold[3]	1,328,813	20,472	1.52	1,718,344	19	0.4
Min inv inc taxes every other year[3]	1,294,450	30,552	2.31	1,804,842	10	3.8
Constant percent of assets	1,319,083	40,475	2.98	1,743,755	0	0.0
Constant 5 pct of moving avg assets[2]	1,300,304	41,036	3.06	1,790,573	0	0.5
1992-2011 S&P 500 returns						
Minimum payout threshold	1,404,477	40,326	2.79	1,120,721	7	17.2
Investment tax threshold[3]	1,415,654	7,692	0.54	1,069,095	15	111.4
Min inv inc taxes every other year[3]	1,415,727	25,397	1.76	1,108,908	14	32.7
Constant 5 pct of starting assets	1,429,891	45,033	3.05	1,095,961	5	18.6
Constant 5 pct of moving avg assets[2]	1,438,300	38,630	2.62	1,073,496	8	12.5

All simulations start with assets of $1 million and qualifying distributions of $50,000 in the first year. Spending is evenly allocated over the twelve months of the year. Constant returns are allocated evenly over the 12 months of the year. S&P 500 returns are accrued on a monthly basis.

1. Foundations may not qualify for the lower investment income tax rate in the first year.
2. Moving average is calculated over the previous three years' ending assets.
3. In years with negative returns, the payment was set to the minimum required to avoid the 30% undistributed income penalty. This threshold was met in all years.

Source: S&P 500 return data accessed at http://finance.yahoo.com.



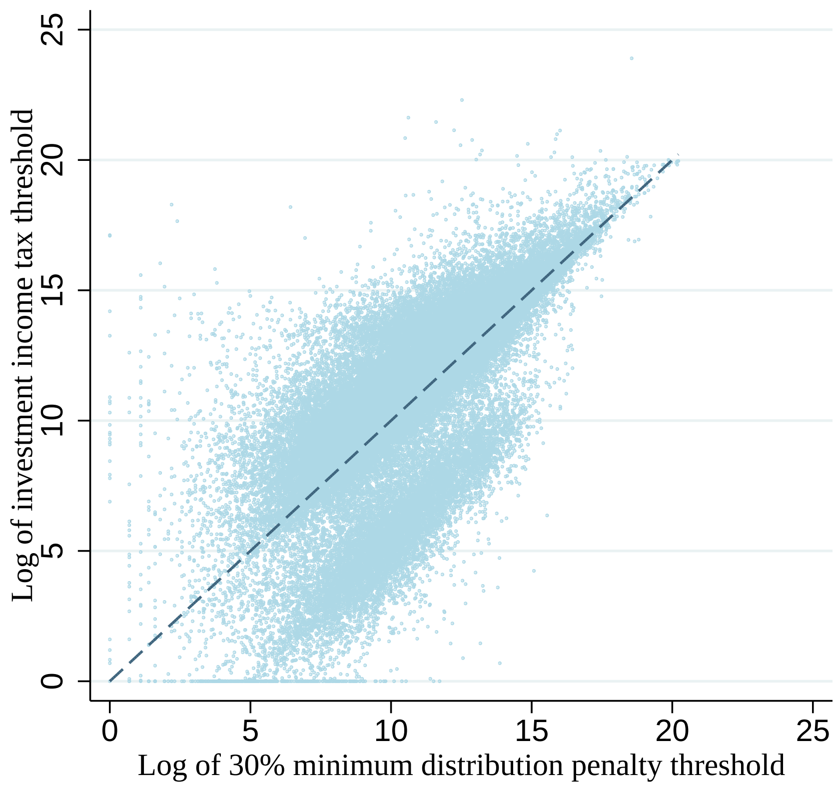
Histogram has been truncated at 3, after which density continues to drop off.
This graph is based on 41185 observations. Observations with minimum payout thresholds of 0 are excluded.
Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics

Figure 3.1. Foundation spending relative to minimum payout and investment tax thresholds

Table 3.2. Summary statistics for foundation financials

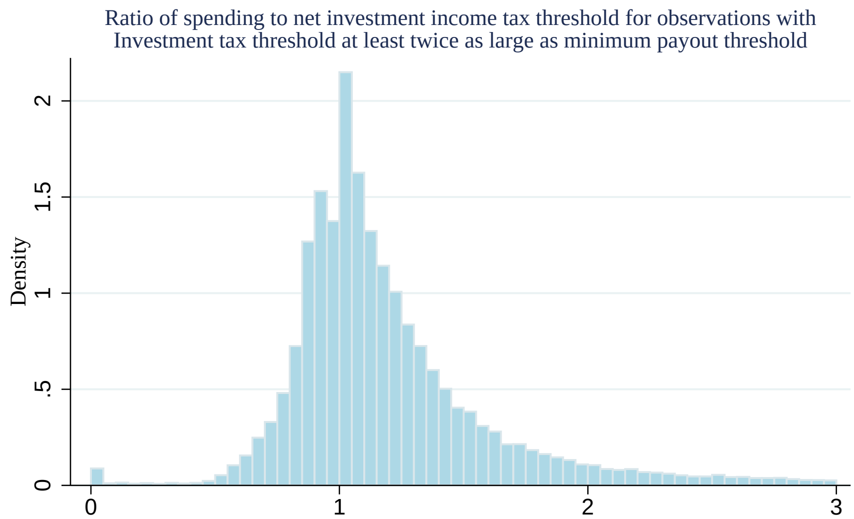
VARIABLES	N	mean	sd	min	max
Fair mkt val of year-end assets ($ mil)	139,262	45	391	0	38,736
Net val of nonchar-use assets ($ mil)	139,262	41	362	0	34,977
Qualifying distributions ($ mil)	139,262	3	26	0	3,653
Net investment income ($ mil)	139,262	3	48	-315	11,275
Fraction of assets in corporate stocks	139,262	0	0	0	1
Foundation qualifies for low tax rate	139,262	0	0	0	1
Savings from low tax rate ($ thou)	139,262	34	479	0	112,748
Investment return	139,262	0	1	-37	363
Capital gains percent	139,262	0	1	0	358

