Internet Appendix

Air Pollution, Behavioral Bias, and the Disposition Effect in China

In this Internet Appendix, we provide additional descriptions of variables in our data sample in Internet Appendix 1 and more information about scientific findings related to air pollution in Internet Appendix 2. Then, in Table IN1, we provide robustness checks to verify that our results are robust to the selection of mutual funds and to their differing channels of distribution. In Table IN3, we report additional robustness checks related to the quasi-experiment of the "Huai-river policy." In Table IN4, we provide further account-level analysis based on Cox proportional hazards models. Finally, Tables IN5 and IN6 provide additional evidence on tests related to the mechanism of air pollution-induced disposition effect. Below, we discuss additional analysis provided in these tables.

Table IN1 presents various robustness checks for the baseline specification, as reported in Panel B of Table 2 in our main text. Recall that in our main text, the city-level disposition effect is constructed based on the probability that investors located in a given city trade all equity mutual funds offered by the fund family. Could this sample selection of funds distort our measurement of the disposition effect? For instance, some funds may have wider coverage than others—could the fund family selectively offer some funds only to investors in some regions, possibly distorting the representativeness of our measurement of the disposition effect? Our strategy for addressing this concern is to conduct a robustness check in which the disposition effect is measured based on only funds traded in almost every city in our sample. More explicitly, in Panel A, we construct a city-level disposition effect based only on investors' trading of the top fund or top three funds offered by the mutual fund family, where top funds are ranked based on their national account coverage. The top fund is offered in every city in our sample, whereas the top three funds are offered in more than 90% of the cities in our sample—the trading of these funds will not be subject to investor-selection issues.

In a similar spirit, another potential concern about the fund sample selection could relate to fund distribution channels. Mutual funds are in general distributed in China through three major channels: banks, brokerage firms, and online distribution. Could some funds be offered only through selected channels, subsequently affecting the measurement of the disposition effect? Our empirical strategy for addressing this concern is, again, to focus on top funds that use all three distribution channels. In Panel B, we separately compute the disposition effect for each of the three distribution channels that top funds may use in each city—including banks, brokerage firms, and online distribution—and then construct the city-level disposition effect as the average value across the three channels. In Panel C, we further examine whether the brokerage channel, which has attracted attention from the previous literature on Chinese stock investors (e.g., Seasholes and Zhu, 2010), could give rise to a different influence of AQI. More specifically, we interact AQI with the dummy variable that takes the value of one when the brokerage channel is used and zero otherwise.

We can see that our results remain robust in these scenarios, suggesting that our results are not contaminated by potential fund selection issues or the selection of distribution channels through which fund families distribute funds. Our results are robust regardless of how we change the selection of funds in our sample. In other words, in our sample, all funds have broad coverage that is uncorrelated with the potential influence of air pollution on behavioral bias. Moreover, the use of the brokerage channels does not affect the sensitivity of investor behavior to air pollution. Although many investors physically visit brokerage firms to trade stocks early on, the popularity of online and cell phone-app based trading among fund investors over the last decade has helped eliminate the potential difference across different distribution channels. This potentially explains the insignificant difference across different distribution channels.

We next further confirm the negative influence of air pollution on performance by conducting a path analysis (e.g., Wright, 1934; Pevner, Xie, and Xin, 2015) in which we treat air pollution as the exogenous variables that can influence investors' trading performance either directly or indirectly though the mediating variable of the disposition effect. The analysis is reported in Table IN2 in this appendix.

Since the goal of the path analysis is to examine whether AQI can lead to underperformance, we first construct the measure of counterfactual performance in the spirit of Table 2 (Panel B) that can help us identify investors' potential underperformance. More explicitly, for each aggregate city account, we classify the aggregate trading of its investors into two categories: a net buy or a net sell of a fund. When a fund is sold in aggregate (i.e., a net sell), we compute its counterfactual performance as the return or market-adjusted return in the 20-working-day period following the sale date (our results are robust to this horizon). In the case of a net buy, the counterfactual performance is defined as the negative of actual returns generated by the fund during the same period. The counterfactual performance, therefore, measures the part of fund returns that the aggregate investors of a city have missed due to their actual selling/buying decisions. Because the counterfactual performance measures what investors miss, we should observe a positive relationship between AQI and counterfactual performance if air pollution leads to trading underperformance.

In Panel A, we pool the counterfactual performance for all city-fund trading on each date, whereas in Panel B we focus on the trading of the top fund (again in terms of national account coverage) of the family. We first estimate the *direct path* by regressing counterfactual performance on the degree of air pollution (*Log(AQI)*), with all control variables as included in Column 2 of Panel B Table 2. In Model (1) of Panel A, for instance, the first line reports that the coefficient between the counterfactual return and the AQI is 20.3 bps. This coefficient confirms that more severe pollution is generally associated with larger underperformance. As for the magnitude, a one-standard-deviation change in air pollution is roughly associated with an annualized counterfactual return of 2.12%, i.e., $20.3 \times (260/20) \times 0.801 = 212\text{bps} = 2.12\%$, where the ratio $260/20$ roughly translates 20 working day returns into annualized ones based on the assumption that a year contains approximately 260 working days, and 0.801 is the one-standard-deviation change in air pollution. The latter number is constructed as follows. Since we use the logarithm of AQI as our main independent variable, its one-standard-deviation change is computed as the difference between the logarithm of the average level of AQI and that of AQI one standard deviation below its mean value, i.e., $ln(\mu_{AQI}) - ln(\mu_{AQI} - \sigma_{AQI}) = 0.801$, where μ_{AQI} and σ_{AQI} refer to the mean value and the standard deviation of AQI in our sample.

We then employ a structural equation model (SEM) to estimate the indirect effect of haze on counterfactual performance through the disposition effect. The equations in the SEM include a regression of the mediation variable on selling-date air pollution (i.e., the line labeled "*I. P(Selling Date AQI, Path),*" where "*Path*" refers to the mediating variable of the disposition effect) and a regression of counterfactual performance on the mediating variable (i.e., the line labeled "*II. P(Path, Counterfactual Ret)*" in the table), again with all control variables as included in Model (2) of Panel B Table 2. In particular, the line labeled "*Indirect effect, P(Haze, Path)*P(Path, Counterfactual Ret)*" reports the indirect effect of air pollution on counterfactual performance mediated through the path of the disposition effect, which synchronizes the impact of air pollution on the path and that of the path on performance. The significance of the indirect effect (i.e., the product coefficient) is assessed using the Sobel (1982) test. We can see that the indirect effect is highly significant. In terms of economic magnitude, the contribution from the indirect path is approximately 11.6% of that from the direct effect (i.e., $2.355/20.3 = 11.6\%$, where 2.355 is the coefficient of the indirect effect) or 10.4% of the total effect (2.355/(20.3+2.355)=10.4%), which is also economically sizable.

