Dirty Air and Clean Investments: The impact of pollution information on ESG investment

Raymond Fisman* Pulak Ghosh† Arkodipta Sarkar‡ Jian Zhang§

January 30, 2023

Preliminary draft: Please do not circulate or cite without permission

Abstract

We study the link between exposure to pollution information and investment portfolio allocations, exploiting the rollout of air quality monitoring stations during 2006-2019 in India. Using a triple-difference framework, we show that retail investors' investments in "brown" stocks become more negatively related to local air pollution after a monitoring station appears nearby. Since green stocks do not outperform brown stocks over this period, we suggest that our findings are likely driven by investor tastes and pollution salience rather than a shift in expected returns.The effect of pollution information on investment choices is most prominent amongst tech-savvy investors who are most plausibly "treated" by real-time pollution data, and by younger investors who tend to be more sensitive to environmental concerns. Overall, our results provide micro-level support for the view that salience of environmental conditions affect investors' tastes for green versus brown investments.

*Raymond Fisman is at Boston University. email:: rfisman@bu.edu

†Pulak Ghosh is at the Indian Institute of Management Bangalore. email: pulak.ghosh@iimb.ac.in

‡Arkodipta Sarkar is at the National University of Singapore. email: asarkar@nus.edu.sg

§Jian Zhang is at Hong Kong University email: zhangj1@hku.hk

1 Introduction

Globally, investment in so-called ESG (environmental, social, governance) funds has grown by over 500 percent, from US$4.8 trillion in 2010 to US$29.2 trillion in 2021, nearly three times the rate of growth of assets under management more generally.[1] Much of the growth and attention has focused on the "E" in ESG, with sustainable investment seen as one mechanism for disciplining firms that generate negative environmental externalities.

Much of the focus and attention on sustainability in investment appears timed to increasing concerns about pollution generally and climate change specifically. Yet there is little empirical evidence on the forces that have led to the rise in ESG investment Hong, Karolyi and Scheinkman (2020). It may be driven by "supply"—a proliferation of ESG funds that provide ready investment opportunities. Alternatively, increased ESG investment may result from a shift in investor preferences due to the greater relevance and salience of environmental issues resulting from, e.g., global accords such as the Paris Agreement or increased frequency of extreme climate events, e.g., hurricanes, forest fires, floods, droughts, and heat waves.

Yet identifying a role for investor tastes is a challenge – news about climate change often serves as a common shock which in turn may be confounded by concurrent events that may in turn also impact portfolio allocations. If one wishes to exploit panel variation, one requires shocks that impact investor tastes but do not, for example, affect perceived returns of green versus brown investments. Additionally, one requires highly disaggregated investor data—with linkages to geography or some other source of exposure to environmental conditions—in order to take advantage of any panel variation that one might exploit.

In this paper we document that greater awareness of pollution leads investors to invest less in "brown" industries, using a difference-in-differences framework applied to geocoded data on the portfolios of Indian retail investors. We take advantage of the introduction of continuous air quality monitoring stations (CAAQMSs) throughout India during 2007-2021. We posit that the arrival of a

[1] See https://www.unpri.org/about-us/about-the-pri, accessed January 8, 2023.

monitoring station makes pollution more salient for those located nearby, who receive easy access to air quality data. Whereas pollution may have simply been ignored prior to the appearance of CAAQMSs, we assert that investors may have a taste for greener investments in response to more extreme pollution once air quality information becomes available in an easy-to-access (via smartphone) format. Helpful for identification, the monitoring stations' readings were advertised as relevant for a range of 20 kilometers, offering a natural way of defining investors who are "treated" with readier access to pollution information, which we compare to the portfolio allocations of "control" investors that are more distant from monitoring stations (but still close enough to the treatment population that they plausibly serve as a relevant benchmark).

We utilize a comprehensive database of trading records from the National Stock Exchange of India (NSE), which we use to construct the portfolios of approximately 19 million investors during 16 years.We may trace any trades in these portfolio to the individual's (anonymized) permanent account number, and the associated postal index number (PIN) code for the account holder. We may thus calculate, with a high degree of precision, the distance between an investor's address and the nearest CAAQMS, and evaluate how the sensitivity of their portfolio to the Air Quality Index (AQI) changed after the monitoring program made this information widely available. (As we explain below, we may observe—albeit imperfectly—air quality prior to the creation of CAAQMSs via satellite data, though importantly, data from these stations were not made available to the public in real-time. This allows us to distinguish the role of salience from shifts in actual pollution that might be correlated with the creation of monitoring stations.)

Specifically, we examine the sensitivity of an investors' holdings in "brown" stocks, as defined by Choi, Gao and Jiang (2020), to the extent of air pollution before versus after the opening of a nearby monitoring station. In an event study specification, we show that the brown share of investors in an impacted PIN code is stable in the two years leading up to the appearance of station, then experiences an immediate and sustained decline. Our point estimates indicates that, for a one standard deviation decline in the (within-monitoring-station) AQI, there is a 0.5 percentage point (1.25 percent of the mean) decrease in the share of brown stocks in affected investors' portfolios after a monitoring station appears (whereas the sensitivity is approximately zero beforehand). We use a "donut hole" approach to define a benchmark set of investors, located in PIN codes in the 40-60 kilometer range around a given monitoring station—these investors are sufficiently distant from the station that its reading had less relevance (the range communicated to the public was just 20 kilometers) yet close enough to 'treated' investors to offer a plausible control group. For this group of "donut hole" investors we observe no change in AQI sensitivity after a monitoring station opens. In an additional set of results, we explore heterogeneity in investors' responses to exposure to air quality information.

Given that AQI information from monitoring stations was delivered primarily via a smartphone app, we suggest that salience may have been greater amongst tech savvy investors who were also more apt to trade via mobile phone. Thus, we begin by splitting the sample based on whether a trader most often executed transactions via mobile device, the internet, or by some other means (i.e., through either a trading kiosk or a broker). We find a far greater shift in responsiveness to pollution after a monitoring station appears amongst mobile-based investors; there is also a greater sensitivity for internet-based investors relative to those using more traditional methods.

We next explore heterogeneity by age. Beyond a moderately higher degree of tech savvy, young investors may also have greater environmental awareness given that they disproportionately bear the consequences of climate change and environmental degradation. We thus posit that the young will be most responsive to being confronted with pollution information.[2] We find that AQI information has the greatest effect on investment sensitivity for young investors, and the weakest effect for the elderly.

Finally, we split the sample by gender. While the motivation for this heterogeneity test is less straightforward than our other sample splits, we suggest that it links to the broader literature on the "environmental gender gap"—women express greater concern for environmental issues than men (e.g., Xiao and McCright, 2015). We extend this line of research to show that this gender gap applies as well to male versus female investment: the portfolios of women are more sensitive to AQI than men, once this information becomes readily available.

Our work connects most directly to the body of research that aims to understand investors' non-

[2]Alternatively, air quality may be more salient for the elderly, given the health consequences, though to the extent that this is the case, our results suggest that it is dominated by the aforementioned effects.

pecuniary concerns generally, and specifically their interest in ESG investments. Closest to our own work is that of Choi, Gao and Jiang (2020), which studies the link between weather and investment in a cross-city framework utilizing largely cross-country variation in temperature. They show that higher temperatures are associated with a decline in prices of "brown" stocks, driven by the trading activity of retail investors in particular. While this work suggests that (retail) investor tastes, driven by increased salience of climate concerns, play an important role in ESG investing, the analysis and conclusions are limited by the coarseness of the data. We are able to tie individual trading behavior to much more localized, high-frequency shifts in environmental conditions that allows for more credible identification, and also for the heterogeneity analyses that require individual-level data.

More broadly, our work sits at the intersection of several large and growing areas of inquiry: research on ESG investments, salience and investor behavior, and the salience of environmental problems and attitudes toward environmental issues.

As described in Hong, Karolyi and Scheinkman (2020), ESG research may generally be classified in one of several categories. Broadly speaking, one branch of research focuses on the extent to which climate (and resultant environmental) risks are incorporated into stock prices (e.g., Görgen et al., 2020; Bansal, Kiku and Ochoa, 2016) and other assets such as homes (Baldauf, Garlappi and Yannelis, 2020; Murfin and Spiegel, 2020) and agricultural land (Hong, Li and Xu, 2019). To the extent that environmental risks become more important over time in ways that are not fully anticipated, there may be differential returns for green versus brown companies. The question of whether socially responsible companies generated higher returns more generally has been much-studied, but without any clear resolution (e.g., Hong and Kacperczyk, 2009; Berg et al., 2022).

Our own work is much more closely tied to the strand of work that explores the link between the beliefs and preferences of investors, and their green versus brown portfolio allocation decisions. In addition to Choi, Gao and Jiang (2020), which focuses more on retail investors, Alok, Kumar and Wermers (2020) examine institutional investors' responses to climate-related disasters and find that nearby fund managers respond more sharply than more distant ones.

While we focus on pollution salience, there is a much larger literature which examines how inattention and salience impact portfolio allocation decisions, whether driven by the notability of past returns (Bordalo, Gennaioli and Shleifer, 2013; Cosemans and Frehen, 2021); the media (Huberman and Regev, 2001; Jiang et al., 2022); or ESG ratings themselves (Pelizzon, Rzeznik and Hanley, 2021).

Finally, moving beyond finance-focused applications, our work relates to the larger literature in social psychology and economics which explores whether exposure to (idiosyncratic) environmental shocks impact beliefs and attitudes toward climate change (Li, Johnson and Zaval, 2011; Zaval et al., 2014; Lujala, Lein and Rød, 2015). Moving beyond attitudes to actions, Barwick et al. (2019) examine how pollution information—delivered by the same type of real-time monitoring stations that we consider here—impacted online search behavior in China, with more searches related to pollution avoidance behaviors after a monitoring station appears. Our work also documents real behavioral changes, but in a very different domain, and one that has broader social consequences rather than one with private benefits.

2 Institutional Background

This section provides an overview of India's rollout of continuous ambient air quality monitoring stations (CAAQMSs), which we exploit in our empirical design. The introduction of CAAQMSs was a part of a broader National Air Quality Monitoring Programme (NAMP), set up by the Central Pollution Control Bureau (CPCB) in coordination with state-level control boards. Naturally, monitoring stations themselves do not have any direct impact on pollution. Rather, the CAAQMSs served the purpose of understanding—both spatially and temporally—the nature of pollution problems, and communicating this information directly to the public.

A brief history of air quality monitoring in India The CPCB has been systematically collecting pollution data under its national monitoring programme since 1985.[3] Initially, monitoring stations collected data on four key pollutants: sulfur dioxide (SO2), nitrogen dioxide (NO2), PM10 (particulate matter under 10 microns), and PM2.5 via *manual* monitoring stations scattered around the country. These manual stations involved the collection of ambient air over a period of several days, which

[3] Air quality monitoring began earlier, in 1978, in 10 cities – Ahmedabad, Mumbai, Kolkata, Kochi, Delhi, Hyderabad, Jaipur, Kanpur, Chennai, and Nagpur.

was then transported to a central location where the data were analyzed. The resultant (manually generated) report was then archived to the Environment Air Quality Data Entry System (EAQDES). It was then at the discretion of local of local pollution control authorities to upload and/or make this information available. The post-collection process itself could take up to a week, and the data that had been obtained was often limited and, in some cases, completely unavailable or unfit for processing (Pant et al. (2019)).[4] The CPCB itself released the data at a monthly or annual frequency. As the preceding description makes clear, the manually collected pollution data gathered under the earlier NAMP were slow to produce, of questionable quality, and hard for the general public to access (if the data were made available at all).

Both the quality and availability of data shifted markedly with the introduction of CAAQMSs, first piloted in 2006 and later expanded in 2016 (see below for more details on the rollout). The monitoring stations collect information on a wider range of pollutants than earlier instruments.[5] Furthermore, both the collection and analysis of data has been fully automated via "internet of things" technology that facilitates continuous automated data collection, as well as the transfer of pollution data to the central server in real-time every few minutes. At the center, data analysis is also automated via advanced AI, and is ready for use shortly after collection. Pollution data from the CAAQMSs are used to calculate an air quality index (AQI), a standardized metric that incorporates a wider range of pollutants than the earlier NAMP monitoring program. Real-time AQI readings are publicly available online via a smartphone app, with historical data archived on by the CPCB and downloadable on its website.[6] local AQI via For more details on the monitoring program as well as references for AQI measurement, see Pant et al. (2019).

A primary purpose of the continuous air quality monitoring program is creating public awareness of environmental conditions. This objective was greatly facilitated with the advent of CAAQMSs, as the earlier manually-driven system could not provide real-time environmental data to the public. This

[4]For a snapshot of the history of air pollution monitoring before and after CAAQMS look into: Air Pollution Monitoring in India

[5]Measured pollutants including PM2.5, PM10), nitrogen oxides, sulphuric dioxide, benzene, toluene, ethylbenzene, and xylene

[6]It is difficult to track exactly how many people keep track of local air quality via smartphone because, in addition to the CPCB's own app, there are dozens more that provide real-time AQI data for India and internationally. Many are listed as having 100,000+ downloads. The CPCB reports that its own app has been downloaded over 500,000 times.

information is now delivered via public displays, web widgets, alerts, and a proliferation of mobile apps that provide localized information on air quality.

Rollout and location of CAAQMSs Real-time pollution monitoring was first piloted in Delhi in 2006, with the expansion accelerating only in 2016 (see Figure 1). [7]

The decision of where to locate monitoring stations is done by the CPSB in consultation with the state-level pollution control boards. The criteria naturally include consideration of nearby pollution severity and population; however, there is a much longer list of practical concerns that include geographic obstructions, security, cost, and power availability, among many others. While we identify our main results from a difference-in-differences framework, it is still worth noting that within a relatively narrow geographical region, the locations of CAAQMSs (as well as their timing) are dictated in large part by considerations are largely exogenous to factors that might affect the portfolio allocation decisions of individual traders.