Note: Excludes observations with zero assets at beginning or end of year, foundations with fewer than five years worth of observations, initial or final observation in data if outside start or end of dataset, filings marked as being in termination proceedings, filings marked as initial or final, observations for which it is impossible to impute an investment income tax rate, assets at start or end of year are less than $10,000, and observations with zero net noncharitable-use assets. Source: Internal Revenue Service (2011).



Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics

Figure 3.2. Investment income tax relative to minimum payout threshold levels



Histogram has been truncated at 3, after which density continues to drop off.
This graph is based on 36073 observations.
Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics

Figure 3.3. Investment tax threshold bunching for observations with investment tax threshold at least twice as large as the minimum payout threshold

This graph is based on 11856 and 12132 observations for max and min return years, respectively.
Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics

Figure 3.4. Years in which foundations experience maximum and minimum observed investment returns (percentage-wise)

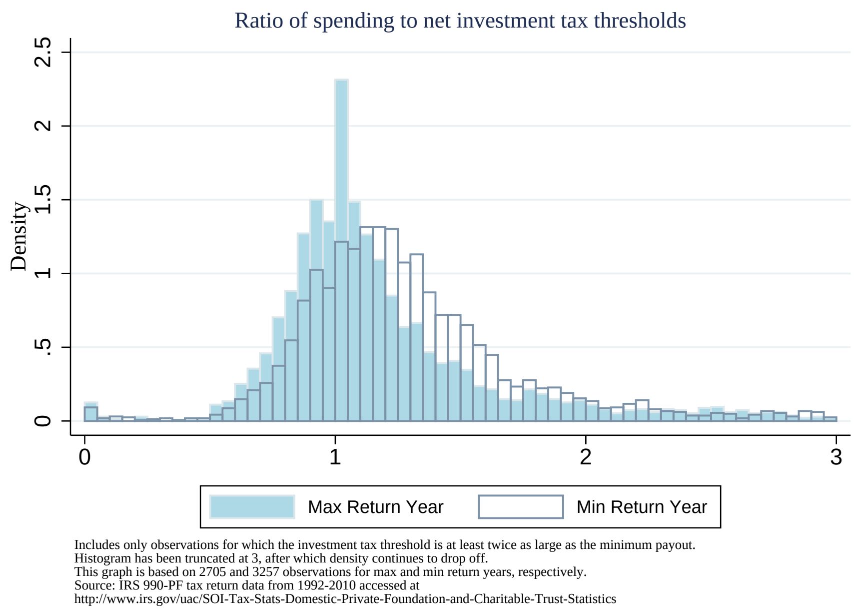


Figure 3.5. Investment tax threshold bunching in years with highest versus lowest returns

Table 3.3. Regressions of log(qualifying distributions) on tax thresholds and net investment income

Log of qualifying distributions	(1)	(2)	(3)	(4)
Log of payout threshold	0.249***		0.248***	0.235***
	(0.015)		(0.014)	(0.013)
Log of inv inc tax threshold		0.082***	0.046***	0.036***
		(0.007)	(0.009)	(0.008)
Log of net inv income				0.104***
				(0.013)
Observations	60363	134643	60272	59636
R^2	0.948	0.883	0.949	0.951
Adjusted R^2	0.939	0.871	0.940	0.942

Standard errors in parentheses
Regressions include year and foundation fixed effects. Errors are clustered at year level.
Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.4. Regressions of indicator for foundations qualifying for low tax rate on tax filing variables

