Review of Economic Studies (2018) **85**, 1609–1634 doi:10.1093/restud/rdy016

Advance access publication 29 March 2018

Something in the Air: Pollution and the Demand for Health Insurance

TOM Y. CHANG

University of Southern California

WEI HUANG

University of International Business and Economics

and

YONGXIANG WANG

University of Southern California and Shanghai Jiao Tong University

First version received February 2017*; Editorial decision May* 2017*; Accepted February* 2018 (*Eds.*)

We find that daily air pollution levels have a significant effect on the decision to purchase or cancel health insurance in a manner inconsistent with rational choice theory. A one standard deviation increase in daily air pollution leads to a 7.2% increase in the number of insurance contracts sold that day. Conditional on purchase, a one standard deviation decrease in air pollution during the cooling-off (*i.e.* cost-free cancellation) period relative to the order-date level increases the return probability by 4.0%. We explore a range of potential mechanism and find the most support for projection bias and salience.

Key words: Projection bias, Salience.

JEL Codes: D12, D91

1. INTRODUCTION

The important decisions that people make have lasting consequences. As such, they require people to predict the utility they will receive in the future from decisions they make today. Although standard economic theory assumes that individuals can accurately make such predictions, evidence from psychology and behavioural economics suggests that people exhibit systematic biases in predicting future utility (see DellaVigna (2009) for a review). One such bias, captured in such clichés as "sleep on it" or "never go grocery shopping on an empty stomach" is that current conditions have an oversized influence on intertemporal decision making.[1]

1. Empathy gaps (Loewenstein, 2005; Ariely and Loewenstein, 2006), projection bias (Loewenstein *et al.*, 2003), salience (Bordalo *et al.*, 2013, 2014; Koszegi and Szeidl, 2013), attribution bias (Haggag and Pope, 2016), and present bias (Laibson, 1997; O'Donoghue, and Rabin, 1999, 2015) are examples of mechanisms for why such might be the case.

The editor in charge of this paper was Nicola Gennaioli.

In this article, we use transaction-level data from a large Chinese insurance company to examine the role that idiosyncratic variation in daily air pollution plays in an individual's decision to purchase or cancel health insurance. Although air pollution, which has an immediate and deleterious effect on one's health, is subject to high day-to-day variability, it is essentially stationary during our study period. Since the health insurance policies we examine do not cover pre-existing conditions and have a 180-day waiting period before coverage begins, the value of the policy is a function of the premiums and the probability of illness in the future. And since premiums vary infrequently and are uniform across cities, daily air pollution levels should be a non-factor in a rational person's decision to purchase or cancel health insurance.

We instead find that both the purchase and cancellation of health insurance policies are significantly influenced by idiosyncratic variation in daily air pollution levels. Specifically, we find that, when air pollution is high, individuals are more likely to purchase insurance contracts. In addition, insurance contracts are more likely to be canceled if air pollution improves during the government-mandated 10-day "regret period", during which individuals can, without cost, cancel their insurance contracts. This cancellation effect is negatively related to air pollution during the cooling-off period (CoP) and is driven by the change in air pollution *relative* to the level at the time of purchase. That is, individuals are more likely to buy insurance when pollution is high and more likely to cancel it if air pollution levels are lower during the CoP relative to the date of purchase.

Controlling for seasonal and regional variation in sales patterns, we find that a one-standard deviation increase in the daily level of PM2.5 in a city, as measured by the Air Quality Index (AQI), leads to a 7.2% increase in the number of insurance contracts sold in that city that day. This effect of pollution on sales is non-linear, with measurable effects occurring only at AQI levels associated with adverse health effects. We also find that a one-standard deviation decrease in the AQI during the CoP *relative* to the order date leads to a 4% increase in the share of insurance contract that are canceled. In contrast, AQI levels have no impact on either sales or cancellations of other insurance products that the company sells. In addition, the results of a distributed lag model show that pollution affects the aggregate level of insurance contracts sold, rather than changing the timing of insurance purchase.

We explore a range of explanations for our results. We first consider a range of rational explanations, including learning broadly defined, and find them unable to fully explain our empirical results. We next consider a range of alternative psychological mechanisms, and find the greatest support for two mechanisms that are essentially indistinguishable in our setting: projection bias and salience.