In Model (2) of Panel A, counterfactual fund returns are adjusted by the market return. The results are very similar to those found in Model (1). In this case, a one-standard-deviation change in air pollution is associated with an annualized counterfactual performance of 1.86% (i.e., $17.85 \times (260/20) \times 0.801 = 186\text{bps}$), and the magnitude of the impact mediated through the indirect channel is approximately 11.8% of that from the direct effect (i.e., $2.099/17.85 = 11.8\%$). In Panel B, we focus on investors' trading on the leading fund of the family in terms of geographic coverage. We basically find very similar results as those in Panel A.

Overall, Table IN2 confirms that investors' trading on highly polluted dates is associated with underperformance, in which the air pollution-induced disposition effect plays an important role. One caveat of the path analysis is that it captures only the average and linear effect of the indirect channel aggregated at the city level, and in this regard may not reflect the full influence of air pollution on particularly severe trading mistakes at the account level that may lead to substantial financial losses, such as trading against momentum. Tables 9 and 10 in the main text, as well as Tables IN5 to IN7, will provide further analysis on the potential economic ground through which air pollution influences investor behavior.

Next, we explain the robustness checks of the RD test in Table IN3. In the first set of tests, we further assess the testing power of the Huai-river policy by conducting a placebo test in which we apply the same RD and two-stage tests to two artificial lines that are 5 degrees north and 5 degrees south of the geographic location of the Huai River. The results are tabulated in Panel A of Table IN3. Intuitively, the RD test should not yield any significant results because the heating policy applies to the river but not to these two artificial lines. Consistent with this intuition, we do not find any significant changes across these lines, suggesting that our tests have the proper power of rejecting non-existent influences of air pollution. Next, in Panel B, we choose different bandwidths (6° and 8°) in estimating the RD test. Recall that a bandwidth of 8° means that we include only cities that are located within 8° of the latitude of the Huai River. Although the RD test leaves some degrees of freedom in choosing the bandwidth (see footnote 6 for more discussion), our results are robust to this threshold.

Panel C applies the main RD test to a special group of investors: migrants from the southern part of China trading in both the southern and northern parts of the Huai River. First, we identify migrant investors based on their city of birth (which we obtain from their national residential identification cards) and the city of their address at the time they open a trading account—migrant investors are those whose birth cities differ from their trading cities. We focus on migrants coming from the southern part of China because this part of China has less air pollution, according to the RD test. When these investors move to different cities located on both sides of the Huai River, we can examine their trading behavior with respect to air pollution introduced by the Huai River Policy. We find that even within this subsample of investors, whose birth regions are relatively homogeneous, residents to the north of Huai River still exhibit a higher disposition effect than residents to the south. This effect is difficult to generate through cultural or regional factors unrelated to air pollution, which further validates our causal interpretation of the RD test.

Panel D further controls for sunshine (daily-average sunshine of a city in a year) and excludes important political event-dates or tier-one cities from our test. We have conducted these robustness checks in our baseline regression of Table 2. We repeat the robustness checks here for two reasons. First, we want to show that such robustness checks also apply to endogeneity tests. Second, and more importantly, the exclusion of tier-one cities could have additional value for the RD test, as it might be difficult to balance the influence of these large cities across the Huai River. In both cases, we find that our main results remain robust. Moreover, when we apply the same robustness checks to DID tests, our results remain the same.

In Panel E, we apply the RD test to trading volume and the fraction of selling-orders. We do not find any difference across the river in this case. In addition to these trading variables, we have also verified that all city-level variables used in our analysis, such as GDP or GDP growth, are not significantly different on the two sides of the river. This absence of a significant difference is consistent with the literature (Almond et al., 2009; Chen et al., 2013). These findings suggest that the higher observed disposition effect in high-polluting regions is neither spuriously correlated with investors' general trading frequency nor attributable to fundamentals of the cities (other than pollution). The only valid interpretation of our results across all RD-related tests is, therefore, that investors in the polluted region make more trading mistakes, not that they alter the volume of their trading activity.

Table IN4 conducts various robustness checks for the results based on the Cox proportional hazards model as reported in Panel B of Table 8. Panel A further controls for investor characteristics—its layout is similar to Panel B of Table 8 except that we include investor characteristics as control variables. We can see that controlling for investor characteristics does not affect the main result that air pollution intensifies the disposition effect.

In Panel B, we first adopt an alternative time window for the classification of news dates (from the announcement day to five days after) and report the results in Models (1) and (2). We can see that the influence of AQI is still concentrated on no-news dates. Next, we further split news dates according to different types of news. In particular, Models (3) and (4) are estimated when selling dates are associated with the announcement of financial news—i.e., quarterly report and dividend policies—or with operational news (such as the turnover of managers and changes in investment policies related to management fees, front-load, and redemptions, etc.), respectively. We still classify news days as the period from the announcement day to three days later. We can see that in both cases, the influence of AQI becomes insignificant. Hence, when investors pay attention to fund fundamentals, they are less exposed to the influence of air pollution.

We finally provide additional tests related to momentum and the magnitude effect as implied by our proposed mechanism in Tables IN5 and IN6. Table IN5 first employs a different empirical specification in examining whether air pollution intensifies the tendency of selling winners. In particular, Panels A1 and A2 of Table 9 in the main text rely on pooled Logit regressions in assessing this impact. Although it is reasonable to use Logit specifications when the dependent variable measures the likelihood of selling, one potential concern is that we cannot control for fund and time-fixed effects in such specifications. To show that our results are robust to these fixed effects, Panels A1 and A2 of Table IN5 adopt a panel (OLS) specification with fund- and time-fixed effects to control for fund characteristics and trends. We can see that the main conclusion that air pollution intensifies the tendency of selling winners remains highly robust in this alternative empirical specification.