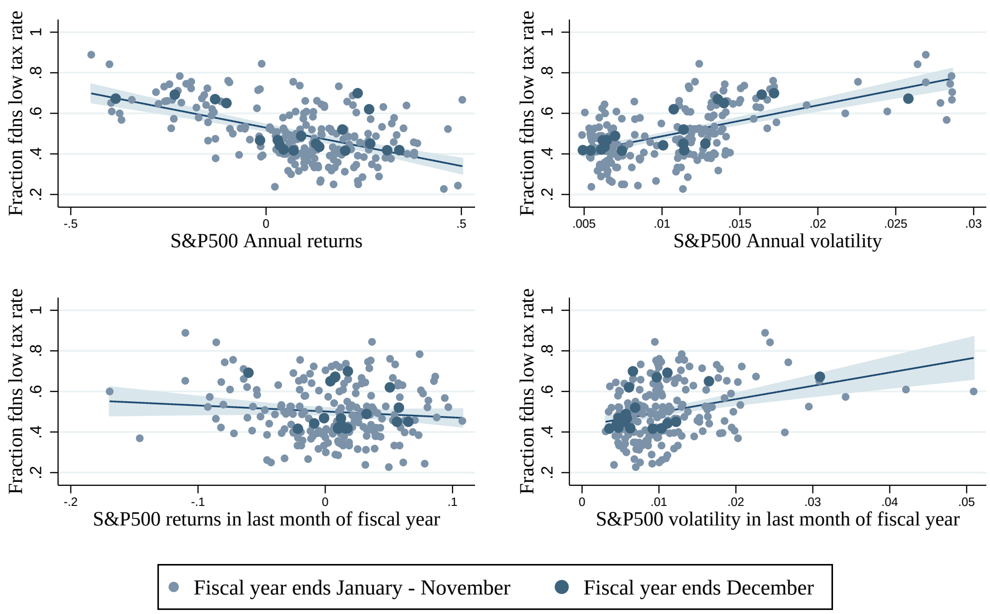
	(1)	(2)	(3)	(4)
Foundation qualifies for low tax rate				
Inv inc threshold > payout threshold	-0.4650***	-0.4647***	-0.4577***	-0.4694***
	(0.0147)	(0.0146)	(0.0127)	(0.0087)
Investment return	0.0027**	0.0027**	0.0123***	0.0104**
	(0.0011)	(0.0010)	(0.0039)	(0.0048)
Lowest return year		0.0148	-0.0179**	-0.0110
		(0.0110)	(0.0072)	(0.0088)
Highest return year		0.0043	0.0362***	0.0410***
		(0.0126)	(0.0081)	(0.0073)
Capital gains percent			-0.0095**	-0.0081
			(0.0040)	(0.0050)
Lowest capital gain year (pct basis)			0.0644***	0.0723***
			(0.0216)	(0.0194)
Highest capital gain year (pct basis)			-0.0255***	-0.0257***
			(0.0081)	(0.0088)
Last year's inv ret				-0.0006
				(0.0061)
Last year's cap gains pct				-0.0013
				(0.0066)
Observations	130708	130708	130708	113814
R^2	0.2431	0.2431	0.2457	0.2711
Adjusted R^2	0.1670	0.1670	0.1698	0.1859

Standard errors in parentheses
Regressions include foundation fixed effects.
Errors have been clustered at the year level.
Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5. Regressions of indicator for foundations qualifying for low tax rate on tax filing variables – with year fixed effects

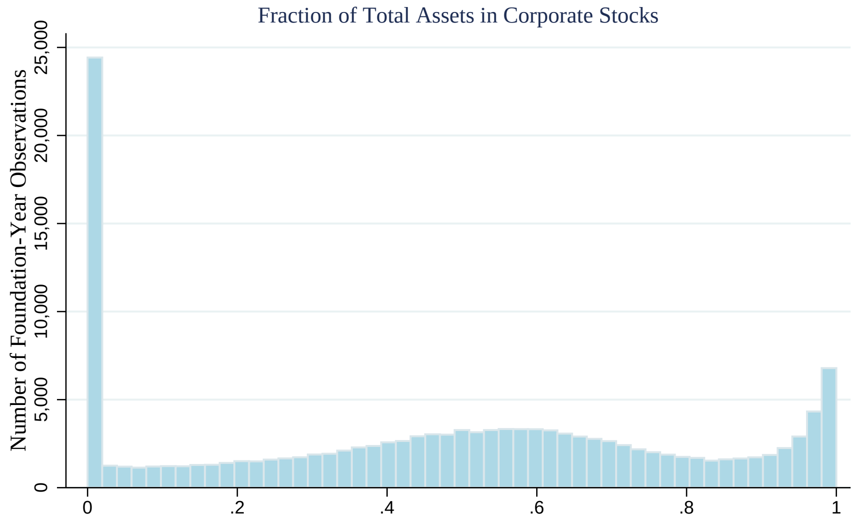
	(1)	(2)	(3)	(4)
Foundation qualifies for low tax rate				
Inv inc threshold > payout threshold	-0.4334***	-0.4339***	-0.4335***	-0.4438***
	(0.0137)	(0.0136)	(0.0136)	(0.0103)
Investment return	0.0041*	0.0035**	0.0217***	0.0162**
	(0.0020)	(0.0015)	(0.0066)	(0.0057)
Lowest return year		-0.0302***	-0.0328***	-0.0228***
		(0.0043)	(0.0041)	(0.0054)
Highest return year		0.0238***	0.0438***	0.0461***
		(0.0062)	(0.0068)	(0.0075)
Capital gains percent			-0.0187**	-0.0138**
			(0.0068)	(0.0060)
Lowest capital gain year (pct basis)			0.0066	0.0100
			(0.0070)	(0.0074)
Highest capital gain year (pct basis)			-0.0283***	-0.0268***
			(0.0059)	(0.0075)
Last year's inv ret				0.0021
				(0.0052)
Last year's cap gains pct				-0.0040
				(0.0056)
Observations	130708	130708	130708	113814
R^2	0.2614	0.2619	0.2621	0.2854
Adjusted R^2	0.1871	0.1875	0.1877	0.2016