Our results are important for several reasons. First, the setting we study allows us to differentiate between a wide range of possible explanations for an important and long-run decision in an unusually clean manner. Specifically, the 180-day waiting period before the insurance benefits start helps to rule out several potential confounds, and the nature of health insurance sales by this firm makes it very unlikely that there was a supply-side response to idiosyncratic fluctuations in air pollution levels.[2] In addition, the existence of a relatively short CoP allows us to test not only the relationship between air pollution and the decision to purchase insurance, but also how changes in air pollution affect the decision to cancel it.

Our article is most similar to Conlin *et al.* (2007), Simonsohn (2010), and Busse *et al.* (2015) who show, respectively, that idiosyncratic variation in weather affects the return probability of cold-weather items, college enrollment, and the type of automobile purchased. We complement this literature by presenting direct evidence of impact of cancellation-date conditions on

2. In conversation with the firm's senior marketing manager we were told that they had not considered that air pollution might have an effect on in insurance sales, and that they do not engage in high-frequency marketing efforts.

cancellations, and by providing evidence that consumers are overly influenced by something other than weather.

Second, insurance is one of the world's largest industries, eclipsed only by real estate, finance, and government services. In 2014, insurance premiums in the U.S. exceeded $2 trillion, with $839 billion attributed to health insurance premiums in the private insurance market alone. Healthcare spending is also a huge part of the economy, accounting for 5.7% of China's GDP and 9% of the GDP for all Organization for Economic Co-operation and Development (OECD) countries.[3] Further, given the ongoing debate regarding health insurance coverage, understanding how individuals make insurance decisions has important implications for generating effective policy in this domain. Our results provide strong empirical evidence on the importance of psychological effects in the market for health insurance in China and, potentially, insurance markets more widely.

Our results also document an unanticipated consequence of rising air pollution levels in the developing world. This finding contributes to a small but rapidly growing literature documenting the impact of air pollution on non-health outcomes: labour productivity (Graff Zivin and Neidell, 2012; Hanna and Oliva, 2015; Chang *et al.*, 2014, 2016), student test scores (Lavy *et al.*, 2014), and crime (Herrnstadt and Muehlegger, 2015).

Finally, our results provides evidence in support of the hypothesis put forth in Loewenstein *et al.* (2003) that "cooling-off laws" might be effective "as devices for combating the effects of projection bias". Indeed, our results suggest that the efficacy of CoPs in attenuating the effects of intertemporal behavioural biases more generally are determined, in part, by autocorrelation in the driving state variable (*i.e.* if the intertemporal behavioural bias is caused by a slowly changing state variables and has negative welfare consequences, consumers might benefit from longer CoPs.)[4]

The article proceeds as follows. The subsequent section we review some basic information on the relationship between air pollution and health along with our empirical strategy. Section 3 describes the data used in the article. Section 4 presents our empirical results on the effect of daily air pollution levels on the decision to purchase and cancel insurance contracts. Section 5 explores several potential mechanisms for our empirical findings. Section 6 concludes.

2. AIR POLLUTION AND EMPIRICAL STRATEGY

Our empirical strategy exploits the relationship between air pollution and human health. A large body of toxicological and epidemiological evidence suggests that exposure to air pollution harms health (see EPA (2004)). The health risks related to exposure to air pollution arise primarily from changes in pulmonary and cardiovascular functioning (Seaton *et al.*, 1995) and can manifest in respiratory episodes, such as asthma attacks, and cardiovascular events, such as heart attacks (Dockery and Pope, 1994; Pope, 2000). Exposure to air pollution also leads to more subtle effects, such as changes in blood pressure, irritation in the ear, nose, throat, and lungs, and mild headaches (Ghio *et al.*, 2000; Pope, 2000; Auchincloss *et al.*, 2008). Figure 1 presents the air pollution levels, as expressed in AQI levels, and the relevant heath effects, as per the U.S. Environmental Protection Agency (EPA). Some of symptoms associated with high levels of air pollution are experienced immediately (*e.g.* watery eyes, scratchy throat, shortness of breath), while others can arise within a few hours after exposure. This immediacy of this physical response is important for our empirical design which exploits high-frequency variation in air pollution levels.

3. http://www.oecd.org/els/health-systems/health-data.htm.

4. Importantly, since we do not know whether individuals were better off with or without insurance in our setting, we cannot say whether the effect we document increases or decreases welfare.