There is no clear threshold for how the relevance of AQI readings diminish with the distance from a monitoring station. In the most urban areas, pollution can be highly localized, and for this reason—and also because of population density—India's major metropolitan areas are covered by multiple stations. For example, Delhi has 41 monitoring stations to cover its 20 million residents spread across nearly 1500 square kilometers. There are 21 stations for Mumbai (population 21 million, area 603 sq km); and 7 stations for Kolkata (population 15 million, area 205 sq km).

Once one gets outside of a handful of major metropolitan areas, however, coverage is much sparser. For example, Jodhpur, a city of 1.6 million, has a single monitoring station, and there are no other stations within 60 kilometers of it. The entire state of Jammu and Kashmir has just a single monitoring station in its largest city of Srinagar (population 1.7 million).

As we describe below, we will take traders within a 20 kilometer radius of a CAAQMS to be "treated" with AQI information when the monitoring station opens,This range is based on conversations with CPCB officials. and take investors in a 40-60 kilometer "donut" as a benchmark or control group. It is important to note that our results do not hinge on this particular treatment-control definition—

[7] See, e.g., Gulia et al. (2022) on the timing of CAAQMSs. For the current map of stations, see https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing, last accessed December 16, 2022.

we consider also a narrower radius of 15 or 10 kilometers in robustness checks. Also, given the preponderance of overlapping monitoring stations in India's very largest cities, we also show that the results also hold when we exclude the small subset of traders located in Mumbai, Delhi, and Kolkata.

3 Data

We use several databases in our analysis that allow us to connect pollution, monitoring, and trading behavioral at a granualr level. The databases include individual-level stock holdings over time for Indian investors as well as basic geographic and demographic information; geocoded information on the timing of the installation of CAAQ monitoring stations, and information on local pollution inferred from satellite images.

Stock Trading Data: We take advantage of comprehensive data on stock trading from the National Stock Exchange (NSE) of India, one of the largest stock exchanges in the world. We obtained the data on the universe of trading records from the NSE for the period of 2004-2020. The data allow us to observe a number of features for each transaction – the anonymized Permanent Account Number (PAN) of the trader, the transaction date, the security ticker, total shares purchased or sold, and the execution price. We limit our analysis to transactions involving stocks included in the Prowess Database (similar to CRSP in the U.S.) maintained by the Centre for Monitoring Indian Economy (CMIE).[8] In addition, we retain only securities that are common shares of domestic stocks and exclude trading activities related to ETFs and foreign stocks.[9] As noted above, for each retail investor, we further obtain their geographical location at the six-digit PIN Code level, which is essential to match our trading data to information on the location and opening of CAAQMSs. Overall, the initial sample includes equity transactions for 19 million unique investors across India.

We provide the geographic distribution of domestic retail investors in Figure A2. Unsurprisingly, there is a particularly high concentration of investors in Delhi, Kolkata, and Mumbai. These are India's three most populous cities, and also major financial centers.[10] We aggregate investors' trading activities

[8]Prowess is the standard database employed by researchers in work on Indian equity markets. See, e.g.,(Khanna and Palepu, 2000; Goldberg et al., 2010; Balasubramaniam et al., 2022; Bau and Matray, 2022)

[9]None of the ETFs in our sample have an explicit ETF orientation.

[10]In Appendix Table A5, we show that our baseline results are virtually unchanged when we exclude investors from these three areas.

of green stocks based on their location to match the variation in CAAQMS introduction and conduct our main analysis at the PIN code level. Intuitively, each regional account describes the trading activities of a representative regional investor who buys and sells shares of a stock.

Identifying Brown Stocks: We classify firms based on whether it is in a "*"brown"* or *"green"* industry, following Choi, Gao and Jiang (2020), which in turn uses the definition of the Intergovernmental Panel on Climate Change (IPCC) to classify five sectors as sources of high emissions – Energy; Transport; Buildings; Industry (such as chemicals and metals); and Agriculture, Forestry, and Other Land Use. Then, following Krey et al. (2014), each sector is classified into sub-categories, each of which is hand matched with industry names provided by Datastream.

We use the industry classifications provided in Prowess database to hand match to Datastream industries, in order to classify it as green or brown based on the Choi, Gao and Jiang (2020) list. Figure 3 plots the evolution of brown and green (i.e., not brown) stocks over our sample period. In panels (a) and (b) we show the number of stocks and market capitalization, and in panel (c) we show brown stocks' share of market capitalization.Brown stocks' market share varied between about 30 and 40 percent between 2004 and 2020, peaking in 2010 then dropping steadily for the next decade.

Pollution Monitoring Stations – CAAQMSs: We obtained geo-coded data on the location and timing of pollution monitoring stations from India's Central Pollution Control Board. The first CAAQMS appeared in Delhi in 2006, and by 2021 there were 311 stations spread across the country (see Figure 1). We classify a PIN Code as in the "treatment" group if its centroid is located within a 20 kilometer radius of a CAAQMS under the assumption that, when a monitoring station appears, investors in the PIN Code gain ready access to real-time air quality data. We leave a buffer of a "donut hole" region comprised of PIN codes 20-40 kilometers around the station, and then classify PIN Codes that have centroids located 40-60 kilometers from each station as the "control" group. We classify a PIN code as control only when its centroid is not within the treated region of any other CAAQMS. Thus, the treatment and control groups are mutually exclusive and the sample is created such that there is no overlap across the groups, with every CAAQMS having set of treated and a control PIN Codes. Figure A1 illustrates our assignment approach for a monitoring station in Jodhpur, which provides a relative straightforward case, given that the city hosts only a single station. India's largest cities have multiple stations that are generally within 60 kilometers of each other, which creates some complications in treatment and control assignments, which we address as described above. However, as also noted earlier, we sidestep these complications in a robustness test in which we exclude Mumbai, Delhi, and Kolkata from our analysis.

Satellite Data on Pollution: We are able to measure local air pollution at the PIN-Code level in a consistent manner both before and after the introduction of CAAQMSs. Specifically, we take advantage of data generated by NASA's Moderate Resolution Imaging Spectroradiometer on its Terra satellite. These readings are used by NASA to generate aerosol optical depth (AOD) data that is a widely-used proxy for particulate matter concentration (see, e.g., Van Donkelaar, Martin and Park, 2006). Past work has shown a very high correlation between AOD measurements and actual particulate matter concentrations (e.g., Tsai et al., 2011), particularly when aggregated to the monthly level (Kumar, Chu and Foster, 2008). The AOD data are available at a frequency of 30 minutes in a 10-by-10 kilometer grid. Our main pollution data are average monthly readings for the grid location that contains the PIN Code's centroid

One natural concern is the endogenous timing of CAAQMS installations, i.e., monitoring stations may have been placed in areas where pollution is worsening. In Figure 4 we provide an event plot of pollution (as captured by AOD) around installation dates, and do not observe any significant shift. (We also report the difference-in-difference regression results in Appendix Table A1, which similarly show no relationship between the appearance of monitoring stations and the level of air pollution.

Sample Description and Summary Statistics: Our regression analysis a imarginally at the PIcode-month level. Our main sample consists (41.2 versus 41.9%)f 1,859 distinct "treated" PIN codes and 5,254 "control (40.4%)" PIN codes, over the period January 2004 to June 2020. The panel is unbalanced – in 65 percent of PIN-code-month observations, stock holdings are zero so that our main outcome, share of brown stocks, is undefined. Some monitoring roll-out occurs during 2017-2019, close to our sample period end. To ensure that the underlying sample for each station is a relatively balanced panel, we have additional filtering and focus on the +/- 4 years time window around the rollout time. The total number of observations for the baseline specification at the PIN-code-month level is 499,036.

Table 1 presents summary statistics for the primary variables we use in our analysis. Aggregated across our entire sample period, we observe that a typical retail investor holds roughly 41.8% of their portfolio in brown stocks, marginally above the overall share of brown stocks in the Indian market (see Figure 3). Of relevance for the heterogeneity analyses we present in Section 5.2, Table 1 also provides portfolio brown shares across investor groups. We observe that female investors hold a marginally lower fraction of assets in brown stocks relative to their male counterparts (41.2 versus 41.9%), while the share of brown stocks is notably lower for young investors (40.4%), relative to mid-aged and old investors (41.9 and 42.8% respectively).

4 Empirical Strategy

In our empirical analyses, we will examine the link between pollution exposure and and investment in 'brown' investments, before versus after pollution data become readily available via a CAAQ monitoring station. As delineated above, we will further take advantage of fine-grained information on investors' geographies to add a third layer of differences, based on whether an individual is within 20 kilometers of a station, or sufficiently distant thus estimate the sensitivity of brown investment to pollution exposure as a function of availability of real-time AQI data (PIN codes 40-60 kilometers from the monitoring station).

The empirical design of this study follows Barwick et al. (2019) and exploits the staggered set up of the pollution monitoring stations. The regression specification thus takes the following form:

$$
\begin{aligned}
Brown\,Share_{p(m)t} = \beta \cdot Treat_p \cdot AQI_{pt} \cdot Post_{mt} + \sum_k \beta_k \cdot Lower\ Order\ Interactions \\
+X_{mt} + \alpha_p + \alpha_{mt} + \varepsilon_{pt} \qquad (1)
\end{aligned}
$$

$Brown\,Share_{p(m)t}$ denotes the average share of brown stocks of a retail investor in PIN code p in monitoring station area m at time t. A station area m contains PIN codes within 20 kilometers of the monitoring station – treatment PIN codes – and 40-60 kilometers from the station – control PIN codes. $Treat_p$ is an indicator variable that denotes whether a PIN code is "treated" by a nearby (within 20 kilometers) monitoring station; AQI_{pt} is the Air Quality Index in PIN code p in time t. We remind the reader that AQI is measured for both treated and control PIN codes using satellite data, based on information that was not readily available to Indian citizens. The treatment variable captures exposure to this information in real-time as a result of monitoring stations. Lower Order Interactions includes the interactions: $Treat_p \cdot AQI_{pt}$, $Treat_z \cdot Post_{mt}$, $Post_{mt} \cdot AQI_{pt}$, and $Treat_p$, AQI_{pt} and $Post_{pt}$ dummies. α_p is a set of PIN code fixed effects that control for time invariant factors of a PIN code, so that identification is from time-series variation in whether a nearby monitoring station has opened. In some specifications we additionally include α_{pt}, a set of the station area $\times$ time fixed effects, which collectively control for time varying characteristics in a region around monitoring stations, so that identifying variation comes from differences in treatment and control groups at the same point in time. The coefficient of interest β measures the changes in the relationship between pollution exposure and investment in brown stocks with the appearance of a monitoring station, for nearby (treated) versus more distant (control) PIN codes.

We also implement a dynamic version of regression specification 1 explore dynamics around the treatment date, in particular to check whether there existed any differences between the control and treatment prior to the appearance of CAAQMSs.

$$Brown\,Share_{p(m)t} = \sum_{k=-6,k\neq 0}^{6} \beta_k \cdot Treat_p \cdot AQI_{pt} \cdot D(t=k)_{mt} + X_{pt} + \alpha_p + \alpha_{mt} + \varepsilon_{pt} \tag{2}$$

$D(t=k)_{mt}$ is a dummy variable that takes a value 1 if the month t belongs to the k^{th} quarter since the monitoring station appeared.

Our empirical design allows us to alleviate several key threats to identification. First, while monitoring stations were placed following several pre-defined guidelines (see Section 2 for details), one might still be concerned about endogeneity of site selection. In the specification 1, we compare geographically proximate control and treatment PIN codes around the same monitoring station, and can further observe in our dynamic specification whether there is any differential pre-trend prior to the appearance of monitoring stations. We will also show that there is no difference in observed air quality between the treated and control groups around the monitoring station "treatment" date. We also show in appendix table A2that there is no change in some other confounding factor like economic activity (measured as entry of new firms and night time light images) that could differently impact trading in the treated and control areas post the introduction of CAAQMS.

Further, we are interested in the difference in the sensitivity of brown investment to pollution before versus after the appearance of a monitoring station, rather than the direct effect of the monitoring stations itself. Given that, as noted immediately above, we find that pollution does not change around the introduction of monitoring stations, we can credibly attribute the change in brown-invesment-to-AQI sensitivity to the salience of pollution information (Barwick et al. (2019)).

Finally, our setting allows us to pair each treated set of investors to a control group, rather than relying solely on the staggered roll-out of monitoring stations. Our specification is thus less vulnerable to potential biases in difference-in-differences estimation, as highlighted by recent methodological findings (see Baker, Larcker and Wang, 2022; Sun and Abraham, 2021, among others).

5 Results

We begin by presenting our results graphically. In Figure 6 we depict the point estimates and 95% confidence intervals for the coefficients generated by a variant of specification (2) above, in which we split the sample into investors situated in treated and control PIN codes—those less than 20 kilometers from a monitoring station and those 40-60 kilometers from the same station, respectively.

From in identification perspective, it is comforting that our measure of brown share investments is essentially flat in the two years (eight quarters) preceding the installation of a monitoring station. At that point there is a clearly discernible shift in the sensitivity of brown investment share to AQI in treated PIN codes—higher AQI (i.e., worse pollution) is associated with a lower brown share after a monitoring station arrives. We observe no such shift in control PIN codes.

To assess whether the result is being driven by a particular part of the pollution distribution, we provide a binned scatterplot showing the relationship between observed AQI and brown share holdings (adjusted for the time-varying mean in a station area across both control and treatment PIN codes), before versus after the arrival of a monitoring station. We present these results in Figure 5. Consistent with the patterns in the preceding figure, we observe no correlation between measured AQI and brown share investment in either group in the pre-monitoring-station period. In the post-monitoring period, there is a clear negative association between AQI and brown investment share in "treated" geographies, but not control ones.