Standard errors in parentheses
Regressions include foundation and year fixed effects.
Errors have been clustered at the year level.
Source: IRS 990-PF tax return data from 1992-2010 accessed at http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



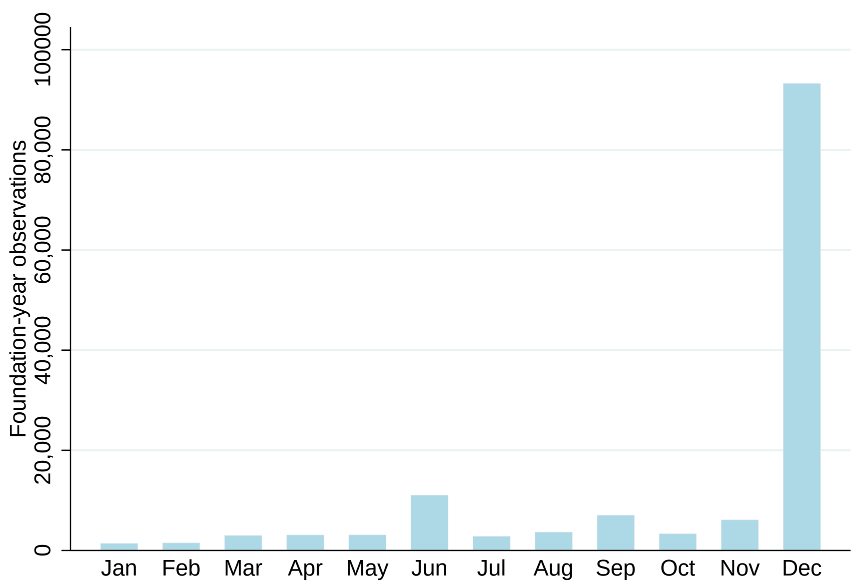
Sample includes only foundations represented in data all 19 years. 64% of observations are for foundations with fiscal years ending in December.
Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics
S&P500 values accessed at http://finance.yahoo.com.

Figure 3.6. Scatter plots of fraction of foundations qualifying for low tax rate relative to stock market returns and volatility



Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics
Fraction equals Investments -- corporate stock (II.10b) divided by Total Assets (II.16) for the beginning of year (column (a)).

Figure 3.7. Observed asset allocations to corporate stocks



Source: IRS 990-PF tax return data from 1992-2010 accessed at
http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics

Figure 3.8. Distribution of fiscal year ends

Table 3.6. Regressions of fraction of foundations qualifying for low investment tax rate

	(1)	(2)	(3)	(4)	(5)
Fraction of fdns qualifying for low tax rate					
Annual S&P500 return	-0.39***				-0.22***
	(0.06)				(0.07)
Annual S&P500 volatility		15.65***			15.00***
		(1.64)			(2.66)
Monthly S&P500 return			-0.32		-0.17
			(0.24)		(0.16)
Monthly S&P500 volatility				6.68***	-4.73***
				(1.80)	(1.49)
Observations	228	228	228	228	228
R^2	0.2790	0.3719	0.0114	0.1036	0.4488
Adjusted R^2	0.2758	0.3691	0.0070	0.0996	0.4389

Standard errors in parentheses
Includes only foundations with a mix of high and low tax rates and with observations for every year. Regressions are clustered at the year-quarter level.
Source: IRS 990-PF tax return data from 1992-2010 accessed at http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics and http://finance.yahoo.com
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7. Regressions of fraction of foundations bunching around investment tax threshold, with varying bunching definitions

	(1)	(2)	(3)	(4)	(5)	(6)
	$\pm 15\%$	$\pm 10\%$	$\pm 5\%$	$\pm 2\%$	-2 to $+5\%$	0 to $+5\%$
Annual S&P500 ret	0.24**	0.27***	0.18**	0.02	0.09	0.09**
	(0.10)	(0.08)	(0.07)	(0.04)	(0.05)	(0.05)
Annual S&P500 vol	-17.89***	-13.27***	-7.22***	-2.93**	-4.96**	-3.93**
	(3.79)	(2.96)	(2.53)	(1.41)	(1.93)	(1.66)
S&P500 ret in last month of fiscal yr	-0.29	-0.46*	-0.20	-0.01	-0.02	-0.07
	(0.35)	(0.27)	(0.18)	(0.13)	(0.15)	(0.13)
S&P500 vol in last month of fiscal yr	11.11***	8.23***	2.11	-0.83	0.69	0.84
	(2.63)	(2.27)	(1.68)	(1.02)	(1.26)	(1.19)
Observations	228	228	228	228	228	228
R^2	0.2122	0.2126	0.1758	0.0672	0.1077	0.1025
Adjusted R^2	0.1981	0.1984	0.1610	0.0505	0.0917	0.0864

Standard errors in parentheses
Includes only foundations with a mix of high and low tax rates and with observations for every year. Regressions are clustered at the year-quarter level.
Source: IRS 990-PF tax return data from 1992-2010 accessed at http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics and http://finance.yahoo.com
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