We next provide additional tests to extend the counterfactual analysis. In the main text, the counterfactual analysis focuses on investors' selling activities in the first two weeks of any calendar month or those in the first two weeks of any post-announcement period. In Panels B1 and B2, we expand the counterfactual analysis to funds sold in all days (MOM_{-t} becomes the fund return realized in the one-month period prior to the sale date t) and to funds sold in an extended 20 working day period of the post-announcement period. We first find that consistent with Table 9, winners sold on polluted days can generate significant counterfactual returns in these cases. Second, the magnitude of the interaction term $MOM_{-t} \times ln(AQI_t)$ is on par with but smaller than that reported in the main text. For instance, the coefficient is 12.7 in Model (3) of Table 9, Panel B1 and 7.4 in the corresponding model in this table. This difference is reasonable because Table 9 focuses on the more prominent calendar month momentum while this table examines a more generalized version of momentum, which is defined as fund return realized in the one-month period prior to the sale date.

As a complement to counterfactual analysis, Panels C1 and C2 of this table further confirm that the calendar month return and post-announcement-period return of funds can be predicted by the previous month's fund return and announcement-period fund return, respectively. In other words, both time-series momentum and post-announcement price drift are highly significant in our data, confirming that the selling winner typically incurs opportunity costs by missing the momentum profitability that can be generated by past winners.

Finally, we conduct robustness checks on the regression discontinuity analysis and tabulate the full specification of the magnitude test, as reported in Table 10. Panel A1 of Table IN6 applies the regression discontinuity analysis to different holding horizons when we follow Ben-David and Hirshleifer (2012) to use 3rd-, 4th-, and 5th-degree polynomials in fitting the probability of selling for both positive and negative ranges of returns. The observation remains the same across all specifications with different degrees of polynomials that the sign effect of $I(ret > 0)$ is significantly positive in short holding horizon, insignificant in the mid-horizon, and significantly negative in the long holding horizon. In Panel A2 of Table IN6, we include the influence of air pollution. Again, our main finding that air pollution does not intensify sign effect remains highly robust.

In Panel B of Table IN6, we tabulate the full specification of the magnitude test as reported in Panel B of Table 10. As discussed in the main text, our main finding is that air pollution significantly enhances the intensity of selling larger gains. Although in mid-holding periods, air pollution also seems to enhance the selling of larger losses (i.e., triggering a V-shaped disposition effect), this effect is not pervasive in other horizons. Hence, the most important impact of air pollution is to trigger the selling of more salient gains, which is largely consistent with realization utility models such as Barberis and Xiong (2012).

Additional References for the Internet Appendix

Sobel, M. E., 1982. Asymptotic confidence intervals for indirect effects in structural equation models. Sociological Methodology, 13, 290-312.

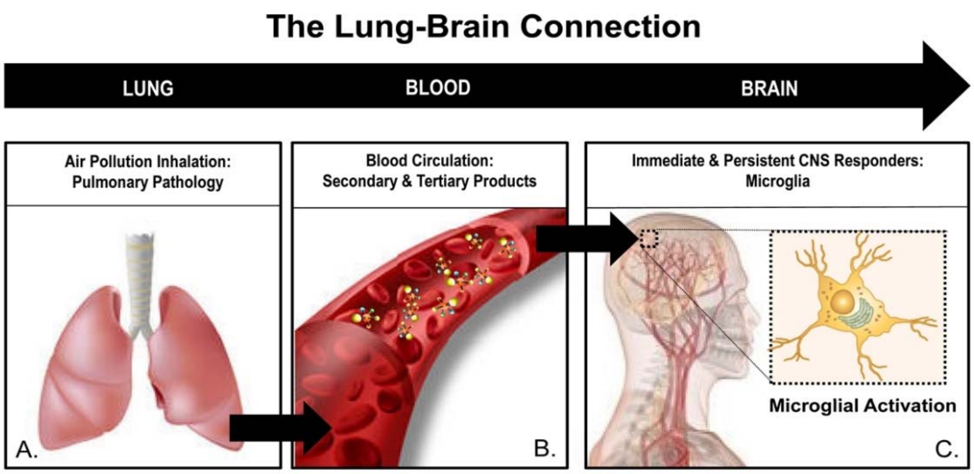
Internet Appendix 1. Additional Description of the Database

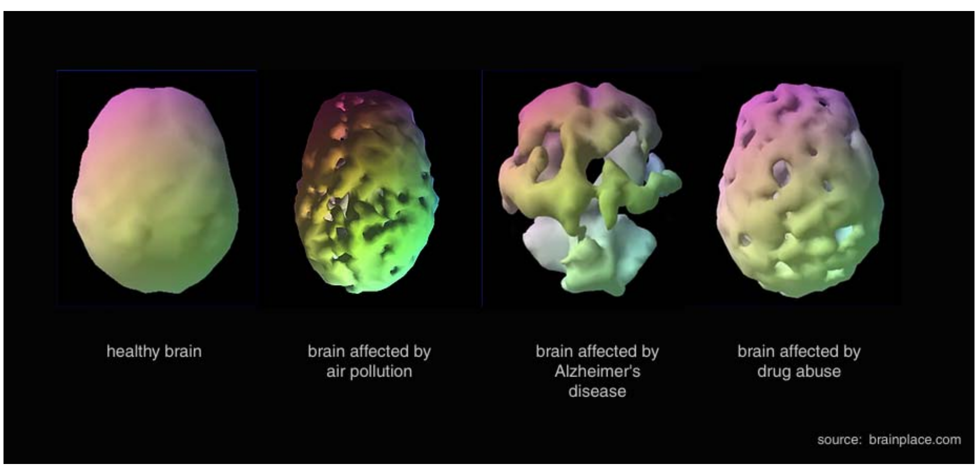
The dataset consists of three main parts: investor account-level information, dividend distribution information and investor trade information. The investor account-level information describes an individual investor's account, including the investor's unique national identity (e.g., date of birth, age, gender, and location) and the account status (e.g., application date, confirmation date, and bonus type). Dividend distribution information includes the amount distributed to each investor based on his/her fund holdings, and this table includes the fund code, investor ID, investor location, dividend date and bonus type. Finally, the investor trade information includes all of the necessary information regarding an investor's purchases and sales of the fund, including an investor's trade type, trade fees and the channel used to purchase shares.