We show our main result, based on 1, in Table 2, with a variety of fixed effects. As explained in the preceding section, our primary interest is in the coefficient on the three-way interaction of AQI $\times$ Treated $\times$ Post, which reflects the differential response of investors to AQI after a monitoring station appears in nearby versus distant PIN codes. The point estimate of -2.4 in our favored specification, which includes controls for investor and weather characteristics as well as Station $\times$ Year-Month fixed effects, may be interpreted as follows: after a monitoring station appears, an AQI increase of 0.27 (the within-station interquartile range) is associated with a decline of the brown share in investors' portfolios in affected areas of 0.65 percentage points (1.6 percent relative to the mean brown share of 41 percent). If we set $Treated = 0$, there is relationship between AQI and brown share either before or after the introduction of a monitoring station, consistent with the patterns shown in Figure 6. And when we evaluate the effect of Treated $\times$ Post at "low" levels of AQI (e.g., 0.31, the 10th percentile), we observe a statistically insignificant coefficient in our favoured specification using full set of fixed effects and controls (table A3) implying that there is no observable effect of the introduction of CAAQMS itself on the holding of brown shares.

We present a dynamic version of this analysis in Figure 7. The pattern shows that the Treated $\times$ AQI coefficient is very closely zero for all pre-periods, and then falls to a very consistent value of between -2 and -3 in the post-periods.

In Appendix Tables A4 - A5, we present a pair of robustness checks for the main results. In Appendix Tables A4 we restrict the "treated' group to those within 15 or 10 kilometers of a monitoring station; the point estimates in our preferred within-station specification are virtually unchanged. Appendix Table A5 excludes PIN codes in the largest metropolitan areas—as explained in Section 2, these PIN codes offer a less straightforward delineation of treatment and control assignments. We observe substantially larger point estimates on the three-way interaction.

The question naturally arises of whether the portfolio adjustments are made as a result of beliefs about the returns of green versus brown stocks, or a distaste for holding brown stocks. While it is beyond the scope of our paper to evaluate whether tastes or beliefs drive investors' portfolio changes, we can at least evaluate whether, during the period that monitoring stations were opening, brown stocks under-performed in India or elsewhere (in case the trading behavior we document itself put downward pressure on prices). We show in Appendix Figure A4 the cumulative returns of brown and green portfolios over the period January 2000 to December 2019 (see Section 3 for details on the portfolios' construction). Over any relevant horizon, the two portfolios perform quite similarly, though with the brown portfolio actually generating higher returns. This pattern suggests that, to the extent that beliefs rather than tastes explains our main results, it stems from erroneous expectations of negative brown stock returns.

5.1 Attention to the level of pollution versus sensitivity to changes

In the preceding analysis, we explore how the *sensitivity* of brown-share investment to air quality changes with the arrival of a monitoring station. This formulation evaluates the extent to which investors, who face a given level of pollution, are discouraged from holding brown stocks when air quality is particularly poor. A related but distinct view of pollution salience is that investors who are generally confronted by information about poor air quality as a result of a monitoring station will permanently shift their portfolios away from brown stocks. In this formulation, we replace the time-varying measure of AQI in specification (1) with a time-variant PIN code level measure of average pre-treatment pollution, AQI_p.

These results appear in Table 3. The coefficients on the three-way interaction term are somewhat larger than those in Table 2, which contains our main results. However, the relevant variation is also narrower—the between PIN-code interquartile range in average AQI is 0.24, as compared to the overall variance in month-to-month AQI that we study in our main results, which as noted above has an interquartile range of 0.27.

5.2 Heterogeneity by investor type

We provide some exploratory analyses based on several attributes of traders that potentially relate to propensity to access to AQI information and/or concern for the environment, including age, gender, and whether investors primarily execute trades via mobile, internet, or more traditional means.

We naturally do not have random assignment or any close approximation to it for these attributes, and as such the results should be interpreted with caution. Still, the patterns are interesting to consider as an extension to our main results given that, based on intuition as well as past research, one may have prior expectations about which groups may have greater sensitivity to pollution salience.

We begin by comparing sensitivity to AQI based on whether an investor predominantly makes trades via mobile, internet or broker or other "traditional" method. As observed in the introduction, given that the public predominantly obtained AQI updates via smartphone, it is natural to speculate that more tech savvy investors—who use their mobile devices to execute trades—would be more exposed to pollution updates.

In Figure 8, panel A, we show event plots for responsiveness to AQI information disaggregated by investors' primary means of trading. As may readily be seen in the graph, the largest effect is observed among investors who use mobile devices to trade. The confidence intervals of the other two groups—internet and physical trading—are largely overlapping, though the effect of monitoring stations is marginally greater for internet-based traders. the corresponding tables are in Appendix Tables A6

We next turn to explore heterogeneity by age. Before turning to these results, it is worth noting that, while younger investors are more frequently classified as "mobile" traders, the differences are surprisingly modest relative to middle-aged investors: 13 percent of younger investors are classified as mobile, as compared to 10 percent of middle-aged investors (elderly investors are indeed far less likely to use mobile devices for trades—only 3.7 percent of the elderly in our sample are classified as mobile).

More importantly, putting aside technological concerns, there are reasons to expect differential responsiveness to pollution information by age. First, a vast literature documents a very strong negative correlation between environmental concerns and age—unsurprising, given that the young will disproportionately bear the consequences of climate change and environmental degradation.[11] Of more direct relevance, recent survey-based evidence finds higher stated interest in ESG amongst younger investors and lower interest amongst older investors (Haber et al., 2022). An alternative hypothesis is that the elderly might be *more* sensitive to information on air pollution, because they are far more vulnerable to the effects of air pollution (see Gouveia and Fletcher, 2000; Fischer et al., 2003).

In Figure 8, panel B, we revisit our main event plot, disaggregating the sample into young (18-29), middle-aged (30-55), and elderly (above 56) investors; the corresponding tables are in Appendix Tables A7. We observe a substantially greater shift in brown-investment-to-AQI sensitivity amongst the young, relative to the other two age groups. While we cannot put too strong an interpretation around this finding, we see this particular result as reinforcing the above-cited evidence on the age distribution of interest in ESG investing, which may be of direct practical relevance.

Finally, we split the sample by gender. As noted in the introduction, there is an ex ante rationale for a differential response given past work on a "gender environmentalism gap" (e.g., Xiao and McCright, 2015). We illustrate the differential response by gender in the event plots Figure 8 panel C. The gender difference is striking—women exhibit a responsiveness that is 2-3 times greater than that of men. Observe that the male point estimates are more precisely estimated, and also much closer to the full-sample estimates, which reflects the fact that most Indian retail investors are men. For completeness, we also provide the tabular version of the gender split in Appendix Table A8, which provides the same message as the event plots.

[11] See, e.g., Liere and Dunlap (1980), for an early and well-cited review which describes age as the "predominant" individual attribute that is correlated with environmental concerns; a more recent review article by Sanchez-Sabate and Sabaté (2019) similarly finds an important role for age, focused specifically on environmental concerns and meat consumption.

Our final "heterogeneity" test explores whether we see a comparable shift in responsiveness for institutional investors as observed in our main sample of retail investors. We present these results in A3. If anything, institutional investors' response to the availability of AQI data moves in the opposite direction to that of retail investors, though we cannot reject that institutions are simply unaffected by the appearance of real-time pollution data. There are various potential explanations for this non-result. The most natural is that institutional investors may be less sensitive to "taste-based" shifts in investing. However, we also may simply have a less-precise mapping of AQI to relevant location, since the PIN code for an institutional investor reflects their place of work rather than residence. More broadly, we interpret this non-result with caution, relative to our main findings on retail investors.

6 Conclusion

We document that exposing investors to ready information about air pollution heightens the sensitivity of their "brown" investments to air quality. We interpret these findings through the lens of salience, in the spirit of Bordalo, Gennaioli and Shleifer (2013) among others—ready access to air quality data makes this information a more salient input into investment decisions.

As we noted in our discussion of the results, this shift comes despite the fact that returns for a long-short Green-Brown portfolio does not generate any excess returns—if investors adjust their portfolios in the expectation of higher returns, the shift is not justified by realized outcomes. That said, we cannot identify whether the shift we document is driven by mistaken beliefs, or shifts in investor tastes as a result of greater attentiveness to environmental problems. Distinguishing between these two explanations is one important future direction, and one that we plan to pursue going forward.

References

Alok, Shashwat, Nitin Kumar, and Russ Wermers. 2020. “Do fund managers misestimate climatic disaster risk.” *The Review of Financial Studies*, 33(3): 1146–1183.

Baker, Andrew C, David F Larcker, and Charles CY Wang. 2022. “How much should we trust staggered difference-in-differences estimates?” *Journal of Financial Economics*, 144(2): 370–395.

Balasubramaniam, Vimal, John Y Campbell, Tarun Ramadorai, and Benjamin Ranish. 2022. “Who Owns What? A Factor Model for Direct Stock Holding.”

Baldauf, Markus, Lorenzo Garlappi, and Constantine Yannelis. 2020. “Does climate change affect real estate prices? Only if you believe in it.” *The Review of Financial Studies*, 33(3): 1256–1295.

Bansal, Ravi, Dana Kiku, and Marcelo Ochoa. 2016. “Price of long-run temperature shifts in capital markets.” National Bureau of Economic Research.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2019. “From fog to smog: The value of pollution information.” National Bureau of Economic Research.

Bau, Natalie, and Adrien Matray. 2022. “Misallocation and capital market integration: Evidence from India.”

Berg, Florian, Julian F Koelbel, Anna Pavlova, and Roberto Rigobon. 2022. “ESG confusion and stock returns: Tackling the problem of noise.” National Bureau of Economic Research.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. “Salience and asset prices.” *American Economic Review*, 103(3): 623–28.

Choi, Darwin, Zhenyu Gao, and Wenxi Jiang. 2020. “Attention to global warming.” *The Review of Financial Studies*, 33(3): 1112–1145.

Cosemans, Mathijs, and Rik Frehen. 2021. “Salience theory and stock prices: Empirical evidence.” *Journal of Financial Economics*, 140(2): 460–483.

Fischer, P, G Hoek, B Brunekreef, A Verhoeff, and J Van Wijnen. 2003. “Air pollution and mortality in The Netherlands: are the elderly more at risk?” *European respiratory journal*, 21(40 suppl): 34s–38s.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova. 2010. “Imported intermediate inputs and domestic product growth: Evidence from India.” *The Quarterly journal of economics*, 125(4): 1727–1767.

Görgen, Maximilian, Andrea Jacob, Martin Nerlinger, Ryan Riordan, Martin Rohleder, and Marco Wilkens. 2020. “Carbon risk.” *Available at SSRN 2930897*.

Gouveia, Nelson, and Tony Fletcher. 2000. “Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status.” *Journal of Epidemiology & Community Health*, 54(10): 750–755.

Gulia, Sunil, Nidhi Shukla, Lavanya Padhi, Parthaa Bosu, SK Goyal, and Rakesh Kumar. 2022. "Evolution of air pollution management policies and related research in India."

Haber, Stephen, John D Kepler, David F Larcker, Amit Seru, and Brian Tayan. 2022. "ESG Investing: What Shareholders Do Fund Managers Represent?" *Rock Center for Corporate Governance at Stanford University Working Paper*.

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.

Hong, Harrison, Frank Weikai Li, and Jiangmin Xu. 2019. "Climate risks and market efficiency." *Journal of econometrics*, 208(1): 265–281.

Hong, Harrison, G Andrew Karolyi, and José A Scheinkman. 2020. "Climate finance." *The Review of Financial Studies*, 33(3): 1011–1023.

Huberman, Gur, and Tomer Regev. 2001. "Contagious speculation and a cure for cancer: A nonevent that made stock prices soar." *The Journal of Finance*, 56(1): 387–396.

Jiang, Han, Le Lexi Kang, Ziye Zoe Nie, and Hui Zhou. 2022. "Can Old Sin Make New Shame? Stock Market Reactions to the Release of Movies Re-Exposing Past Corporate Scandals." *Stock Market Reactions to the Release of Movies Re-Exposing Past Corporate Scandals (May 1, 2022)*.

Khanna, Tarun, and Krishna Palepu. 2000. "Is group affiliation profitable in emerging markets? An analysis of diversified Indian business groups." *The journal of finance*, 55(2): 867–891.

Krey, Volker, O Masera, G Blanford, T Bruckner, R Cooke, K Fisher-Vanden, H Haberl, E Hertwich, E Kriegler, D Mueller, et al. 2014. "Annex 2-metrics and methodology."

Kumar, Naresh, Allen Chu, and Andrew Foster. 2008. "Remote sensing of ambient particles in Delhi and its environs: estimation and validation." *International Journal of Remote Sensing*, 29(12): 3383–3405.

Liere, Kent D Van, and Riley E Dunlap. 1980. "The social bases of environmental concern: A review of hypotheses, explanations and empirical evidence." *Public opinion quarterly*, 44(2): 181–197.

Li, Ye, Eric J Johnson, and Lisa Zaval. 2011. "Local warming: Daily temperature change influences belief in global warming." *Psychological science*, 22(4): 454–459.

Lujala, Päivi, Haakon Lein, and Jan Ketil Rød. 2015. "Climate change, natural hazards, and risk perception: the role of proximity and personal experience." *Local Environment*, 20(4): 489–509.

Murfin, Justin, and Matthew Spiegel. 2020. "Is the risk of sea level rise capitalized in residential real estate?" *The Review of Financial Studies*, 33(3): 1217–1255.

Pant, Pallavi, Raj M Lal, Sarath K Guttikunda, Armistead G Russell, Ajay S Nagpure, Anu Ramaswami, and Richard E Peltier. 2019. "Monitoring particulate matter in India: recent trends and future outlook." *Air Quality, Atmosphere & Health*, 12(1): 45–58.

Pelizzon, Loriana, Aleksandra Rzeznik, and Kathleen Weiss Hanley. 2021. "The salience of ESG ratings for stock pricing: Evidence from (potentially) confused investors."