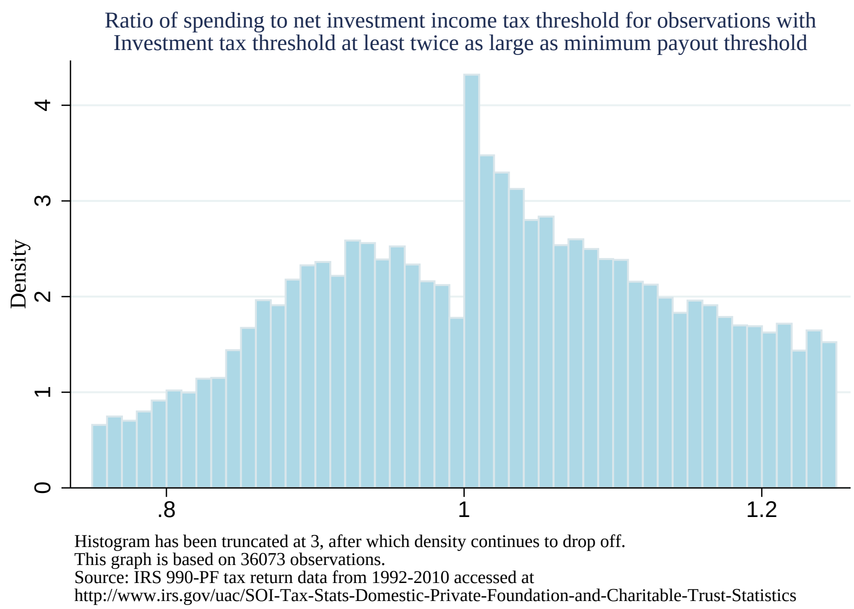


Figure 3.9. Detail of investment tax threshold bunching for observations with relatively large investment tax thresholds

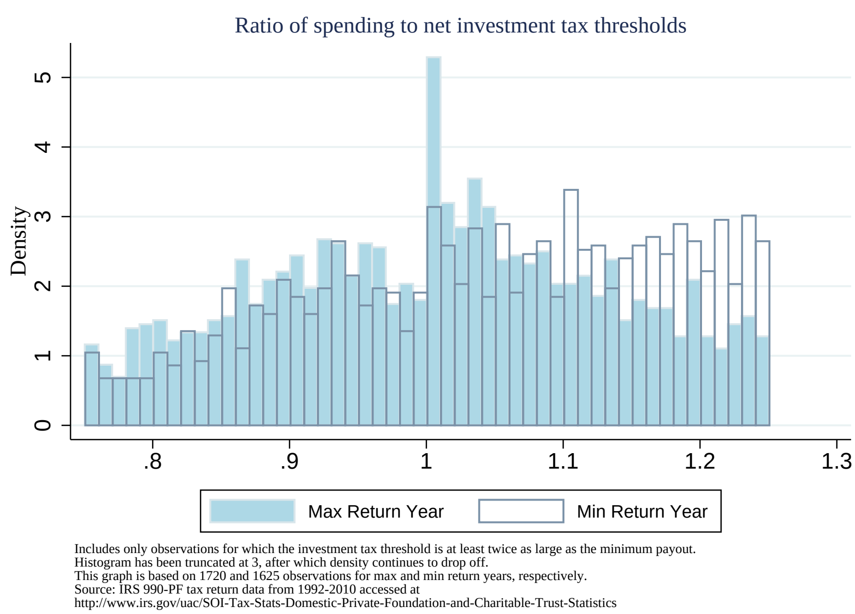


Figure 3.10. Detail of investment tax threshold bunching in years with highest versus lowest returns

Table 3.8. Foundation level regressions on market returns and volatility, using initial asset allocations

	Low tax rate		Bunching (±5%)	
	(1)	(2)	(3)	(4)
Annual S&P500 ret	-0.14**	-0.16***	0.03	0.09
	(0.06)	(0.05)	(0.03)	(0.08)
Annual S&P500 vol	16.90***	3.53	-9.81***	-5.19**
	(2.61)	(2.28)	(1.16)	(2.31)
S&P500 ret in last month of fiscal yr	-0.34	0.06	0.25	0.29
	(0.22)	(0.17)	(0.17)	(0.35)
S&P500 vol in last month of fiscal yr	-7.03***	-2.98**	3.84***	4.28**
	(1.59)	(1.49)	(1.08)	(2.05)
Annual S&P500 ret × Frac of assets in corp stocks (starting)		0.03		-0.08
		(0.06)		(0.09)
Annual S&P500 vol × Frac of assets in corp stocks (starting)		18.01***		-5.93**
		(1.49)		(2.98)
S&P500 ret in last month of fiscal yr × Frac of assets in corp stocks (starting)		-0.53**		-0.05
		(0.24)		(0.32)
S&P500 vol in last month of fiscal yr × Frac of assets in corp stocks (starting)		-5.37***		-0.56
		(1.76)		(1.94)
Observations	121671	121664	38922	38920
R^2	0.2087	0.2101	0.2144	0.2148
Adjusted R^2	0.1225	0.1241	0.0614	0.0617

Standard errors in parentheses
Regressions include foundation fixed effects and are clustered at the year-month level.
Includes only foundations that have a mix of high and low tax rates. Does not include first year observed for each foundation. Bunching regressions exclude observations for which investment tax threshold is less than double the minimum payout threshold.
Source: IRS 990-PF tax return data from 1992-2010 accessed at http://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics and http://finance.yahoo.com
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Bibliography

ABELSON, REED (2000): "Some Foundations Choose to Curb Donations and Pay More Taxes," *The New York Times*, February, 23.