Panel A: Investor Account-level Information	
CustID	Investor's ID
Birth	Investor's date of birth
Gender	1=female, 2=male
Confirm Date	Account confirmation date
Center*	Business center
Channel*	Business channel
Region	Investor province location
Postcode	Investor location postcode
BusinFlag*	Business type
Panel B: Dividend Distribution Information	
CustID	Investor's ID
FundID	Fund code
Regdate	Registration date for the dividend
Exdate	Ex-dividend date
PayDate	Date of payment
Bonustype	0=dividend reinvested, 1=cash dividend
Totalshare	Investor holding of the fund for dividends
Unitprofit	Dividends per share
Totalprofit	Total dividend proceeds
Panel C: Investor Trading Information	
BusinFlag*	Transaction type
Cdate	Confirmation date
Balance	Application amount (cash)
Shares	Application amount (shares)
Confirmbalance	Confirmation amount (cash)
Netvalue	Net value per share (based on date)
Transactionfee*	Total transaction fees
AGIO	Discount percentage on transaction fees
Center	Business center
Channel	Business channel
Agency No.	Channel agency code
Region	Investor province location
City	Investor city location
Postcode	Investor location postcode

Internet Appendix 2. Illustrations of Brain Damages from Two Science Blogs

This Appendix cites two blogs and their posted figures to demonstrate the influence of air pollution on human brains. The first blog (top figure), "Urban air pollution exposure may trigger toxic responses in brain cells and impact neurodegenerative disease pathways" (February 18, 2014) [1], explains the pivotal role microglia plays through which air pollution can immediately affect human brains: "Under normal conditions, microglia primarily serve as the defenders of the central nervous system…But microglia can be dangerous when they are exceptionally 'angry' and are known to leave behind significant bystander damage to neighboring cells. This adverse behavior may lead to the development of any number of neurodegenerative diseases, including Parkinson's disease, Alzheimer's disease, or Gulf War Illness." The second blog (with the figure below), "Can air pollution cause permanent brain damage?" (June 3, 2016)[2], compares the SPECT scan of a brain of a person exposed to air pollution to those of Alzheimer's disease or drug abuse.





[1] https://medicalxpress.com/news/2014-02-urban-air-pollution-exposure-trigger.html
[2] https://u-earthblog.com/2016/06/03/can-air-pollution-cause-permanent-brain-damage/

Table IN1: Robustness Checks on Fund Selection and Distribution Channels (Baseline)

This table conducts various robustness checks for the baseline specifications, as reported in Panel B of Table 2 in our main text. In Panel A, city-level disposition effect is constructed based only on investors' trading of the top fund or top three funds offered by the mutual fund family, the top funds of which are ranked based on their national account coverage. In Panel B, we separately compute the disposition effect for each of the three distribution channels that a fund can use in each city—including banks, brokerage firms, and online distributions—and then construct the city-level disposition effect as the average value across the three channels. Panel C examines whether the brokerage channel may play a special role. Robust t-statistics are reported in parentheses and are based on standard errors clustered by city and date. Superscripts of *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

Panel A: Baseline results when disposition effect is constructed from trading of top funds in coverage

	(1)	(2)	(3)	(4)
	Top Fund (Coverage)		Top 3 Fund(Coverage)	
Log_AQI	0.025***	0.026***	0.038***	0.038***
	(3.01)	(3.02)	(4.15)	(4.12)
Log_GDP		-0.017		-0.044
		(-0.47)		(-1.15)
Log_pop		0.043		0.033
		(1.37)		(1.07)
Log_num_domestic_firm		-0.032*		0.004
		(-1.71)		(0.20)
Log_gov_income		0.021		0.015
		(1.04)		(0.77)
Constant	0.203***	0.183	0.150***	0.439
	(4.92)	(0.33)	(3.34)	(0.78)
Fixed Effect	City, Day-of-the-Week,Month-of-the-Year; S.E. clustered by City-day			
No. of Obs	138,186	138,186	141,988	141,988
R-Sqr	0.02	0.02	0.02	0.02

Panel B: Baseline results when disposition effect is averaged over different distribution channels

	(1)	(2)	(3)	(4)
	Top Fund (Coverage)		Top 3 Fund(Coverage)	
Log_AQI	0.022***	0.022***	0.034***	0.034***
	(2.63)	(2.64)	(3.79)	(3.76)
Log_GDP		-0.021		-0.045
		(-0.57)		(-1.22)
Log_pop		0.042		0.033
		(1.36)		(1.04)
Log_num_domestic_firm		-0.031		0.005
		(-1.64)		(0.23)
Log_gov_income		0.019		0.014
		(0.95)		(0.69)
Constant	0.133***	0.179	0.071	0.406
	(2.99)	(0.33)	(1.49)	(0.73)
Fixed Effect	City, Day-of-the-Week,Month-of-the-Year; S.E. clustered by City-day			
No. of Obs	145,670	145,670	149,571	149,571
R-Sqr	0.02	0.02	0.02	0.02

Panel C: A specific test on brokerage vs. other distribution channels

	(1)	(2)	(3)
	Full Sample	Top Fund Sample	Top 3 Fund Sample
Log_AQI	0.032***	0.023**	0.034***
	(3.05)	(2.45)	(3.39)
D(Brokerage)	-0.053	-0.005	-0.009
	(-0.64)	(-0.07)	(-0.11)
Log_AQI*D(Brokerage)	0.005	-0.001	0.000
	(0.26)	(-0.07)	(0.01)
Log_GDP	-0.069*	-0.022	-0.047
	(-1.81)	(-0.62)	(-1.28)
Log_pop	0.030	0.043	0.034
	(0.94)	(1.42)	(1.09)
Log_num_domestic_firm	0.039*	-0.030	0.005
	(1.80)	(-1.63)	(0.27)
Log_gov_income	0.030	0.017	0.011
	(1.42)	(0.85)	(0.57)
Constant	0.465	0.293	0.541
	(0.83)	(0.54)	(0.99)
Fixed Effect	City, Day-of-the-Week,Month-of-the-Year; S.E. clustered by City-day		
Observations	152,006	145,670	149,571
R-squared	0.02	0.02	0.02

Table IN2: Path analysis of the effects of AQI on counterfactual performance through the disposition effect