Air Quality Index (AQI)	PM2.5 Health Effects Statement	PM 2.5 Cautionary Statement
Good (0-50)	PM2.5 air pollution poses little or no risk.	None
Moderate (51-100)	Unusually sensitive individuals may experience respiratory symptoms.	Unusually sensitive people should consider reducing prolonged or heavy exertion.
Unhealthy for Sensitive Groups (101-150)	Increasing likelihood of respiratory symptoms in sensitive individuals, aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly.	People with heart or lung disease, older adults, and children should reduce prolonged or heavy exertion.
Unhealthy (151-200)	Increased aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly; increased respiratory effects in general population.	People with heart or lung disease, older adults, and children should avoid prolonged or heavy exertion; everyone else should reduce prolonged or heavy exertion.
Very Unhealthy (201-300)	Significant aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly; significant increase in respiratory effects in general population.	People with heart or lung disease, older adults, and children should avoid all physical activity outdoors. Everyone else should avoid prolonged or heavy exertion.
Hazardous (301-500)	Serious aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly; serious risk of respiratory effects in general population.	Everyone should avoid all physical activity outdoors; people with heart or lung disease, older adults, and children should remain indoors and keep activity levels low.

FIGURE 1

U.S. EPA guide to AQI

Our empirical strategy is to use idiosyncratic variation in daily air pollution levels in a city as a health shock to the city's population. Specifically, we assume that AQI levels are negatively related to the contemporaneous aggregate health of the population of that city (*e.g.* the higher the air pollution level, the sicker the population), such that a city's daily AQI level serves as a proxy for the contemporaneous health of its citizens.[5]

Although day-to-day variation in AQI levels is quite high,[6] AQI levels generally follow a cyclical pattern and are correlated with other environmental factors more generally (*e.g.* weather). In Beijing, for example, AQI tends to be lower during the rainy season, when precipitation serves to wash away airborne pollutants, and higher in winter months, when people burn more fossil

5. The relationship between air pollution and health is well understood in China. Indeed, air pollution trails only corruption as the biggest concern among the Chinese public (Pew Research Center, 2015).

6. The within-city day-to-day correlation in AQI levels is less than 0.5.

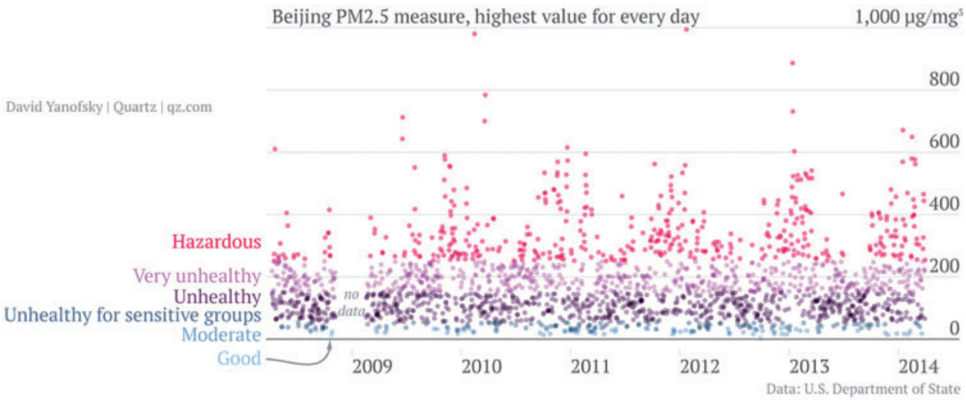

FIGURE 2

Daily PM2.5 levels as measured by the U.S. Embassy in Beijing

fuels for warmth (Figure 2). Although, similar to current temperature, the current AQI level provides some additional information regarding the AQI levels one can expect in the near future, it provides essentially no additional information about the AQI levels one should expect 180 days from now (*e.g.* the waiting period for the health insurance policies that we examine).

3. DATA

The data were obtained from four sources: a large Chinese company that sells a variety of insurance products, the U.S. State Department, and 15 Tianqi, a Chinese weather website, and an online survey. We collected detailed information on over one million insurance contracts. These contracts represent the universe of health insurance policies and include a subset of other insurance products sold by the firm to residents in a small number ($n<5$) of large Chinese cities from 2012 through 2015.[7] For each insurance policy sold, we have the date of purchase, city of residence, contract length, whether the insurance is for the purchaser or for someone else (*e.g.* a family member), and some basic demographic information for the person covered by the insurance policy. We also have cancellation information for policies sold through the end of 2014.