BAHNG, YOUNGJIN, AND DORIS H. KINCADE (2012): "The relaitonship between temperature and sales: Sales data analysis of a retailer of branded women's business wear," *International Journal of Retail & Distribution Management*.

BAKER, MALCOLM, JEREMY C. STEIN, AND JEFFREY WURGLER (2003): "When Does the Market Matter? Stock Prices and the Investment of Equity-Dependent Firms," *Quarterly Journal of Economics*, 118(3), 969–1005.

BEATTY, TIMOTHY K.M., JAY P. SHIMSHACK, AND RICHARD J. VOLPE (2015): "Disaster preparedness and disaster response: Evidence from bottled water sales before and after tropical cyclones," *Working Paper*.

BERNER, MADDY (2015): "College Endowments See Another Year of Growth," online.

BERTRAND, JEAN-LOUIS, XAVIER BRUSSET, AND MAXIME FORTIN (2015): "Assessing and hedging the cost of unseasonal weather: Case of the apparel sector," *European Journal of Operational Research*.

BIGNAMI, PETRA G. (2013): "Sustainable Payout for Foundations: 2013 Update Study," Cambridge Associates LLC.

BLANCHARD, OLIVIER, CHANGYONG RHEE, AND LAWRENCE SUMMERS (1993): "The stock market, profit, and investment," *Quarterly Journal of Economics*, 108(1), 115–136.

BLANCHARD, OLIVIER JEAN, AND STANLEY FISHER (1989): *Lectures on Macroeconomics". Cambridge: The MIT Press*. MIT Press, Cambridge, Massachusetts.

BORIS, ELIZABETH T., LOREN RENZ, ASMITA BARVE, MARK A. HAGER, AND GEORGE HOBOR (2006): "Foundation Expenses & Compensation: How operating characteristics influence spending," Published jointly by the Urban Institute, Foundation Center, and GuideStar.

BUSSE, MEGHAN R., DEVIN G. POPE, JAREN C. POPE, AND JORGE SILVA-RISSO (2014): "The Psychological effect of weather on car purchases," *Quarterly Journal of Economics*.

CAPELLE-BLANCARD, GUNTHER, AND STEPHANIE MONJON (2012): "Trends in the literature on socially responsible investment: Looking for the keys under the lamppost," *Business Ethics: A European Review*, 21(3), 239–250.

CHEN, QI, ITAY GOLDSTEIN, AND WEI JIANG (2007): "Price informativeness and investment sensitivity to stock price," *Review of Financial Studies*, 20(3), 619–650.

CONLIN, MICHAEL, TED O'DONOGHUE, AND TIMOTHY J. VOGELSANG (2007): "Projection Bias in Catalog Orders," *American Economic Review*.

DEEP, AKASH, AND PETER FRUMKIN (2001): "The Foundation Payout Puzzle," Hauser Center for Nonprofit Organizations Working Paper No. 9. Available at SSRN: http://ssrn.com/abstract=301826orhttp://dx.doi.org/10.2139/ssrn.301826.

DELL, MELISSA, BENJAMIN F. JONES, AND BENJAMIN A. OLKEN (2014): "What do we learn from the weather? The new climate-economy literature," *Journal of Eocnomic Literature*.

DING, NING, JERRY T. PARWADA, AND JIANFENG SHEN (2014): "When Does a Stock Boycott Work? Evidence from a Clinical Study of the Sudan Divestment Campaign," .

EMERSON, JED (2003): "Where Money Meets Mission: Breaking Down the Firewall Between Foundation Investments and Programming," *Stanford Social Innovation Review*, 1(2), 38–47.

FABOZZI, FRANK J., K. C. MA, AND BECKY J. OLIPHANT (2008): "Sin Stock Returns," *Journal Of Portfolio Management*, 35(1), 82–94.

GHEZ, GILBERT, AND GARY S. BECKER (1975): "A theory of the allocation of time and goods over the life cycle," *NBER*.

GOLDENBERG, SUZANNE (2014): "Heirs to Rockefeller oil fortune divest from fossil fuels over climate change," *The Guardian*, September 22.

GRAFF ZIVIN, JOSHUA, AND MATTHEW NEIDELL (2014): "Temperature and the

Allocation of Time: Implications for Climate Change," *Journal of Labor Economics*.

HANNAK, ANIKO, ERIC ANDERSON LISA FELDMAN BARRETT SUNE LEHMANN ALAN MISLOVE, AND MIREK RIEDEWALD (2012): "Tweetin' in the Rain: Exploring Societal-Scale Effects of Weather on Mood," *International AAAI Conference on Weblogs and Social Media*.

HASTIE, TREVOR, ROBERT TIBSHIRANI, AND MARTIN WAINWRIGHT (2015): *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC Press.

HEINKEL, ROBERT, ALAN KRAUS, AND JOSEF ZECHNER (2001): "The effect of green investment on corporate behavior," *Journal of Financial and Quantitative Analysis*, 36(4), 431–450.