This table presents the results of path analysis of the relationships among the explanatory variable (air pollution or *Log(AQI)*, the mediating variable (the disposition effect measured at the city level) and the counterfactual performance of city-level trading activities. For each aggregate city account, we first classify the aggregate trading activity of its investors on a particular fund on a given date into two categories: a net buy or a net sell. When a fund is sold in aggregate, we then compute its counterfactual performance as the return or market-adjusted return of the fund in the 20-working-day period following the sale date. In the case of a net buy, the counterfactual performance is computed as the negative of fund return. In both cases, counterfactual performance captures the part of the fund return that the aggregate investors of a city have *missed* due to their trading decisions. In Panel A, we pool the counterfactual performance for all city-fund trading on each date, whereas in Panel B we focus on the trading of the leading fund (in terms of geographic coverage) of the family. We then estimate a structural equal model (SEM) of the direct effect of haze on the counterfactual performance of investors as well as the indirect effect of haze on counterfactual performance explicitly through the disposition effect. The equations in the SEM include a regression of counterfactual performance on haze and mediating variables and regression of the mediation variables on haze, with all control variables as included in Column 2 of Panel B in Table 2. The significance of the indirect effect (i.e., the product coefficient) is assessed using the Sobel (1982) test. Superscripts of *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

Panel A: Path analysis on the direct and indirect influences of AQI on pooled counterfactual returns (20D, in bps) of all funds

	All Funds, Pooled		All Funds, Pooled	
	(1)		(2)	
	Raw Return (bps)	*t-stat*	Mkt-adj Return (bps)	*t-stat*
Direct Path				
P(Selling Date AQI, Counterfactual Ret)	20.306	*4.15*	17.854	*4.10*
Mediated Path (through the disposition effect)				
I. P(Selling Date AQI, Path)	0.030	*9.49*	0.030	*9.49*
II. P(Path, Counterfactual Ret)	78.461	*23.43*	69.915	*23.44*
Indirect effect, P(Haze, Path)*P(Path, Counterfactual Ret)	2.355	*8.79*	2.099	*8.80*
Number of Observations			213182	
Adj. Goodness-to-fit Index			0.92	

Panel B: Path analysis on the direct and indirect influences of AQI on pooled counterfactual returns (20D, in bps) of the leading fund

	The Leading Fund		The Leading Fund	
	(1)		(2)	
	Raw Return (bps)	*t-stat*	Mkt-adj Return (bps)	*t-stat*
Direct Path				
P(Selling Date AQI, Counterfactual Ret)	12.763	*2.34*	11.240	*2.09*
Mediated Path (through the disposition effect)				
I. P(Selling Date AQI, Path)	0.029	*2.62*	0.029	*2.62*
II. P(Path, Counterfactual Ret)	29.782	*9.17*	38.324	*10.92*
Indirect effect, P(Haze, Path)*P(Path, Counterfactual Ret)	0.866	*2.52*	1.115	*2.55*
Number of Observations			70295	
Adj. Goodness-to-fit Index			0.91	

Table IN3: Robustness Checks on the Huai River Policy

This table conducts various robustness checks for the RD test related to the Huai River Policy. Panel A provides the placebo test using two acritical lines that are 5° south and 5° north of the geographic location of the Huai River. Models (1), (2), (3), and (4) adopt the linear-OLS specification (used in Model (4) of Panel A in Table 3), linear-two-stage specification (used in Model (2) of Panel A in Table 4), quadratic-OLS specification (used in Model (4) of Panel B in Table 3), and quadratic-two-stage specification (used in Model (4) of Panel A in Table 4), respectively. Panel B resets the bandwidth of the RD regression to be 6° and 8°. Panel C applies the RD test to a subsample of investors: migrant investors who were born in cities in southern China trading in both the southern and northern banks of the Huai River. Panel D further controls for sunshine (average sunshine dates of a city in one year) and excludes important political event-dates or tier-one cities from our test. In Panel E, we apply the RD test to trading volume and the fraction of selling orders. Robust t-statistics are reported in parentheses and are based on standard errors clustered by city-year. Superscripts of *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

Panel A: Placebo Test Based on Aritificial Huai River Lines				
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	Linear Specification		Quadratic Specification	
A1: Aritificial Line 5 degrees South of the River				
D(North)	-0.098		0.030	
	(-1.19)		(0.54)	
AQI_hat		0.051		-0.005
		(0.51)		(-0.59)
Control	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	574	574	574	574
A2: Aritificial Line 5 degrees North of the River				
D(North)	-0.074		0.006	
	(-0.61)		(0.65)	
AQI_hat		-0.069		0.013
		(-0.61)		(0.67)
Control	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	665	665	665	665

Panel B: Alternative Specifications on Bandwith				
	Bandwidth = 8		Bandwidth = 6	
	OLS	IV	OLS	IV
B1: Linear Specifications				
D(North)	0.194***		0.259**	
	(3.55)		(2.68)	
AQI_hat		0.018***		0.030***
		(2.89)		(2.66)
Control	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	603	603	544	544
B2: Quadratic Specifications				
D(North)	0.199***		0.245***	
	(3.73)		(2.93)	
AQI_hat		0.021***		0.030***
		(3.01)		(2.90)
Control	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	603	603	544	544
Panel C: Subsample Tests based on Migrant Investors Born in Southern China				
	OLS	IV	OLS	IV
	Linear Specification		Quadratic Specification	
D(North)	0.268***		0.297***	
	(2.70)		(2.67)	
AQI_hat		0.063**		0.039**
		(2.03)		(1.99)
Control	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	642	642	642	642
Panel D: Controlling for Sunshine and Excluding Big Political Events and Large Cities				
	OLS	IV	OLS	IV
	Control for Sunshine		Excluding Big Events and Cities	
D(North)	0.185***		0.184***	
	(3.81)		(3.61)	
AQI_hat		0.020**		0.022*
		(1.99)		(1.88)
Control	Yes	Yes	Yes	Yes
Additional Control for Sunshine (Annual Sverage)	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	709	709	676	676
Panel E: Analysis on Trading Volume				
	OLS	IV	OLS	IV
	Log(Total Trading Volume)		Fraction of Sale	
D(North)	0.614		0.036	
	(1.52)		(1.36)	
AQI_hat		0.077		0.004
		(1.36)		(1.05)
Control	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
No. of Obs	709	709	709	709