Sanchez-Sabate, Ruben, and Joan Sabaté. 2019. "Consumer attitudes towards environmental concerns of meat consumption: A systematic review." *International journal of environmental research and public health*, 16(7): 1220.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

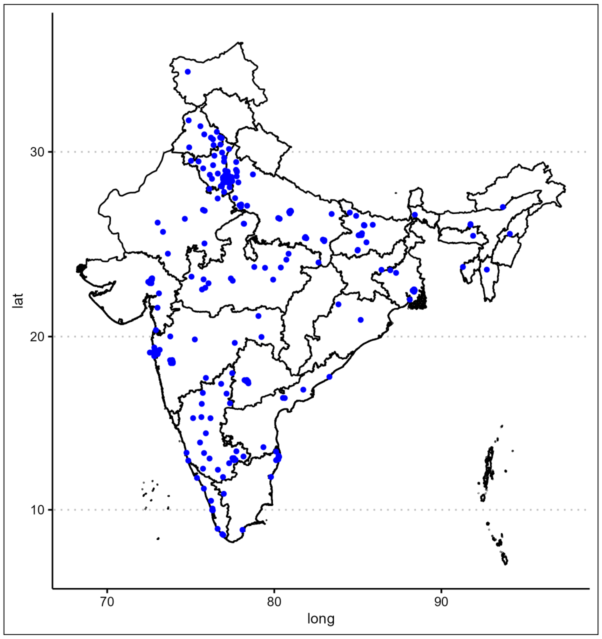
Tsai, Tzu-Chin, Yung-Jyh Jeng, D Allen Chu, Jen-Ping Chen, and Shuenn-Chin Chang. 2011. "Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008." *Atmospheric Environment*, 45(27): 4777–4788.

Van Donkelaar, Aaron, Randall V Martin, and Rokjin J Park. 2006. "Estimating ground-level PM2. 5 using aerosol optical depth determined from satellite remote sensing." *Journal of Geophysical Research: Atmospheres*, 111(D21).

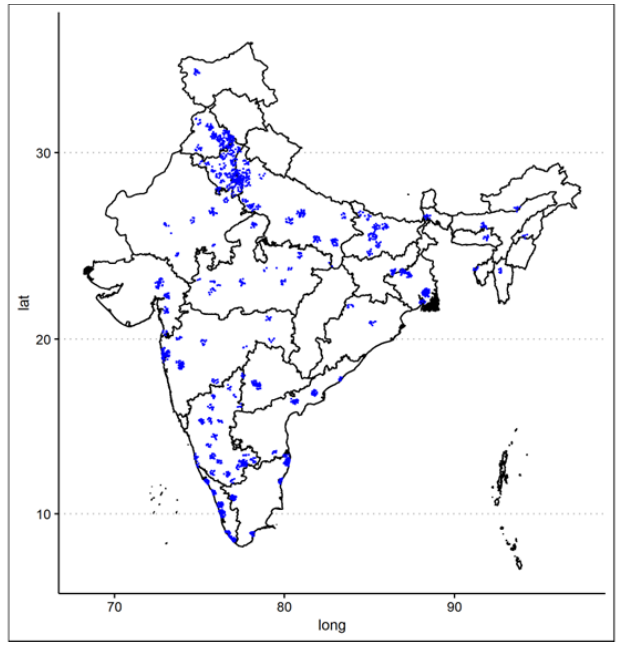
Xiao, Chenyang, and Aaron M McCright. 2015. "Gender differences in environmental concern: Revisiting the institutional trust hypothesis in the USA." *Environment and Behavior*, 47(1): 17–37.

Zaval, Lisa, Elizabeth A Keenan, Eric J Johnson, and Elke U Weber. 2014. "How warm days increase belief in global warming." *Nature Climate Change*, 4(2): 143–147.

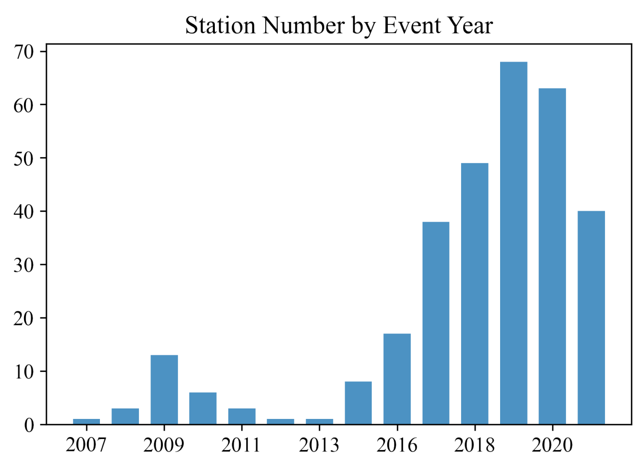
Figure 1: Continuous Ambient Air Quality (CAAQ) Monitoring Station



(a) Geographic Distribution of CAAQ Monitoring Station



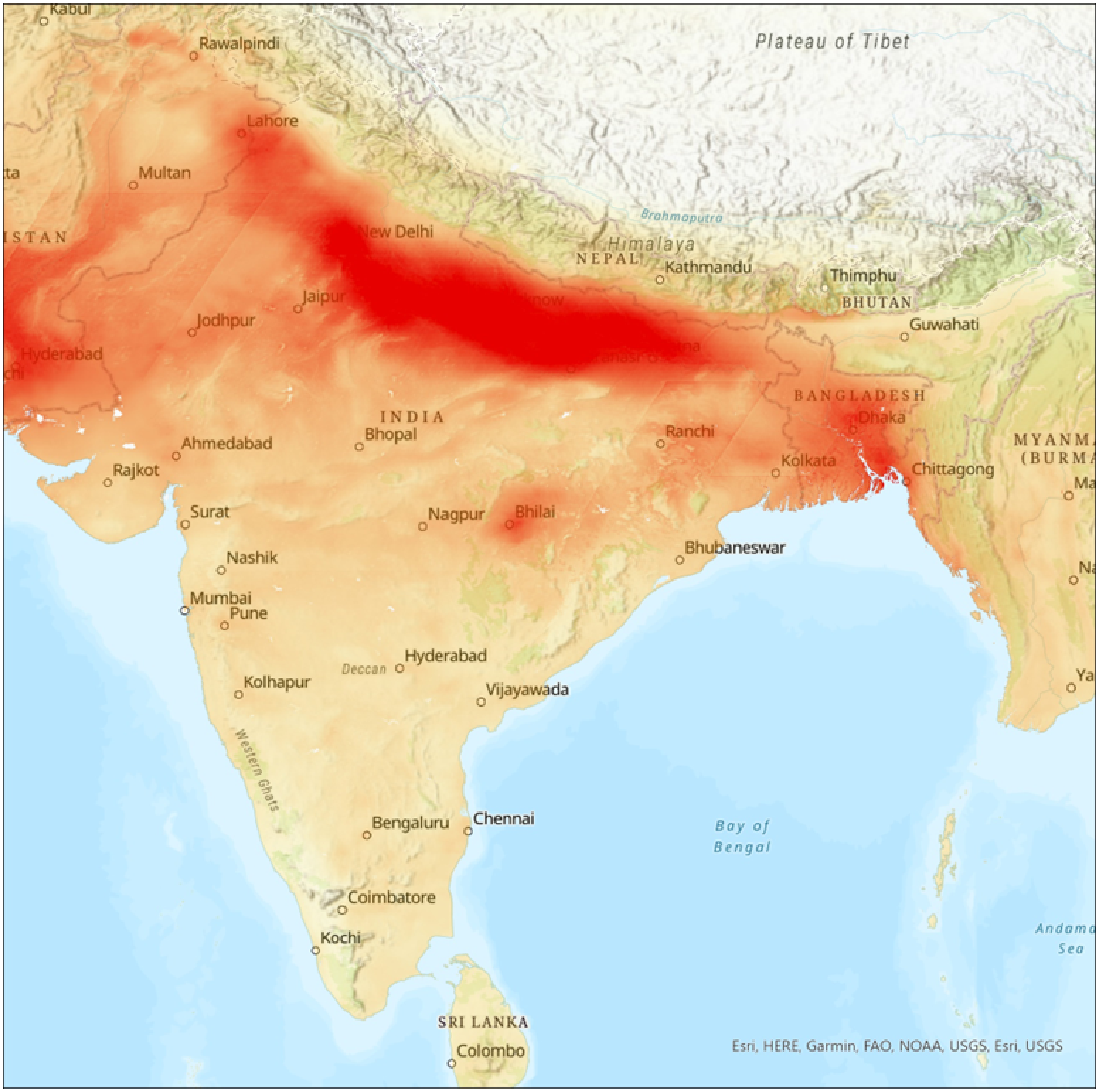
(b) Zip Code covered by CAAQ Monitoring Station



(c) Evolution of CAAQ Monitoring Station

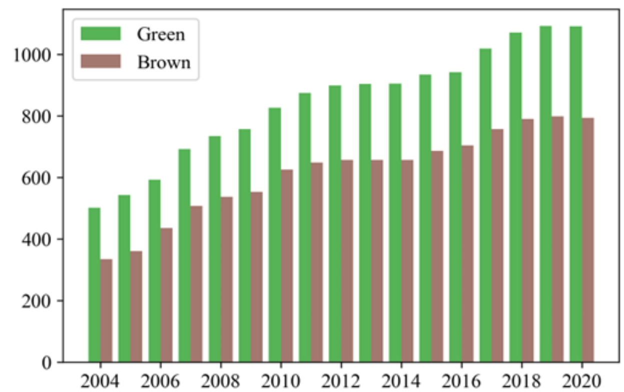
This figure plots the distribution of CAAQ monitoring stations and the zip codes that are covered. Panel A plots the geographic location of the CAAQ monitoring stations across the country. Panel B plots the zip codes that are covered by the CAAQ monitoring stations. Panel C plots the yearly evolution of the CAAQ monitoring stations over the years.

Figure 2: Snapshot of Satellite-retrieved Pollution

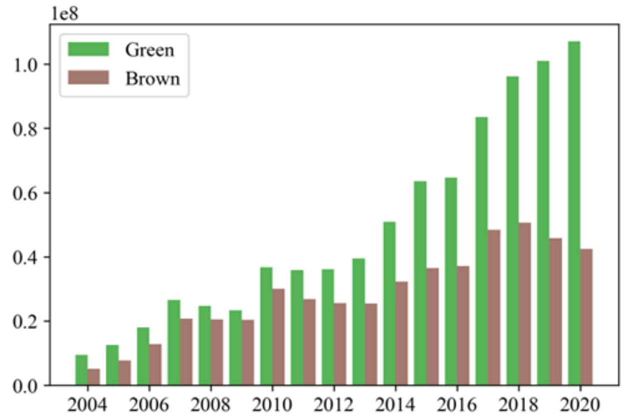


This figure plots the snapshot of pollution that is retrieved from the satellite data. The dark red shows high pollution while lighter colour shows lower level of pollution.

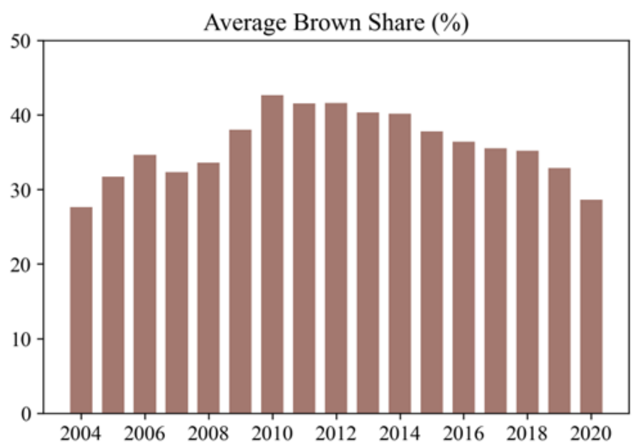
Figure 3: Distribution of Green and Brown Stock



(a) Brown vs Green Stock(stock number)



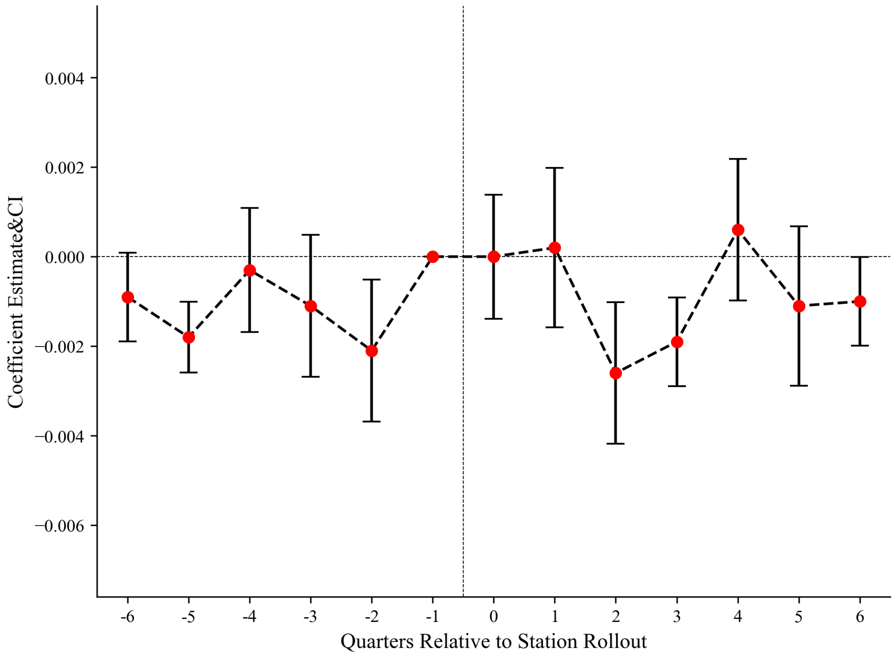
(b) Brown vs Green Stock(market capitalization)



(c) Evolution of the Share of Brown Stocks

This figure plots the distribution of green and brown stock holding over the years. Panel A presents the number of green and brown stocks held by the retail investors over the years. Panel B presents the market capitalization of green vs brown stocks. Panel C presents the market share of brown stocks (in terms of market capitalization) among all publicly traded equities in India.