HIRSHLEIFER, JACK (1966): "Investment decision under uncertainty: Applications of the state-preference approach," *Quarterly Journal of Economics*, 80(2), 252–277.

HONG, HARRISON, AND MARCIN KACPERCZYK (2009): "The price of sin: The effects of social norms on markets," *Journal of Financial Economics*, 93(1), 15–36.

HOWARTH, EDGAR, AND MICHAEL S. HOFFMAN (1984): "A multidimensional approach to the relationship between mood and weather," *British Journal of Psychology*.

INTERNAL REVENUE SERVICE (2011): "Table 1. Domestic Private Foundations: Number and Selected Financial Data, by Type of Foundation and Size of End-of-year Fair Market Value of Total Assets, Tax Year 2011," p://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics (accessed February 20, 2015).

IRVINE, WILLIAM B. (1987): "The ethics of investing," *Journal of Business Ethics*, 6(3), 233–242.

KLEIN, MARIEL A., AND THEODORE R. DELWICHE (2015): "Judge Dismisses Divestment Lawsuit," *The Harvard Crimson*, March 24.

KRAMER, MARK R., AND SARAH E. COOCH (2007): "The Power of Strategic Mission Investing," *Stanford Social Innovation Review*, 5(4), 43–51.

KUMAR, RAMAN, WILLIAM B. LAMB, AND RICHARD E. WOKUTCH (2002): "The End of South African Sanctions, Institutional Ownership, and the Stock Price Performance of Boycotted Firms Evidence on the Impact of Social/Ethical Investing," *Business and Society*, 41(2), 133–165.

LABOVITZ, JOHN R. (1974): "The Impact of the private foundation provisions of the Tax Reform Act of 1969: Early empirical measurments," *The Journal of Legal Studies*, 3, 63–105.

LAZO, JEFFREY K., MEGAN LAWSON, PETER H. LARSEN, AND DONALD M. WALDMAN (2011): "U.S. Economic sensitivity to weather variability," *American Meteorological Society*.

LEVI, ORI, AND ITAI GALILI (2008): "Stock purchase and the weather: Individual differences," *Journal of Economic Behavior and Organization*.

LI, CHENXI, ANDY REINAKER, CHENG ZHANG, AND XUEMING LUO (2015): "Weather and Mobile Purchases: 10-Million-User Field Study," *Working Paper*.

LODERER, CLAUDIO, JOHN W. COONEY, AND LEONARD D. VAN DRUNEN (1991): "The price elasticity of demand for common stock," *Journal of Finance*, 46(2), 621–651.

MACDONALD-SMITH, ANGELA (2014): "Winds blow cool for fossil fuels and iron," *Australian Financial Review*, September 24.

MAUNDER, W. J. (1973): "Weekly weather and economic activities on a national scale: An example using United States retail trade data," *New Zealand Meteorological Service*.

MAXMIN, CHLOE, BOROWSKY SARAH, AND STUDENTNATION (2013): "The Time To Divest: A Response to Harvard President Drew Faust," *The Nation*, October 15.

MCKEOWN, W. B. (1997): "On being true to your mission: social investments for endowments," *Journal of Investing*, 6(4), 71–78.

MCWILLIAMS, ABAGAIL, AND DONALD SIEGEL (1997): "Event studies in management research: Theoretical and empirical issues," *Academy of Management Journal*, pp. 626–657.

MELILLO, JERRY M., TERESE (T.C.) RICHMOND, AND GARY W. YOHE (eds.) (2014): *Climate Change Impacts in the United States: The Third National Climate Assessment*. U.S. Global Change Research Program.

MERTON, ROBERT C. (1987): "A simple model of capital market equilibrium with incomplete information," *Journal of Finance*, 42(3), 483–510.

MEZNAR, MARTIN B., DOUGLAS NIGH, AND CHUCK CY KWOK (1994): "Effect of announcements of withdrawal from South Africa on stockholder wealth.," *Academy of Management Journal*, 37(6), 1633–1648.

——— (1998): "Announcements of withdrawal from South Africa revisited: making sense of contradictory event study findings," *Academy of Management Journal*, pp. 715–730.

MINKEN, HARALD (2008): "Systematic risk, and how it is taken account of in the discount rate of Norwegian transport cost benefit analyses," Madrid. ECTRI TWG G on Transport Economics and Policy.

MORCK, RANDALL, ANDREI SHLEIFER, ROBERT W. VISHNY, MATTHEW SHAPIRO, AND JAMES M. POTERBA (1990): "The stock market and investment: is the market a sideshow?," *Brookings Papers on Economic Activity*, 1990(2), 157–215.

MYERS, BRUCE, AND PAIGE ROBERTS (2008): "National Analysis of Federal Excise Tax on Capital Gains," Cambridge Associates report prepared for Council of Michigan Foundations.

OFFICE OF MANAGEMENT AND BUDGET (2016): "Budget of the U.S. Government, Fiscal Year 2017," Published online at https://www.whitehouse.gov/omb/budget/Overview, accessed on May 11, 2016.

PARSONS, ANDREW G. (2001): "The Association between daily weather and daily shopping patterns," *Australasian Marketing Journal*.

PETAJISTO, ANTTI (2009): "Why do demand curves for stocks slope down?," *Journal of Financial and Quantitative Analysis*, 44(5), 1013.