Table IN4: Robustness Checks for the Cox Proportional Hazards Model

This table conducts various robustness checks for the results based on the Cox proportional hazards model as reported in Panel B of Table 9. Panel A further controls for investor characteristics. In Panel B, we first adopt an alternative time window for the classification of news dates (from the announcement day to five days later) and then apply the original subsample analysis to selling dates with different types of news. In particular, Models (3) and (4) are estimated when selling dates are associated with the announcement of quarterly report and dividend policies (still from the announcement day to three days later) or with other types of announcements (such as the turnover of managers and changes in investment policies related to management fees, front-load, and redemptions, etc.), respectively. All models include quarter fixed-effects. Robust t-statistics are reported in parentheses and are based on standard errors clustered by account and sale date. Superscripts of *, **, and *** indicate significance levels of 10%, 5%, and 1%, respectively.

Panel A: Cox Proportional Hazards Models with Additional Controls for Investor Characteristics

	(1)	(2)	(3)	(4)
	Full Sample		Subsamples of Selling Dates with/without News	
			No-News Dates	News Dates (Day 0-3)
Gain	0.312***	0.178***	0.230***	-0.135
	(56.62)	(3.53)	(3.97)	(-1.44)
Gain*Log(AQI)		0.032***	0.063***	0.018
		(2.67)	(4.60)	(0.80)
Log(AQI)		-0.102***	-0.133***	-0.065***
		(-9.55)	(-10.75)	(-3.44)
Female	-0.003	-0.003	-0.005	0.002
	(-0.70)	(-0.73)	(-0.96)	(0.21)
Age	-0.014***	-0.014***	-0.013***	-0.015***
	(-78.65)	(-78.32)	(-64.94)	(-43.01)
Migrant	-0.004	-0.013	-0.054***	0.058**
	(-0.30)	(-0.93)	(-3.19)	(2.38)
Education_High	0.053***	0.069***	0.075***	0.033***
	(8.63)	(10.90)	(9.74)	(2.95)
Experienced	-0.308***	-0.321***	-0.300***	-0.393***
	(-53.19)	(-54.61)	(-40.10)	(-38.27)
Log_GDP	-0.109***	-0.113***	-0.122***	-0.016
	(-17.08)	(-17.77)	(-15.54)	(-1.42)
Log_pop	0.103***	0.118***	0.138***	-0.012
	(19.86)	(22.03)	(21.59)	(-1.16)
Log_num_domestic_firm	0.070***	0.068***	0.058***	0.058***
	(16.82)	(16.20)	(11.43)	(7.54)
Log_gov_income	0.005	0.005	0.003	0.007
	(1.23)	(1.28)	(0.51)	(0.90)
Quarter FE	Y	Y	Y	Y
N	252,541	252,541	197,549	54,992

Panel B: Cox Proportional Hazards Models on Different News Types and News Horizon

	(1)	(2)	(3)	(4)
	Subsample Analysis by Different News Horizon		News Dates by News Types (0-3)	
	No-News Dates	News Dates (Day 0-5)	Quarterly reports/Dividends	Others
Gain	0.164***	-0.086	0.260**	-0.742***
	(2.74)	(-1.06)	(2.25)	(-4.31)
Gain*Log(AQI)	0.083***	0.020	0.004	0.026
	(5.88)	(1.05)	(0.13)	(0.66)
Log(AQI)	-0.109***	-0.064***	0.014	-0.043
	(-8.57)	(-3.94)	(0.58)	(-1.22)
Log_GDP	-0.103***	0.011	0.015	-0.052***
	(-12.29)	(1.12)	(1.00)	(-2.64)
Log_pop	0.067***	-0.081***	-0.058***	-0.108***
	(10.72)	(-9.29)	(-4.47)	(-5.71)
Log_num_domestic_firm	0.030***	0.015**	0.026***	0.070***
	(6.01)	(2.36)	(2.69)	(5.64)
Log_gov_income	0.021***	0.060***	0.030***	0.100***
	(3.81)	(9.04)	(3.05)	(7.99)
Quarter FE	Y	Y	Y	Y
N	180,540	72,001	38,297	16,695

Table IN5: Propensity to sell and Hypothetical returns

This table presents three sets of robustness checks to Table 9. First, as a robustness check to the Logit specification used in Panel A of Table 11, $D_{i,t} = \beta_1 \times MOM_{-t} + \beta_2 \times MOM_{-t} \times ln(AQI_{i,t}) + \beta_3 \times ln(AQI_{i,t})$, we adopt a panel specification with fund- and time-fixed effects to control for fund characteristics and trends. The results are tabulated in Panel A of this table. Next, in conducting the counterfactual analysis, $Ret_{i,t+1\sim t+20} = \beta_1 \times MOM_{-t} + \beta_2 \times MOM_{-t} \times ln(AQI_t) + \beta_3 \times ln(AQI_t)$, Panels B1 and B2 of Table 11 focus on the actual sales of investors occurring during the first 10 days of a calendar month and the first 10 days of any post-announcement period, respectively. We present a robustness check in this table to expand the sale dates to include all dates for the momentum-related test (Panel B1) and the first 20 days of any post-announcement period (Panel B2). Finally, Panels C1 and C2 examine whether time-series calendar momentum and post-announcement price drift exist in the underline assets of Table 11 (i.e., funds). More explicitly, Models (1) and (2) regress calendar return that can be generated by funds on their previous month return, when lagged or concurrent market return are controlled. Models (3) and (4) regress the cumulative return that can be generated during the first and second 10-working-day post-announcement periods on announcement return, when the market returns in both the announcement period and the post-announcement holding period are controlled. Superscripts of *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

A. Robustness Checks on Selling decision of investors vs. momentum (OLS with f.e.)

A1. Selling decision vs. past-month momentum	Selling in the first 10 days of each calendar month		Selling in all dates	
	(1)	(2)	(3)	(4)
Mom (last month return)	0.097***	-0.223**	0.291***	-0.065
	(5.67)	(-2.29)	(24.08)	(-1.05)
Mom*AQI		0.075***		0.083***
		(3.35)		(5.81)
AQI		0.003		-0.005***
		(1.64)		(-4.25)
Lagged Market Ret	YES	YES	YES	YES
Lagged Market*AQI	NO	YES	NO	YES
Fund FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Observations	55,115,198	55,115,198	112580000	112580000
R-squared	0.005	0.005	0.004	0.004