Figure 4: Trend in Pollution Around (CAAQ) Installation

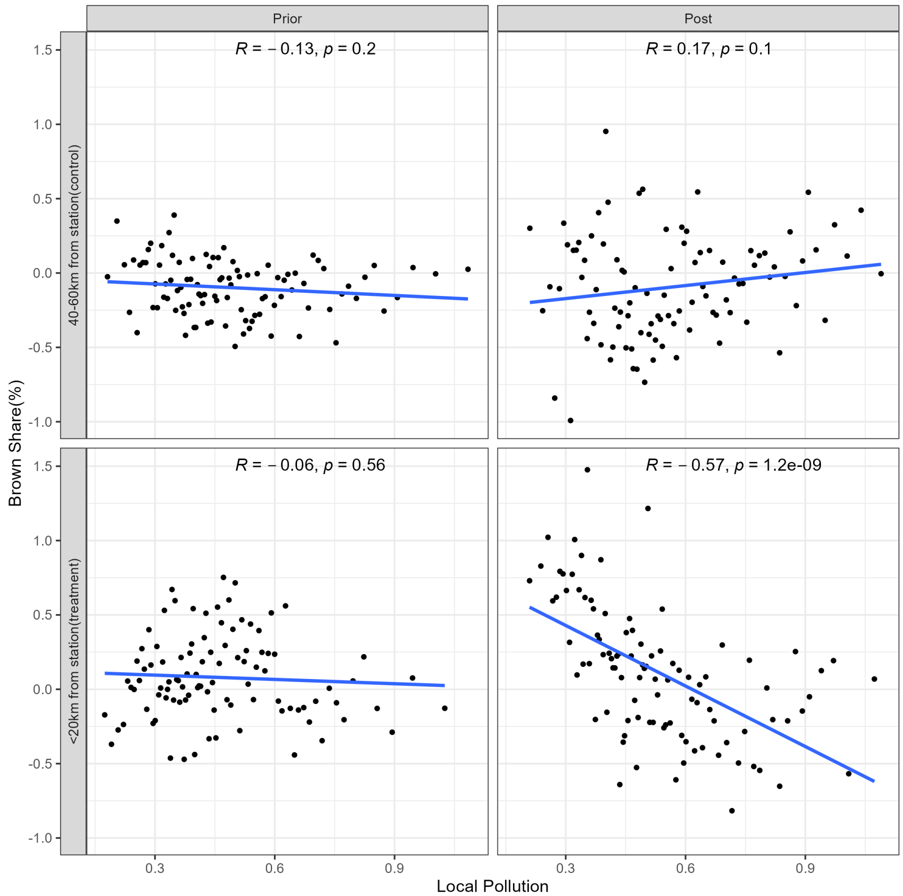


This figure plots trend in pollution following the installation of of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we plot the coefficient $\{\beta_k\}$ from the following specification:

$$Pollution_{z(p)t} = \sum_{k=-6,k\neq-1}^{6} \beta_k \cdot Treat_z \cdot 1(t = k) + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

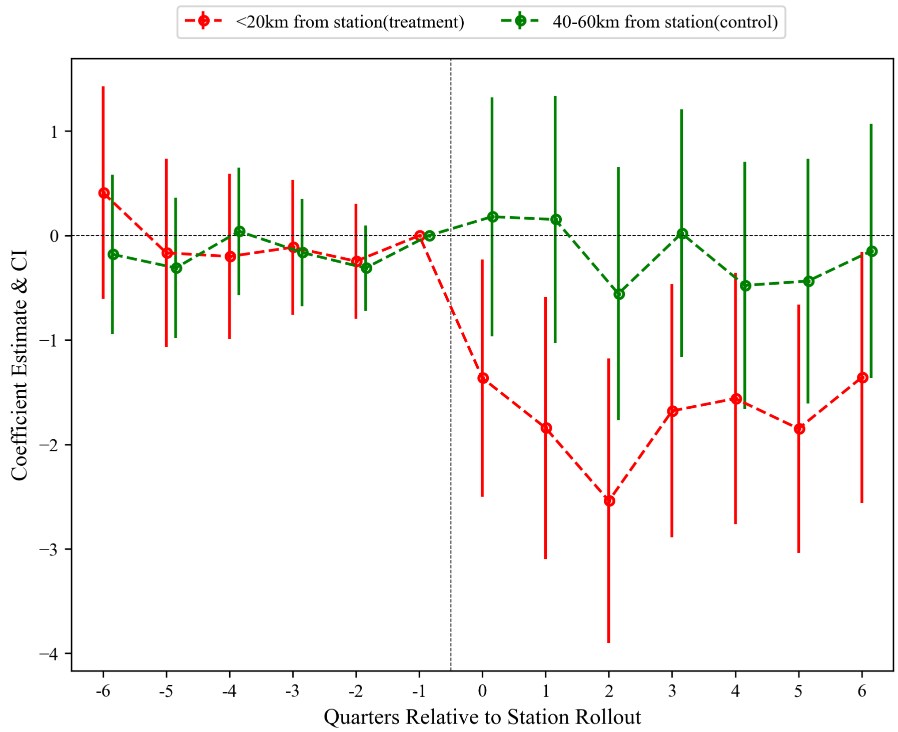
where $Pollution_{z(p)t}$ denotes the average pollution in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Capped spikes drawn with the estimated coefficients $\{\beta_k\}$ indicate 95% confidence intervals obtained from standard errors clustered at the zip-code level.

Figure 5: Correlation of Brown Stock share and Pollution - Control vs Treated



This figure presents the scatter plot of the share of brown stocks against local pollution separately for control and treated group before and after the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. We sort pollution in 100 buckets and plot the average pollution and the average brown share holding. The brown share is adjusted for the time varying mean holding in a station area (control and treatment). The first row are the plots for the control zip codes, i.e. 40km to 60km around the station and the second row are the plots for the control zip codes, i.e zip codes belonging within 20kms of the station. The first column is prior to the CAAQ installation, the second column is after CAAQ installation.

Figure 6: Sensitivity of Brown Share to Pollution - Control vs Treated

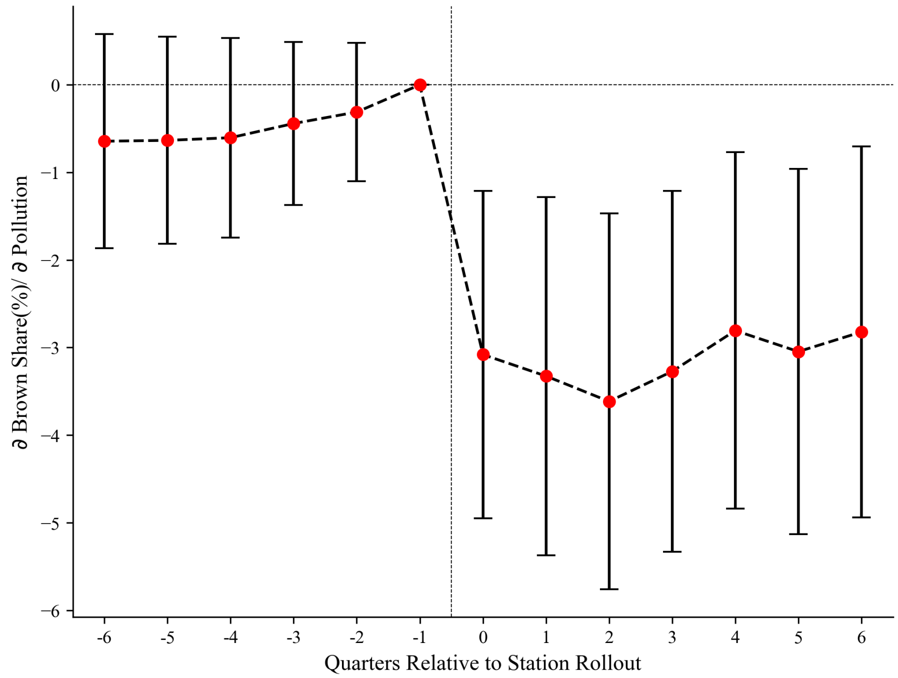


This figure plots the sensitivity of investing share in brown stocks to pollution before and after the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations separately for the control and treatment groups. Specifically we plot the coefficient $\{\beta_k\}$ from the following specification separately for the control and treatment groups:

$$Brown\,Share_{z(p)t} = \sum_{k=-6,k\neq-1}^{6} \beta_k \cdot Pollution_{zt} \cdot 1(t = k) + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Capped spikes drawn with the estimated coefficients $\{\beta_k\}$ indicate 95% confidence intervals obtained from standard errors clustered at the zip-code level.

Figure 7: Sensitivity of Brown Share to Pollution

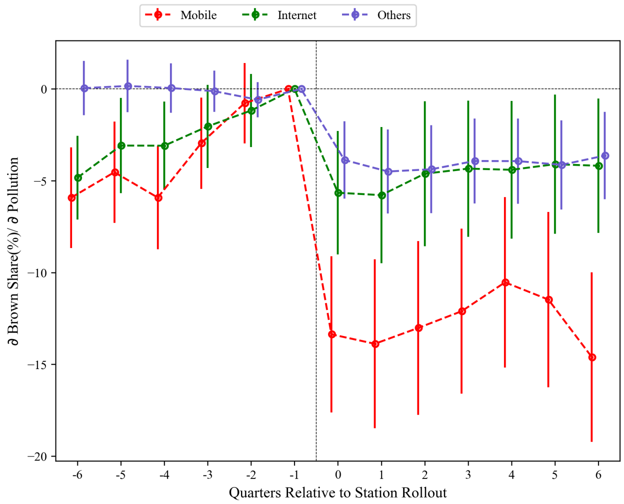


This figure plots the sensitivity of investing share in brown stocks to pollution before and after the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we plot the coefficient $\{\beta_k\}$ from the following specification:

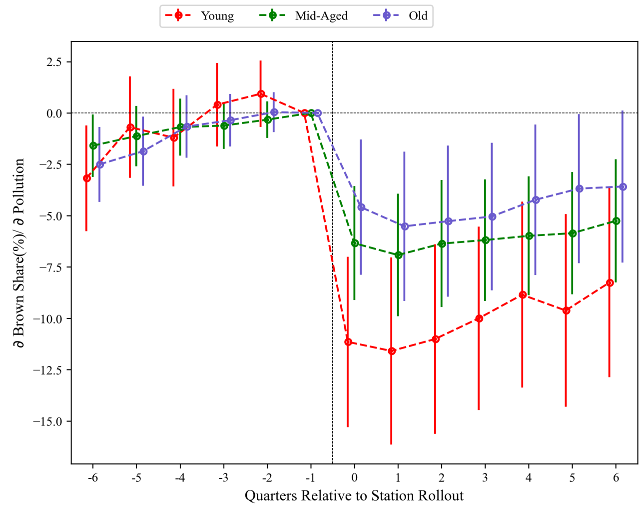
$$Brown\,Share_{z(p)t} = \sum_{k=-6,k\neq-1}^{6} \beta_k \cdot Treat_z \cdot Pollution_{zt} \cdot 1(t=k) + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Capped spikes drawn with the estimated coefficients $\{\beta_k\}$ indicate 95% confidence intervals obtained from standard errors clustered at the zip-code level.

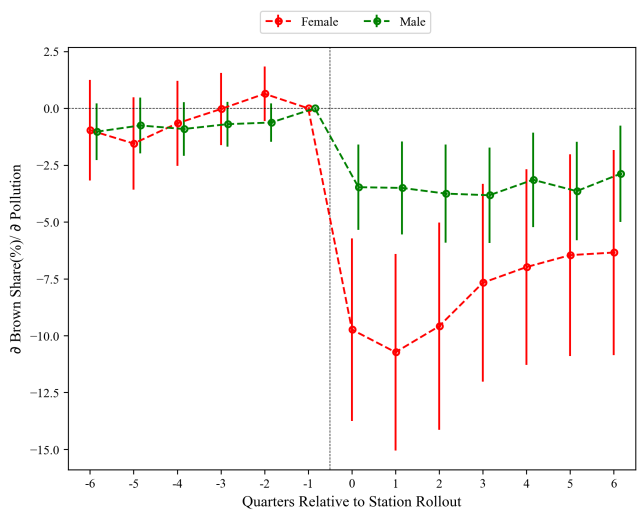
Figure 8: Continuous Ambient Air Quality (CAAQ) Monitoring Station



(a) The Role of New Tech



(b) The Role of Age



(c) The Role of Gender

This figure plots the heterogeneity in the sensitivity of investing share in brown stocks to pollution before and after the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we plot the coefficient $\{\beta_k\}$ from the following specification:

$$Brown\,Share_{z(p)t} = \sum_{k=-6,k\neq-1}^{6} \beta_k \cdot Treat_z \cdot Pollution_{zt} \cdot 1(t=k) + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Panel A plots the heterogeneity based on the trading technology. Panel B plots the heterogeneity based on age and panel C plots the heterogeneity based on gender. Capped spikes drawn with the estimated coefficients $\{\beta_k\}$ indicate 95% confidence intervals obtained from standard errors clustered at the zip-code level.

Table 1: Summary Statistics

	mean	sd	min	p5	p10	p25	p50	p75	p90	p95	max
Average Brown Share Holding%	41.876	9.343	0.000	28.040	32.590	37.670	41.840	45.840	51.190	56.010	100.000
Average Brown Share Holding Female%	41.194	19.085	0.000	4.280	18.080	32.980	41.160	48.400	61.580	76.400	100.000
Average Brown Share Holding Male%	41.935	9.644	0.000	27.640	32.340	37.600	41.900	46.010	51.540	56.580	100.000
Average Brown Share Holding Young%	40.414	17.615	0.000	9.500	20.720	32.030	39.970	47.790	59.270	70.760	100.000
Average Brown Share Holding MidAged%	41.943	11.481	0.000	24.440	30.750	37.180	41.910	46.490	53.100	59.250	100.000
Average Brown Share Holding Old%	42.839	16.066	0.000	15.100	26.390	36.540	42.490	48.450	59.390	70.290	100.000
Pollution AOD	0.543	0.217	0.031	0.264	0.305	0.387	0.505	0.656	0.843	0.954	1.880
No. of Investors(Log)	3.861	1.805	0.693	1.099	1.609	2.485	3.638	4.970	6.541	7.288	10.384
Turnover(Log)	16.846	2.891	0.140	11.407	12.925	15.256	17.195	18.801	20.265	21.028	24.848
Rain	3.171	5.322	0.000	0.000	0.000	0.050	0.940	4.170	9.020	13.140	109.590
Temperature	25.764	4.916	3.290	15.750	18.280	23.410	26.620	29.270	31.230	32.500	36.510

This table reports the summary statistics of the key variable used for empirical analysis in this paper. The summary statistics are for every retail investors of India aggregated at the zipcode-month level. The variables particularly report the average share of brown stocks by a retail investor in general and disaggregated across gender and age cohort in particular.