PILLER, CHARLES (2007): "Gates Foundation to keep its investment approach," *Los Angeles Times*, January 14, sec. A, page 23.

PILLER, CHARLES, EDMUND SANDERS, AND ROBYN DIXON (2007): "Dark cloud over good works of Gates Foundation," January 7, sec. A, page 1.

POSNIKOFF, JUDITH F. (1997): "Disinvestment from South Africa: They did well by doing good," *Contemporary Economic Policy*, 15(1), 76–86.

RENNEBOOG, LUC, JENKE TER HORST, AND CHENDI ZHANG (2008): "Socially responsible investments: Institutional aspects, performance, and investor behavior," *Journal of Banking and Finance*, 32(9), 1723–1742.

RIVOLI, PIETRA (2003): "Making a difference or making a statement? Finance research and socially responsible investment," *Business Ethics Quarterly*, 13(3), 271–287.

ROTH TRAN, BRIGITTE (2015): "Divest, Disregard, or Double Down?," *Working Paper*.

SANSING, RICHARD, AND ROBERT YETMAN (2006): "Governing private foundations using the tax law," *Journal of Accounting and Economics*, 41, 363–384.

SCHWARTZ, JOHN (2014): "Rockefellers, Heirs to an Oil Fortune, Will Divest Charity of Fossil Fuels," *New York Times*, September 22, sec. A, page 3.

SHALIT, HAIM, AND SHLOMO YITZHAKI (1994): "Marginal conditional stochastic dominance," *Management Science*, 40(5), 670–684.

SMITH, AARON (2013): "PETA takes a stake in SeaWorld, demands whales be freed," *CNN Wire*, April 24.

SMITH, KEITH (1993): "The influence of weather and climate on recreation and tourism," *Weather*.

SOLOMON, LEWIS D., AND KAREN C. COE (1997a): "The legal aspects of social investing by non-profit fiduciaries," *Journal of Investing*, 6(4), 112–119.

——— (1997b): "Social investments by nonprofit corporations and charitable trusts: A legal and business primer for foundation managers and other nonprofit fiduciaries," *University of Missouri - Kansas City Law Review*, 66, 213–250.

SPECTOR, MIKE (2009): "Foundation Tax-Policy Change Aims to Encourage Giving," *The Wall Street Journal*, March 24.

SPIES, KORDELIA, FRIEDRICH HESSE, AND KERSTIN LOESCH (1997): "Store atmosphere, mood and purchasing behavior," *International Journal of Research in Marketing*.

STARR-MCCLUER, MARTHA (2000): "The Effects of Weather on Retail Sales," *Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board*.

STEELE, A.T. (1951): "Weather's effect on the sales of a department store," *Journal of Marketing*.

SWEATTE, NATASHA (2014): "PETA buys SeaWorld stock in an attempt to end killer whale captivity for good," *The CW, San Diego 6*, August 19.

TEOH, SIEW HONG, IVO WELCH, AND C. PAUL WAZZAN (1999): “The Effect of Socially Activist Investment Policies on the Financial Markets: Evidence from the South African Boycott,” *Journal of Business*, 72(1), 35–89.

TUCKER, PATRICIA, AND JASON GILLILAND (2007): “The effect of season and weather on physical activity: a systematic review,” *Public health*.

U.S. SECURITY AND EXCHANGE COMMISSION (1998): “Final Rule: Amendments to Rules on Shareholder Proposals,” http://www.sec.gov/rules/final/34-40018.htm (accessed April 26, 2013).

US SIF THE FORUM FOR SUSTAINABLE AND RESPONSIBLE INVESTMENT (2014): “Unleashing the Potential of US Foundation Endowments: Using Responsible Investment to Strengthen Endowment Oversight and Enhance Impact,” online.

WAGNER, RICHARD E. (1977): “Tax policy toward private foundations: Confused principles and unfortunate legislation,” *Policy Studies Journal*, 5, 314–319.

WILHELM, IAN, AND GRANT WILLIAMS (2009): “Senators Propose Measure to Encourage Foundations to Give More,” *The Chronicle of Philanthropy*, March 25.

WINES, MICHAEL (2014): “Stanford to Purge $18 Billion Endowment of Coal Stock,” *New York Times*, May 6.

WOLFERS, JUSTIN, AND ERIC ZITZEWITZ (2006): “Five Open Questions about Prediction Markets,” in *Information Markets: A New Way of Making Decisions*, ed. by Robert W. Hahn, and Paul C. Tetlock, pp. 13–36. AEI-Brookings Joint Center for Regulatory Studies, Washington, D.C.

WOOD, DAVID, AND LISA HAGERMAN (2010): “Mission investing and the philanthropic toolbox,” *Policy and Society*, 29(3), 257–268.

YODER, TIMOTHY R., NOEL ADDY, AND BRIAN P. MCALLISTER (2011): “Tax-Motivated Increases in Qualifying Distributions by Private Foundations,” *Journal of the American Taxation Association*.

YODER, TIMOTHY R., AND BRIAN P. MCALLISTER (2012): “Do Private Foundations Increase Current Distributions to Qualify for a 50 Percent Tax Rate Reduction?,” *Journal of the American Taxation Association*, 34, 45–66.