A2. Selling decision vs. announcement momentum	Selling in the first 10 days after the announcement period		Selling in 11-20D after the announcement period	
	(5)	(6)	(7)	(8)
Announcement Period Ret	7.134***	-6.749***	6.105***	4.840***
	(35.98)	(-7.00)	(32.54)	(4.96)
Announcement Period Ret*AQI		3.317***		0.305
		(14.72)		(1.32)
AQI		-0.007**		-0.013***
		(-1.97)		(-3.69)
Ann-Period Market Ret	YES	YES	YES	YES
Ann-Period Market Ret*AQI	NO	YES	NO	YES
Fund FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Observations	14367157	14367157	14100073	14100073
R-squared	0.004	0.004	0.003	0.004

B. Robustness Checks on Counterfactural Returns based on all MOM selling or post-announcement selling in 20 days

B1. Counterfactural 20D-return (%) vs. MOM (if investors do not sell in all days)

	(1)	(2)	(3)
Selling-Day AQI	1.280***		1.054***
	(26.37)		(22.64)
Mom (Prior to Selling)		3.892***	-27.894***
		(12.67)	(-9.49)
Mom*AQI			7.407***
			(10.89)
Constant	-0.608**	4.876***	0.063
	(-2.43)	(39.25)	(0.26)
Lagged Market Ret	YES	YES	YES
Lagged Market*AQI	NO	NO	YES
Fund FE	YES	YES	YES
Observations	380,262	380,262	380,262
R-squared	0.031	0.029	0.039

B2. Counterfactural 20D-return (%) vs. Announcement MOM (if investors do not sell in the first 20D immediately after the announcement period)

	(1)	(2)	(3)
Selling-Day AQI	2.089***		1.286***
	(25.57)		(17.87)
Ann-Period Ret (Prior to Selling)		58.444***	-62.192***
		(23.40)	(-5.40)
Ann-Period Ret*AQI			26.445***
			(10.15)
Constant	-5.986***	2.506***	-3.024***
	(-16.97)	(71.06)	(-9.89)
Ann-Period Market Ret	YES	YES	YES
Ann-Period Market *AQI	NO	NO	YES
Fund FE	YES	YES	YES
Observations	87,295	87,295	87,295
R-squared	0.200	0.199	0.205

C. Time-series Momentum and Post-announcement Price Drift of the Underline Funds

C1. Time-series Momentum of Funds (Calendar month return regressed on past-month return)

	(1)	(2)
Mom (last mon ret)	0.097**	0.100**
	(2.09)	(2.14)
Constant	0.011***	0.011***
	(2.98)	(3.03)
Lagged Market	YES	YES
Concurrent Market	NO	YES
Observations	465	465
R-squared	0.014	0.015

C2. Post-announcement Return Drift of Funds (Cumulative Return of first/second 10-day period after announcement vs. announcement period return)

	Ret (1-10D) (3)	Ret (11-20D) (4)
Announcement Period Ret	0.151**	0.107**
	(2.24)	(2.13)
Constant	0.004*	0.006***
	(1.69)	(3.19)
Ann-Period Market Ret	YES	YES
Testing Period Market Ret	YES	YES
Observations	118	118
R-squared	0.694	0.858

Table IN6: Robustness checks on sign realization preference and the influence of AQI

This table provides robustness checks to the regression discontinuity approach on sign realization preference and the magnitude realization test as reported in Table 10. In Panel A1, returns since purchase are restricted in a small region within 0.1 standard deviations around zero, and we expand from 3rd-degree polynomials to also include 4th-degree and 5th-degree polynomials in regression. Other specifications are the same as Table 2 of Ben-David and Hirshleifer (2012). In Panel A2, we further examine the influence of air pollution by interacting it with the dummy variable indicating the positive sign of returns. In Panel B, we expand the specifications Table 12 (Panel B) to include interaction between air pollution and the sign dummy variable of returns. Robust *t*-statistics are reported in parentheses and are based on standard errors clustered by investor. Superscripts of *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

Panel A1. Discontinuity Analysis on sign realization preference (Dependent variable = I{Sell} x 100; range = 0.1 stdev around zero)

	1 to 20			21 to 250			>250		
VARIABLES	3rd	4th	5th	3rd	4th	5th	3rd	4th	5th
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
I(ret>0)	0.333***	0.341***	-0.156*	-0.017	-0.042	0.035	-0.046**	-0.063**	-0.067**
	(4.21)	(4.18)	(-1.96)	(-0.56)	(-1.33)	(1.12)	(-2.02)	(-2.42)	(-2.45)
I(ret=0)	-0.284***	0.069	-0.697***	-0.229***	-0.213***	-0.171***	-0.127***	-0.125***	-0.121***
	(-5.55)	(1.30)	(-13.37)	(-10.90)	(-9.54)	(-7.76)	(-7.24)	(-6.21)	(-5.80)
Sqrt(Time)	-0.043***	-0.028***	-0.014**	-0.024***	-0.014***	-0.015***	0.002***	0.003***	0.003***
	(-5.47)	(-4.67)	(-2.33)	(-19.42)	(-11.40)	(-11.93)	(4.26)	(4.72)	(4.64)
Polynomials	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomials with sqrt(time)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomials with positve and negative indicator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	963,721	963,721	963,721	1,854,455	1,854,455	1,854,455	1,366,419	1,366,419	1,366,419
R-squared	0.012	0.012	0.013	0.001	0.002	0.002	0.000	0.000	0.000