Table 2: The Impact of Pollution Information Program on Sustainable-Investing-Pollution Gradient

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Pollution×Treated×Post	-1.5319**	-2.1017**	-1.7251*	-2.5007***	-1.3634*	-2.0087**	-1.5687	-2.4050**
	(0.7715)	(0.8397)	(0.9762)	(0.9617)	(0.7675)	(0.8365)	(0.9739)	(0.9607)
Pollution×Treated	0.7576*	1.1183***	0.9427*	1.2966***	0.6620*	1.0331**	0.8627*	1.2234**
	(0.3914)	(0.4283)	(0.4958)	(0.4835)	(0.3893)	(0.4270)	(0.4945)	(0.4826)
Pollution×Post	0.2348	0.8557	0.7933	2.0042	-0.0026	0.7614	0.7592	1.8593
	(0.5419)	(0.6750)	(0.6974)	(1.6520)	(0.5387)	(0.6729)	(0.6951)	(1.6451)
Treat*Post	1.8162***	1.7779***	1.4863**	2.0022***	1.2809***	1.4548***	1.2390*	1.7036***
	(0.4947)	(0.5348)	(0.6469)	(0.5913)	(0.4902)	(0.5318)	(0.6437)	(0.5888)
Pollution	-0.2475	-0.6442*	-0.4243	-1.0167	-0.1799	-0.5311	-0.4084	-0.9371
	(0.2557)	(0.3894)	(0.7409)	(0.7642)	(0.2530)	(0.3883)	(0.7372)	(0.7607)
Post	-0.3977	-0.4760	-0.4484		-0.1441	-0.3940	-0.4295	
	(0.3487)	(0.4183)	(0.4553)		(0.3464)	(0.4169)	(0.4534)	
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.537	0.546	0.599	0.560	0.540	0.548	0.600	0.561

This table studies the elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\ Order\ Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table 3: Alternate Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Pollution×Treated×Post	-2.6090*	-3.8613**	-3.5925*	-4.7804**	-2.3219	-3.7559**	-3.3119	-4.6336**
	(1.5432)	(1.6652)	(2.0674)	(1.9596)	(1.5340)	(1.6571)	(2.0593)	(1.9539)
Pollution×Post	0.4799	2.3032	2.7224	3.5447	-0.0299	2.0192	2.6125	3.3286
	(1.1447)	(1.4301)	(1.7695)	(2.3284)	(1.1372)	(1.4229)	(1.7624)	(2.3172)
Treat×Post	2.4038***	2.7464***	2.5514**	3.2359***	1.7964**	2.4156**	2.2334*	2.9098***
	(0.8876)	(0.9535)	(1.2271)	(1.0957)	(0.8820)	(0.9491)	(1.2219)	(1.0920)
Post	-0.5382	-1.2725	-1.5464		-0.1326	-1.0882	-1.4843	
	(0.6677)	(0.8068)	(1.0276)		(0.6631)	(0.8028)	(1.0232)	
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.537	0.546	0.599	0.560	0.540	0.548	0.600	0.561

This table reports the result from an alternate specification to study the importance of pollution monitoring stations and the consequent information on pollution on the decision of holding brown stocks in the portfolio. Here we classify pincodes based on the average pollution prior to the introduction of the Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zT} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_z$ is the average pollution in zip code z in prior to the introduction of the monitoring stations. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy and $Pollution_z$ is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Panel A reports result for young investors, panel B reports results for middle-aged investors and panel C reports results for elderly investors. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table 4: Placebo: Places with Highest Pollution Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Pollution×Treated×Post	1.2012	1.5156	0.0650	1.6537	1.2989	1.6416	0.1700	1.8181
	(1.5706)	(1.5915)	(0.9312)	(1.9442)	(1.4992)	(1.5257)	(0.9324)	(1.8672)
Pollution×Treated	-1.0939	-1.2164	-0.5017	-1.4563	-1.2148	-1.2534	-0.5502	-1.5059
	(0.9054)	(0.9133)	(0.7102)	(1.0610)	(0.8696)	(0.8843)	(0.7149)	(1.0312)
Pollution×Post	0.6898	0.6805	0.8963	4.3664	0.5482	0.5396	0.9114	3.7765
	(0.8680)	(1.0822)	(1.0698)	(4.6851)	(0.8607)	(1.0797)	(1.0768)	(4.5625)
Treat×Post	-0.3828	-0.7775	0.8015	-1.1842	-0.9477	-1.0896	0.5640	-1.4689
	(1.5895)	(1.6085)	(1.1036)	(1.8910)	(1.5208)	(1.5475)	(1.0796)	(1.8187)
Pollution	0.4226	0.2240	-1.0665	-1.1288	0.3942	0.2654	-1.1467	-1.0127
	(0.5239)	(0.6890)	(1.2487)	(1.8982)	(0.5341)	(0.6939)	(1.2605)	(1.8587)
Post	-0.2899	-0.6033	-1.7539		-0.2374	-0.5153	-1.7213	
	(0.8836)	(1.0744)	(1.1214)		(0.8831)	(1.0753)	(1.1320)	
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y
Observations	44,875	44,783	43,345	44,724	44,875	44,783	43,345	44,724
R-squared	0.612	0.618	0.660	0.625	0.615	0.619	0.661	0.626

This table studies the elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations after the most polluted cities. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Internet Appendix

Figure A1: **Illustrative Example of the Empirical Method**

This figure illustrates the emprical method using Jodhpu as an example. Pincodes within the inner 20-km circles are considered “treated”, while regions between the 40-km and 60-km circles are considered “control” units. The estimate compares changes in outcome between the inner “treated” and the outer “control” donut.

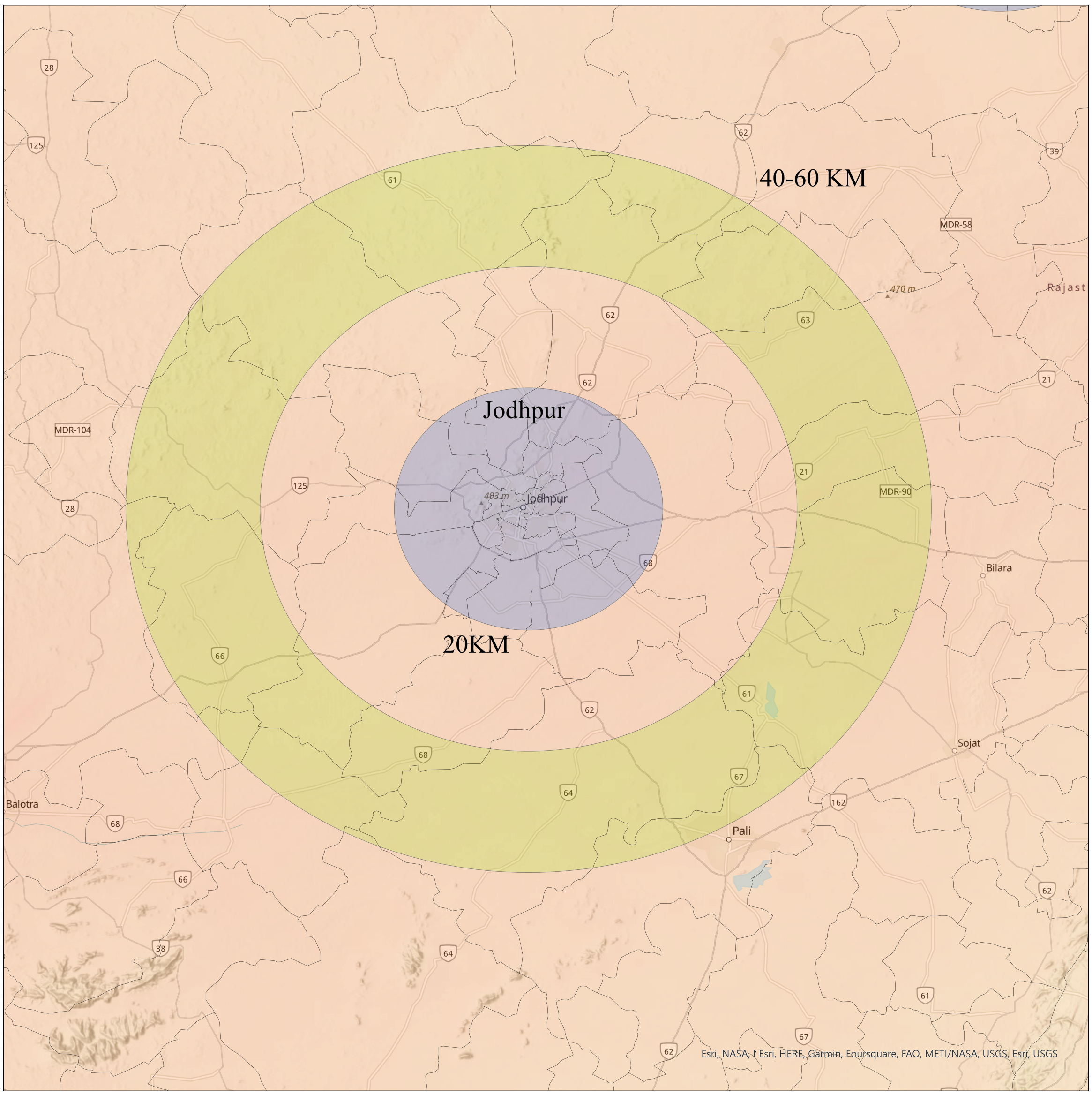


Figure A2: **Geography of NSE investors**

This figure plots the geographical distribution of retail investors across districts who trade at NSE from January 2004 to June 2020.

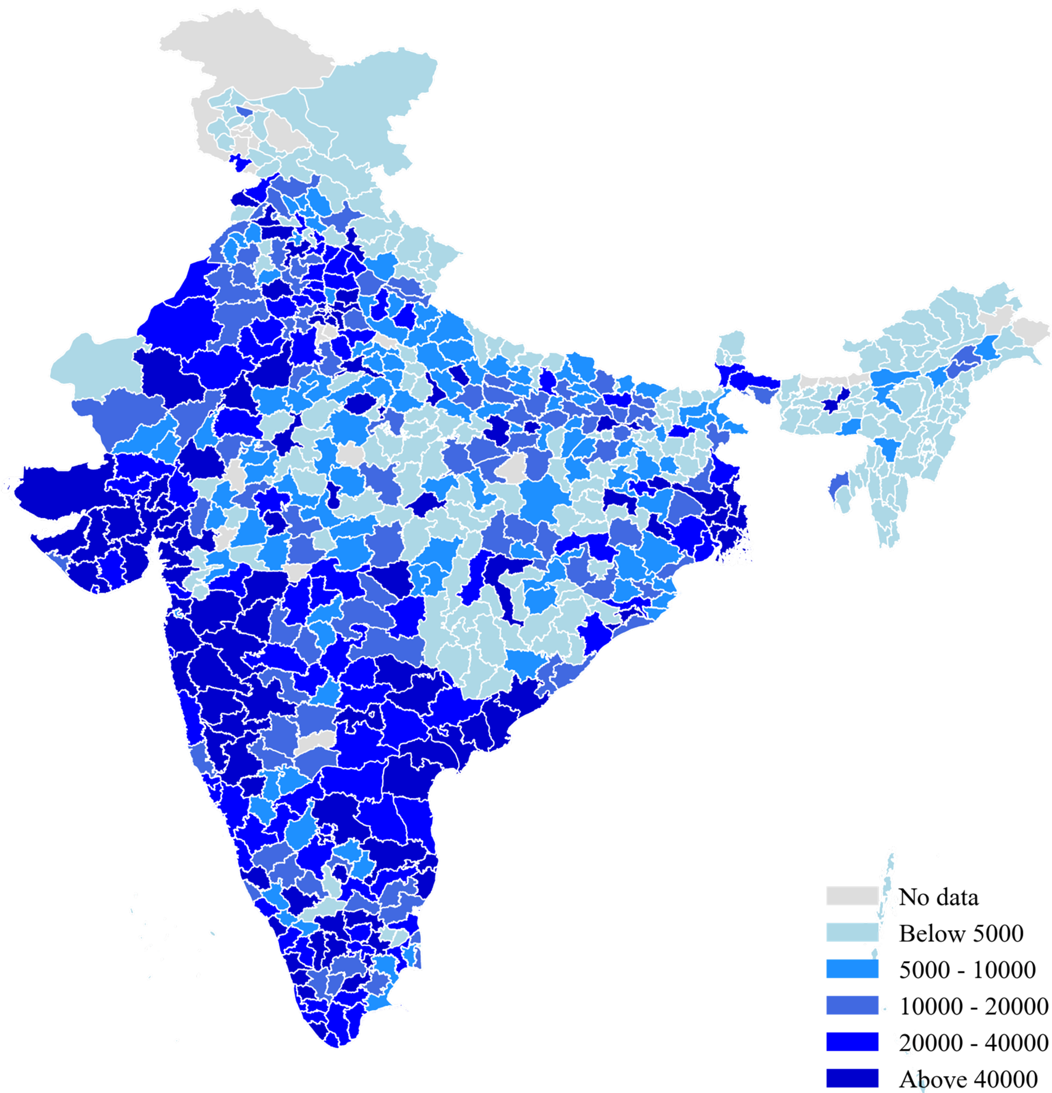
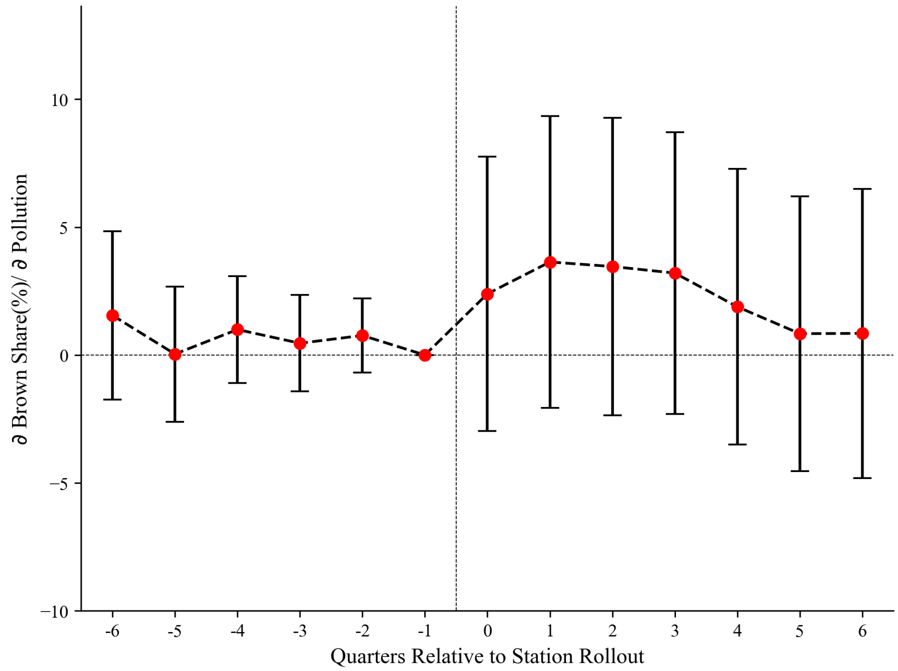