Panel A2. Sign realization preference with air pollution

	1 to 20			21 to 250			>250		
VARIABLES	3rd	4th	5th	3rd	4th	5th	3rd	4th	5th
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
I(ret>0)	0.775**	0.857**	0.301	0.004	-0.007	0.076	0.018	0.003	-0.002
	(2.21)	(2.41)	(0.83)	(0.05)	(-0.08)	(0.83)	(0.31)	(0.05)	(-0.04)
I(ret=0)	-2.526***	-2.070***	-2.900***	-0.228***	-0.226***	-0.179***	-0.022	-0.019	-0.017
	(-10.00)	(-7.96)	(-11.07)	(-3.51)	(-3.47)	(-2.76)	(-0.51)	(-0.43)	(-0.37)
I(ret>0)*Logaqi	-0.095	-0.117	-0.103	-0.005	-0.008	-0.009	-0.015	-0.015	-0.015
	(-1.19)	(-1.46)	(-1.28)	(-0.24)	(-0.40)	(-0.48)	(-1.20)	(-1.24)	(-1.21)
I(ret=0)*Logaqi	0.525***	0.495***	0.508***	-0.000	0.003	0.002	-0.024***	-0.025***	-0.024***
	(8.93)	(8.39)	(8.61)	(-0.02)	(0.21)	(0.14)	(-2.61)	(-2.64)	(-2.63)
Logaqi	-0.515***	-0.487***	-0.503***	-0.004	-0.004	-0.003	0.027***	0.027***	0.027***
	(-8.80)	(-8.28)	(-8.55)	(-0.31)	(-0.31)	(-0.25)	(3.04)	(3.04)	(3.03)
Sqrt(Time)	-0.039***	-0.026***	-0.011*	-0.024***	-0.014***	-0.015***	0.002***	0.002***	0.002***
	(-4.98)	(-4.26)	(-1.83)	(-19.46)	(-11.42)	(-11.95)	(4.11)	(4.58)	(4.50)
Polynomials	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomials with sqrt(time)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomials with positve and negative indicator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	963,721	963,721	963,721	1,854,455	1,854,455	1,854,455	1,366,419	1,366,419	1,366,419
R-squared	0.012	0.013	0.014	0.001	0.002	0.002	0.000	0.000	0.000

Panel B: The Ben-David and Hirshleifer (2012) Magnitude Test (Dependent variable = I{Sell} x 100) Full Specification									
	Short term periods (1 to 20 days)			Mid-term periods (21 to 250 days)			Longer periods (> 250 days)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ret+	5.108***	5.242***	-4.237***	4.749***	8.053***	5.395***	2.044***	3.294***	0.890
	(58.52)	(15.38)	(-4.80)	(86.12)	(48.36)	(10.69)	(21.97)	(8.59)	(1.07)
Ret-	1.616***	-0.619	1.428	-0.376***	-0.823***	3.269***	-0.596***	-2.904***	-3.710***
	(8.96)	(-0.96)	(0.86)	(-5.21)	(-3.53)	(4.65)	(-5.92)	(-7.81)	(-4.13)
I(ret>0)	0.090***	-0.012	0.009	0.236***	0.097***	-0.041	0.050***	0.265***	0.472***
	(7.15)	(-0.52)	(0.14)	(30.83)	(6.77)	(-0.99)	(3.80)	(8.29)	(6.78)
I(ret=0)	-0.676***	-0.909***	-1.462***	-0.812***	-1.461***	-2.304***	-0.568***	0.028	-0.537
	(-26.13)	(-15.60)	(-5.24)	(-23.45)	(-10.63)	(-6.58)	(-8.82)	(0.07)	(-0.77)
Ret+*Logaqi			2.221***			0.605***			0.594***
			(11.74)			(5.53)			(3.24)
Ret-*Logaqi			-0.511			-0.926***			0.199
			(-1.36)			(-6.14)			(0.95)
I(ret>0)*Logaqi			-0.005			0.031***			-0.052***
			(-0.31)			(3.50)			(-3.35)
I(ret=0)*Logaqi			0.130**			0.189**			0.132
			(2.11)			(2.56)			(1.37)
Logaqi			-0.102***			-0.051***			0.034***
			(-7.56)			(-7.19)			(2.82)
Ret+*sqrt(Time)		-0.093	-0.129		-0.362***	-0.357***		-0.044***	-0.049***
		(-0.95)	(-1.31)		(-20.66)	(-20.38)		(-3.06)	(-3.36)
Ret-*sqrt(Time)		0.588***	0.635***		0.045*	0.030		0.085***	0.083***
		(2.90)	(3.16)		(1.95)	(1.29)		(6.86)	(6.59)
I(ret>0)*sqrt(Time)		0.037***	0.035***		0.015***	0.016***		-0.008***	-0.007***
		(5.51)	(5.32)		(11.12)	(11.60)		(-6.84)	(-6.26)
I(ret=0)*sqrt(Time)		0.088***	0.085***		0.068***	0.070***		-0.024	-0.025
		(4.80)	(4.62)		(5.52)	(5.73)		(-1.30)	(-1.38)
sqrt(Time)	-0.080***	-0.099***	-0.096***	-0.037***	-0.036***	-0.037***	-0.003***	0.004***	0.004***
	(-31.24)	(-16.38)	(-15.86)	(-81.83)	(-32.78)	(-33.44)	(-8.87)	(5.57)	(4.97)
log(price)	-0.565***	-0.567***	-0.550***	-0.143***	-0.143***	-0.141***	0.335***	0.334***	0.331***
	(-52.36)	(-52.66)	(-51.23)	(-24.52)	(-24.53)	(-24.04)	(27.25)	(27.08)	(26.59)
volatitlity+	-5.119***	-4.964***	-4.996***	-10.259***	-10.238***	-10.250***	-12.775***	-13.378***	-13.280***
	(-14.71)	(-14.18)	(-14.28)	(-38.50)	(-38.28)	(-38.30)	(-24.81)	(-25.42)	(-25.27)
volatitlity-	-3.896***	-4.326***	-4.466***	-0.635*	-0.712**	-0.723**	-13.858***	-13.456***	-13.407***
	(-6.27)	(-6.93)	(-7.12)	(-1.89)	(-2.11)	(-2.13)	(-20.69)	(-20.16)	(-20.06)
Constant	-1.914***	-1.860***	-1.431***	-2.301***	-2.305***	-2.077***	-2.830***	-3.023***	-3.156***
	(-122.79)	(-80.26)	(-23.01)	(-304.61)	(-194.30)	(-61.78)	(-194.55)	(-126.21)	(-59.84)
Observations	4,357,608	4,357,608	4,357,608	16,326,851	16,326,851	16,326,851	20,158,791	20,158,791	20,158,791
Pseudo R2	0.0321	0.0324	0.0330	0.0317	0.0321	0.0322	0.00948	0.00997	0.0100