Figure A3: Sensitivity of Brown stocks to Pollution: Institutional Investors



This figure plots the sensitivity of trading in brown stocks by institutional investors to pollution before and after the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we plot the coefficient $\{\beta_k\}$ from the following specification:

$$Brown\,Share_{z(p)t} = \sum_{k=-6,k\neq-1}^{6} \beta_k \cdot Treat_z \cdot Pollution_{zt} \cdot 1(t=k) + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Capped spikes drawn with the estimated coefficients $\{\beta_k\}$ indicate 95% confidence intervals obtained from standard errors clustered at the zip-code level.

Figure A4: **Returns on value-weighted green and brown portfolios**

This figure plots the green and brown portfolios' cumulative returns over January 2000 to December 2019.

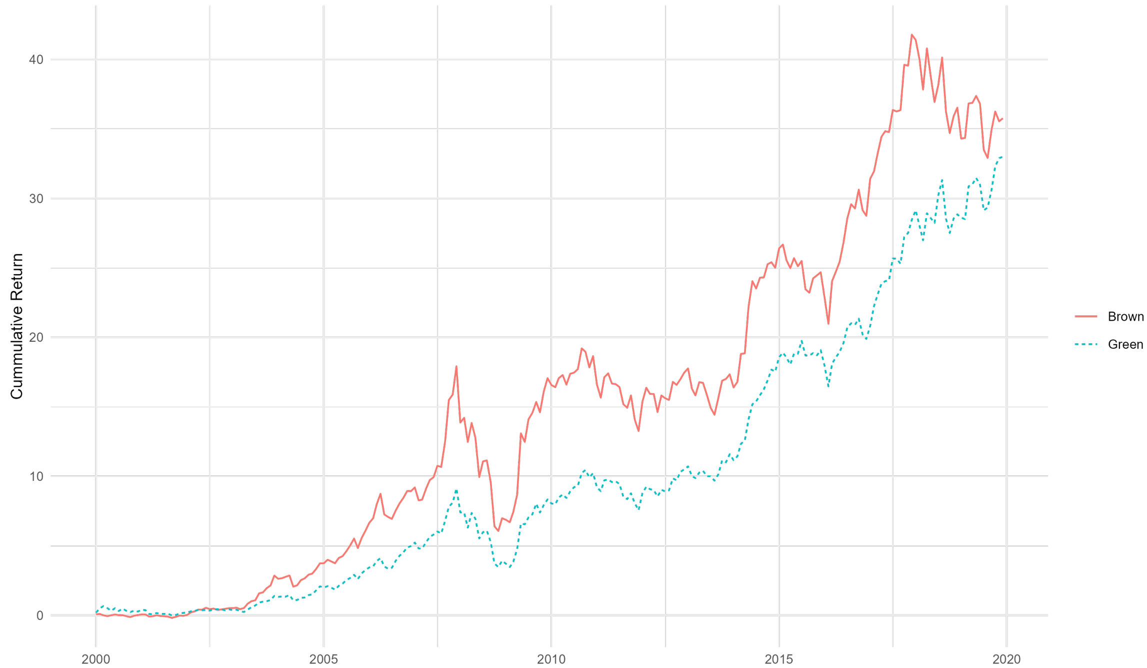


Table A1: The Impact of Monitoring Stations on Pollution CAAQMs

	(1)	(2)	(3)	(4)	(5)	(6)
	Treated		Control		Full Sample	
VARIABLES	Pollution	Log(Pollution)	Pollution	Log(Pollution)	Pollution	Log(Pollution)
Post	0.0004	0.0004	0.0063	0.0036		
	(0.0056)	(0.0034)	(0.0050)	(0.0029)		
Treat*Post					0.0002	0.0000
					(0.0003)	(0.0002)
Station FE	Y	Y	Y	Y		
Year-month FE	Y	Y	Y	Y		
Station*Year-month					Y	Y
Observations	11,489	11,489	14,801	14,801	1,237,258	1,237,258
R-squared	0.619	0.636	0.660	0.686	0.968	0.971

This table studies the elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\ Order\ Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table A2: Changes in the Economic Variables Before and After CAAQMs Rollout

	(1)	(2)	(3)	(4)
VARIABLES	Firm Entry		NightLight Density	
Treat×Post	-0.0208	-0.0399	-0.0943	0.1823
	(0.0420)	(0.0484)	(0.1462)	(0.1416)
Post	0.0983***		1.2520***	
	(0.0346)		(0.0851)	
Observations	32,278	31,904	100,429	100,222
R-squared	0.891	0.908	0.978	0.988
Pincode	Y	Y	Y	Y
Year-month	Y		Y	
Station*Year-month		Y		Y

This table presents the balance tests on how local economic variables before and after the program.We focus on two economic conditions:firm entry and nighttime light density (as a high-frequency proxy for local economic activity). The standard errors are clustered at pincode level.

Table A3: Lowest 10 percentile of Pollution

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Treat×Post	2.3980***	1.4843	0.4892	1.5795	2.2011**	1.3805	0.4634	1.4799
	(0.8526)	(0.9577)	(1.3611)	(1.0081)	(0.8566)	(0.9606)	(1.3636)	(1.0176)
Post	-0.8925	-0.0015	0.3486		-0.8831	0.0296	0.3300	
	(0.7408)	(1.0133)	(1.2936)		(0.7390)	(1.0111)	(1.2936)	
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State*Year-month		Y				Y		
District*Year-Month			Y				Y	
Station*Year-month				Y				Y
Observations	47,996	47,646	46,299	46,280	47,996	47,646	46,299	46,280
R-squared	0.527	0.554	0.589	0.580	0.528	0.554	0.590	0.580

This table studies the elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations after the most polluted cities. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table A4: Pollution Information Program and Sustainable-Investing-Pollution Gradient: Alternative Cutoffs

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var. = Brown Share(%)					
	0-15km		0-10km		0-5km	
Pollution×Treated×Post	-2.4871**	-2.4562**	-2.1983**	-2.1936**	-2.3425*	-2.4248*
	(1.0794)	(1.0731)	(0.9942)	(0.9916)	(1.2787)	(1.2790)
Pollution×Treated	1.2592**	1.2124**	1.0571**	1.0184**	0.9053	0.9353
	(0.5721)	(0.5703)	(0.5062)	(0.5051)	(0.6619)	(0.6636)
Pollution×Post	1.7429	1.5412	1.9651	1.7777	1.4645	1.2084
	(1.6916)	(1.6867)	(1.6967)	(1.6896)	(1.7015)	(1.6974)
Treat*Post	2.2785***	1.8957***	2.0728***	1.7479***	3.0523***	2.6479***
	(0.6891)	(0.6860)	(0.6172)	(0.6149)	(0.8326)	(0.8295)
Pollution	-1.0701	-0.9782	-1.0692	-0.9822	-0.8467	-0.7381
	(0.7839)	(0.7808)	(0.7843)	(0.7803)	(0.7837)	(0.7812)
Observations	448,711	448,711	471,942	471,942	424,220	424,220
R-squared	0.560	0.561	0.561	0.562	0.559	0.561
Local Trading Controls(No. of Investors, Turnover)	N	Y	N	Y	N	Y
Local Weather Controls(Rainfall, Temperature)	N	Y	N	Y	N	Y
Pincode	Y	Y	Y	Y	Y	Y
Station×Year-month	Y	Y	Y	Y	Y	Y

This table presents the robustness checks that apply alternative cutoffs to define the treated regions(i.e. 15km, 10km or 5km) Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\ Order\ Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}, Treat_z \cdot Post_{pt}, Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table A5: Robustness: Removing Large Metropolitan Cities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Pollution×Treated×Post	-2.1120**	-2.4027***	-2.2073**	-2.7009***	-1.9294**	-2.2359**	-2.0348*	-2.5553**
	(0.8453)	(0.8945)	(1.0572)	(1.0153)	(0.8392)	(0.8907)	(1.0551)	(1.0140)
Pollution×Treated	0.9969**	1.2614***	1.0387*	1.3680***	0.9004**	1.1522**	0.9532*	1.2803**
	(0.4221)	(0.4503)	(0.5301)	(0.5049)	(0.4192)	(0.4487)	(0.5286)	(0.5038)
Pollution×Post	0.1721	0.9363	0.7994	2.0449	-0.0588	0.8390	0.7824	1.9175
	(0.5628)	(0.7067)	(0.7514)	(1.7408)	(0.5585)	(0.7041)	(0.7490)	(1.7352)
Treat×Post	2.0514***	1.8278***	1.6328**	2.0253***	1.5107***	1.4806***	1.3770**	1.7110***
	(0.5220)	(0.5529)	(0.6668)	(0.6082)	(0.5171)	(0.5497)	(0.6638)	(0.6055)
Pollution	-0.2487	-0.8179**	-0.8067	-1.2876	-0.1796	-0.7043*	-0.7760	-1.2134
	(0.2595)	(0.3988)	(0.7667)	(0.7874)	(0.2566)	(0.3978)	(0.7626)	(0.7845)
Post	-0.3067	-0.4740	-0.3652		-0.0664	-0.3942	-0.3574	
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y
Observations	473,794	473,794	463,091	473,392	473,794	473,794	463,091	473,392
R-squared	0.543	0.552	0.604	0.565	0.545	0.553	0.605	0.566

This table studies the elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations after the most polluted cities. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table A6: Heterogeneous Response: The Role of New Tech

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Panel A Mobile								
Pollution×Treated×Post	-10.7735***	-9.4363***	-5.1107**	-10.0833***	-9.8264***	-8.1616***	-3.5364*	-8.7934***
	(1.9153)	(1.9185)	(2.1207)	(2.0372)	(1.7966)	(1.8201)	(2.0493)	(1.9375)
Pollution×Treated	4.8522***	4.6318***	2.9987**	5.1342***	4.3007***	3.8543***	2.1723*	4.4245***
	(1.0580)	(1.0634)	(1.2188)	(1.1473)	(1.0030)	(1.0180)	(1.1801)	(1.0985)
Pollution×Post	3.1203***	5.1302***	2.5964*	8.9500***	1.4685*	4.5933***	2.2305*	8.4612***
	(0.9486)	(1.2164)	(1.3550)	(2.7926)	(0.8861)	(1.1517)	(1.3124)	(2.6016)
Treat×Post	12.9445***	11.3718***	7.7129***	11.0515***	10.0700***	8.9844***	5.4540***	8.7014***
	(1.1798)	(1.1734)	(1.3764)	(1.2434)	(1.1064)	(1.1068)	(1.3098)	(1.1709)
Pollution	-1.8260***	-2.5441***	-2.6354	-4.1372***	-1.2076**	-1.8414**	-2.6377	-3.8740***
	(0.5198)	(0.8183)	(1.7691)	(1.4622)	(0.4946)	(0.7996)	(1.7468)	(1.3984)
Post	-2.9850***	-2.9823***	-0.5127		-1.3903**	-2.4777***	-0.3157	
	(0.6263)	(0.7764)	(0.9066)		(0.5821)	(0.7308)	(0.8710)	
Observations	458,195	458,195	449,041	457,475	458,195	458,195	449,041	457,475
R-squared	0.537	0.544	0.584	0.558	0.551	0.556	0.593	0.569
Panel B Internet								
Pollution×Treated×Post	-2.3938*	-2.0753	-0.1046	-2.1225	-1.3632	-0.6745	1.6337	-0.7254
	(1.3431)	(1.3713)	(1.6224)	(1.5657)	(1.2542)	(1.2931)	(1.5453)	(1.4907)
Pollution×Treated	1.0541	1.1284	0.3731	1.0717	0.4981	0.2773	-0.5375	0.3111
	(0.8159)	(0.8369)	(0.9828)	(0.9316)	(0.7783)	(0.8055)	(0.9470)	(0.8979)
Pollution×Post	1.1029	0.9433	-0.6106	2.4031	-0.6511	0.3293	-0.9957	1.8304
	(0.8378)	(1.0640)	(1.2167)	(2.4755)	(0.7682)	(0.9978)	(1.1556)	(2.3359)
Treat×Post	6.1666***	5.1533***	2.9628***	5.1726***	3.0900***	2.5430***	0.4642	2.6172***
	(0.8469)	(0.8563)	(1.0681)	(0.9428)	(0.7777)	(0.7992)	(0.9899)	(0.8788)
Pollution	-0.7728*	-1.2740*	-0.6419	-0.9402	-0.1841	-0.5105	-0.6540	-0.6357
	(0.4666)	(0.7519)	(1.7177)	(1.3677)	(0.4366)	(0.7286)	(1.7006)	(1.3052)
Post	-2.5536***	-1.3962**	-0.2880		-0.8536	-0.8394	-0.0862	
	(0.5606)	(0.6837)	(0.8331)		(0.5200)	(0.6465)	(0.8001)	
Observations	458,195	458,195	449,041	457,475	458,195	458,195	449,041	457,475
R-squared	0.318	0.326	0.384	0.344	0.346	0.351	0.404	0.367
Panel C Others								
Pollution×Treated×Post	-3.7532***	-3.3125***	-3.1694***	-3.9838***	-3.5697***	-3.1159***	-2.9570***	-3.7997***
	(0.9025)	(0.9118)	(1.0471)	(1.0536)	(0.8917)	(0.9052)	(1.0414)	(1.0505)
Pollution×Treated	1.9360***	1.7529***	1.5789***	2.0369***	1.8462***	1.6267***	1.4652**	1.9382***
	(0.4980)	(0.5073)	(0.5736)	(0.5782)	(0.4930)	(0.5046)	(0.5706)	(0.5763)
Pollution×Post	-0.2603	0.9517	1.2091	1.5062	-0.5863	0.8888	1.1989	1.4614
	(0.5775)	(0.7506)	(0.7713)	(1.8482)	(0.5733)	(0.7472)	(0.7687)	(1.8376)
Treat×Post	3.3146***	2.4408***	2.2234***	2.7923***	2.7499***	2.0706***	1.9205***	2.4532***
	(0.5752)	(0.5833)	(0.7053)	(0.6503)	(0.5655)	(0.5771)	(0.6995)	(0.6455)
Pollution	0.0343	-0.7341*	-0.1670	-0.9110	0.1269	-0.6293	-0.1662	-0.8697
	(0.2851)	(0.4326)	(0.7864)	(0.8712)	(0.2815)	(0.4314)	(0.7826)	(0.8658)
Post	-0.3282	-0.8229*	-0.9313*		-0.0048	-0.7520	-0.9238*	
	(0.3828)	(0.4740)	(0.5161)		(0.3796)	(0.4719)	(0.5142)	
Observations	458,195	458,195	449,041	457,475	458,195	458,195	449,041	457,475
R-squared	0.579	0.587	0.632	0.599	0.581	0.588	0.632	0.600
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y

This table studies the heterogeneity in elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations across investors who adopt different technologies(i.e. mobile phone, internet, or others). Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Panel A reports result for young investors, panel B reports results for middle-aged investors and panel C reports results for elderly investors. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table A7: Heterogeneity by Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Panel A: Young								
Pollution×Treated×Post	-7.6956***	-7.5119***	-4.0558*	-7.4996***	-7.2385***	-7.2501***	-3.5523*	-7.2053***
	(1.7136)	(1.8791)	(2.0766)	(2.1505)	(1.6967)	(1.8679)	(2.0804)	(2.1402)
Pollution×Treated	4.7633***	4.5331***	2.3919**	4.5530***	4.4639***	4.2918***	2.1335**	4.3370***
	(0.8503)	(0.9264)	(1.0698)	(1.0176)	(0.8430)	(0.9219)	(1.0703)	(1.0142)
Pollution×Post	1.2748	4.1153***	2.6381*	8.5982***	0.6218	3.8153***	2.4669*	8.1816**
	(1.0897)	(1.3509)	(1.4947)	(3.2930)	(1.0789)	(1.3428)	(1.4902)	(3.2588)
Treat×Post	9.0517***	7.9542***	3.7073***	7.7473***	7.5922***	6.9799***	2.9232**	6.8099***
	(1.0597)	(1.1592)	(1.3438)	(1.2773)	(1.0511)	(1.1546)	(1.3425)	(1.2718)
Pollution	-0.8682*	-1.0958	-2.0970	-3.0610**	-0.6337	-0.7527	-2.0431	-2.8261*
	(0.5033)	(0.7680)	(1.5733)	(1.5163)	(0.4975)	(0.7639)	(1.5743)	(1.5012)
Post	-2.1175***	-2.9941***	-1.6641*		-1.4281**	-2.7304***	-1.5575	
	(0.7098)	(0.8548)	(0.9844)		(0.7036)	(0.8492)	(0.9801)	
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.413	0.421	0.475	0.436	0.419	0.425	0.478	0.440
Panel B: Middle-Aged								
Pollution×Treated×Post	-4.3099***	-4.2561***	-3.3478**	-4.6076***	-4.1376***	-4.1517***	-3.1677**	-4.5061***
	(1.0383)	(1.2334)	(1.3268)	(1.4315)	(1.0372)	(1.2323)	(1.3265)	(1.4323)
Pollution×Treated	2.2820***	2.3265***	1.6279**	2.4103***	2.1610***	2.2377***	1.5457**	2.3452***
	(0.5009)	(0.5870)	(0.6464)	(0.6666)	(0.4987)	(0.5865)	(0.6456)	(0.6668)
Pollution×Post	1.4836**	3.1383***	1.2850	4.7501**	1.2284	3.0173***	1.2334	4.5519**
	(0.7564)	(0.9116)	(0.8864)	(2.2639)	(0.7553)	(0.9122)	(0.8831)	(2.2623)
Treat×Post	4.9029***	4.4461***	3.1775***	4.5964***	4.3414***	4.0800***	2.8837***	4.2507***
	(0.6634)	(0.7885)	(0.8823)	(0.8870)	(0.6596)	(0.7872)	(0.8800)	(0.8872)
Pollution	-0.9166***	-1.7525***	-0.3839	-2.6637***	-0.8118**	-1.6169***	-0.3748	-2.5584**
	(0.3414)	(0.4825)	(0.9036)	(1.0138)	(0.3393)	(0.4822)	(0.9010)	(1.0113)
Post	-1.7295***	-2.2726***	-0.8859		-1.4681***	-2.1727***	-0.8586	
	(0.4852)	(0.5712)	(0.5928)		(0.4844)	(0.5708)	(0.5902)	
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.517	0.523	0.577	0.537	0.519	0.524	0.578	0.538
Panel C: Elderly								
Pollution×Treated×Post	-3.0747**	-2.8609*	-2.3076	-2.3993	-2.8516**	-2.7450*	-2.1144	-2.2736
	(1.3565)	(1.5133)	(1.7195)	(1.7483)	(1.3486)	(1.5055)	(1.7148)	(1.7401)
Pollution×Treated	1.7793***	1.5970**	1.0460	1.5045*	1.6519***	1.4941**	0.9562	1.4229*
	(0.6345)	(0.7093)	(0.8086)	(0.7927)	(0.6319)	(0.7076)	(0.8067)	(0.7907)
Pollution×Post	0.1366	1.4693	0.8878	4.9199*	-0.1823	1.3225	0.8364	4.7108*
	(0.8559)	(1.0367)	(1.0605)	(2.6007)	(0.8524)	(1.0328)	(1.0600)	(2.5885)
Pollution	-0.4018	-0.9519*	-1.0527	-1.9558*	-0.3056	-0.7814	-1.0355	-1.8264*
	(0.3856)	(0.5438)	(0.9847)	(1.1008)	(0.3823)	(0.5430)	(0.9828)	(1.0970)
Treat*Post	4.1886***	3.5314***	1.3263	3.2251***	3.4716***	3.0811***	1.0163	2.7959***
	(0.8435)	(0.9461)	(1.1569)	(1.0590)	(0.8387)	(0.9427)	(1.1542)	(1.0544)
Post	-0.7679	-1.2809**	-0.3453		-0.4259	-1.1528*	-0.3130	
	(0.5469)	(0.6348)	(0.6892)		(0.5435)	(0.6321)	(0.6891)	
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.686	0.690	0.726	0.699	0.688	0.691	0.727	0.700
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y

This table studies the heterogeneity in age specific elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Panel A reports result for young investors, panel B reports results for middle-aged investors and panel C reports results for elderly investors. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.

Table A8: Heterogeneity by Gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var. = Brown Share(%)							
Panel A Female								
Pollution×Treated×Post	-5.5557***	-5.9249***	-5.3078**	-6.8918***	-5.0272***	-5.6203***	-4.6996**	-6.5368***
	(1.6487)	(1.8170)	(2.0709)	(2.1762)	(1.6136)	(1.7939)	(2.0457)	(2.1486)
Pollution×Treated	3.9326***	4.0832***	3.1231***	4.3156***	3.5886***	3.8020***	2.8119***	4.0561***
	(0.7856)	(0.8635)	(1.0062)	(0.9939)	(0.7717)	(0.8555)	(0.9969)	(0.9842)
Pollution×Post	0.6563	1.6732	0.3787	1.8378	-0.1091	1.3247	0.1725	1.3396
	(1.0618)	(1.2965)	(1.4057)	(3.0656)	(1.0369)	(1.2716)	(1.4014)	(3.0043)
Treat×Post	8.4607***	7.6291***	4.9929***	8.0025***	6.7536***	6.4824***	4.0432***	6.8711***
	(1.0325)	(1.1401)	(1.3225)	(1.2922)	(1.0089)	(1.1245)	(1.3010)	(1.2729)
Pollution	-1.1405**	-1.4396**	-1.6196	-1.3263	-0.8588*	-1.0255	-1.5579	-1.0516
	(0.4693)	(0.6536)	(1.3332)	(1.3136)	(0.4579)	(0.6435)	(1.3237)	(1.2865)
Post	-1.5929**	-0.6825	-0.2197		-0.7818	-0.3725	-0.0939	
	(0.6720)	(0.7953)	(0.9116)		(0.6567)	(0.7808)	(0.9086)	
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.664	0.669	0.701	0.679	0.670	0.674	0.703	0.683
Panel B Male								
Pollution×Treated×Post	-1.8036**	-2.1735***	-2.0207**	-2.4068**	-1.6314**	-2.0787**	-1.8625*	-2.3108**
	(0.7722)	(0.8349)	(0.9716)	(0.9539)	(0.7676)	(0.8314)	(0.9695)	(0.9519)
Pollution×Treated	0.8740**	1.1210***	0.9980**	1.2251**	0.7727*	1.0351**	0.9196*	1.1548**
	(0.3972)	(0.4315)	(0.5057)	(0.4865)	(0.3951)	(0.4303)	(0.5043)	(0.4857)
Pollution×Post	0.4904	1.4064**	1.2662*	3.1620*	0.2451	1.3054*	1.2268*	3.0093*
	(0.5563)	(0.6996)	(0.7274)	(1.6902)	(0.5523)	(0.6966)	(0.7247)	(1.6814)
Treat×Post	2.2516***	2.0861***	1.8353***	2.2038***	1.7008***	1.7508***	1.5835**	1.8974***
	(0.4992)	(0.5331)	(0.6512)	(0.5872)	(0.4937)	(0.5298)	(0.6484)	(0.5845)
Pollution	-0.3372	-0.8750**	-0.4119	-1.6909**	-0.2614	-0.7558*	-0.3974	-1.6071**
	(0.2630)	(0.4067)	(0.7743)	(0.7843)	(0.2599)	(0.4055)	(0.7706)	(0.7797)
Post	-0.6303*	-0.9231**	-0.8077*		-0.3691	-0.8358*	-0.7853	
	(0.3610)	(0.4358)	(0.4794)		(0.3581)	(0.4339)	(0.4774)	
Observations	499,036	499,036	488,681	498,310	499,036	499,036	488,681	498,310
R-squared	0.538	0.547	0.599	0.560	0.541	0.548	0.600	0.562
Local Trading Controls(No. of Investors, Turnover)	N	N	N	N	Y	Y	Y	Y
Local Weather Controls(Rainfall, Temperature)	N	N	N	N	Y	Y	Y	Y
Pincode	Y	Y	Y	Y	Y	Y	Y	Y
Year-month	Y				Y			
State×Year-month		Y				Y		
District×Year-Month			Y				Y	
Station×Year-month				Y				Y

This table studies the heterogeneity in gender specific elasticity of the share of brown stock held in a portfolio with information on pollution following the installation of Continuous Ambient Air Quality (CAAQ) Monitoring Stations. Specifically we report the coefficient β from the following specification:

$$Brown\,Share_{z(p)t} = \beta \cdot Treat_z \cdot Pollution_{zt} \cdot Post(pt) + \sum_{k=1}^{6} \beta_k \cdot Lower\,Order\,Interactions + X_{z,t} + \alpha_z + \alpha_{p,t} + \varepsilon_{z,t}$$

where $Brown\,Share_{z(p)t}$ denotes the average share of brown stocks of a retail investor in zip code z belonging to a station area p in time t. A station are p is contains zip codes belonging within 20kms of the station – treatment zip code – and 40km to 60km around the station – control zip codes. $Treat_z$ takes a dummy 1 if a zip code belonged to the treatment group. $Pollution_{zt}$ takes is the pollution in zip code z in time t. Lower Order Interactions includes the interactions: $Treat_z \cdot Pollution_{zt}$, $Treat_z \cdot Post_{pt}$, $Post_{pt} \cdot Pollution_{zt}$, and $Treat_z$, $Pollution_{zt}$ and $Post(pt)$ dummies. Meanwhile $Treat_z$ dummy is absorbed with zip code fixed effect and $Post(pt)$ dummy is absorbed in Station*Year-month fixed effects. $X_{z,t}$ is the set of controls in zip code z in time t. α_z is the zip code fixed effect and $\alpha_{p,t}$ is the station area × time fixed effect. Columns 1-4 report results without controls while columns 5-8 reports results with controls. Panel A reports result for female investors while panel B reports results for male investors. Column 1(5) pincode and year-month fixed effects. Column 2(6) include pincode and state*year-month fixed effects. Column 3(7) include pincode and district*year-month fixed effects. Column 4(8) include (CAAQ) Monitoring Stations*year-month fixed effects. The standard errors are clustered at pincode level.