Providing near-universal health insurance coverage has been a major goal of the Chinese government, and recent reforms have brought them close to this goal. As of 2009, approximately 90% of the population has health insurance through the government. This coverage is accomplished through three insurance programmes: the Urban Employee Basic Medical Insurance (UEBMI), the Urban Resident Basic Medical Insurance (URBMI), and the New Rural Cooperative Medical System (NRCMS). The benefit level of the insurance provided through these programmes, however, is quite low, both in terms of the share of expenses covered and the cap on total lifetime covered expenses. As such, the market for secondary private health insurance to help cover this gap is growing quickly, especially among China's burgeoning middle class. Because of the low cap on lifetime expenses, supplemental health insurance is considered especially important in covering expenses due to significant adverse health events, such as cancer.

7. Due to the sensitive nature of the sales data, we cannot reveal the identities of the cities in our sample or provide disaggregated statistics on sales patterns for the various insurance products in our sample.

The health insurance contracts in our data consist of this type of supplemental private health insurance.

For policies provided by the firm, there is a 180-day waiting period between the date of purchase and the effective start date of insurance coverage. In addition, there is a pre-existing condition clause that prevents the covered individual from receiving benefits if his or her illness is the result of a condition that existed before the date of purchase. Finally, these insurance contracts are subject to a law that requires a 10-day "regret period", during which consumers can cancel their insurance contracts without any penalty.

From the U.S. State Department, we have hourly measures of PM2.5, collected by air quality monitors located on U.S. Embassy compounds in the relevant cities.[8] The PM2.5 level is expressed in terms of an AQI, following the U.S. EPA formula (EPA, 2004). The AQI values are designed to help inform health-related decisions by mapping pollution levels to round-number breakpoints that correspond to categories of health impact (Figure 1). While we cannot provide full details of the pollution levels, as they could be used to determine the identity of the cities in our sample, we can state that the mean daily AQI in our composite sample is 125.6, with a standard deviation of 98.4. Although this level of air pollution is typical for a large Chinese city, it would be considered quite high in the U.S.[9]

Weather information was retrieved for each city in our sample from 15 Tianqi. These data included daily low and high temperatures, precipitation, and a dummy variable for snowfall. After merging the weather data with the AQI and order information by city and date, we dropped observations for city and date combinations for which AQI information was unavailable or appeared unreliable.[10] As shown in Table 1, this left us with a sample of 579,303 health insurance contracts sold across 2,577 city*days, with an average of 224.8 sold in each city each day. The mean contract in our sample is for a period of 31.6 years.[11] Approximately half of the time, an individual purchases insurance for him or herself. Otherwise, an individual purchases insurance for a family member (generally, a spouse or child). The average age of the covered individual is 25.4 years, and just over half of covered individuals are female. The cancellation rate during the 10-day government mandated CoP is 2.8%.

Finally, we use data from a short online survey designed to test whether daily variation in air pollution has an effect on beliefs about pollution levels in the future. In addition to asking for some basic demographic information, the survey asked if whether they thought that air pollution in their city would be better, the same, or worse in a year's time. The survey was conducted in the summer of 2016 on WeChat, the dominant social media platform in China. Messages were sent to all the members of five large, randomly selected WeChat groups. The message said that in exchange for completing a brief survey, they would receive a "WeChat Hongbao" or virtual "red envelop" which contained a random gift of between 1 and 25 RMB. This was a relatively generous Hongbao, and generated a response rate of over 70%.[12]

8. See, for example, http://beijing.usembassy-china.org.cn/aqirecent3.html for more details on the U.S. State Department's air quality monitoring programme in China.

9. Although the statistics are not comparable, as different technologies are used to measure air pollution at different temporal resolutions, as an illustrative example, the EPA reports that the median AQIs in Cambridge, Massachusetts, and Los Angeles, California, in 2015 were 46 and 77, respectively.

10. Three city by date observations were deemed unreliable: one observation had an AQI of zero while two observations had an AQI $>$ 800.

11. For the 25.3% of health insurance policies sold with what the firm refers to as "lifetime" contracts (*i.e.* policy period is for the life of the covered individual), the contract length was set to 85 years, the maximum length allowed for non-lifetime contracts. For the vast majority of contracts (over 98%), premiums are paid on a yearly basis.

12. Of the 641 surveys sent out, 461 were successfully completed.

TABLE 1
Summary statistics

	Mean	*Std. Dev.*	*Min.*	*Max.*	*Obs.*
Date characteristics					
AQI $PM_{2.5}$	125.6	98.4	0.04	731	2,577
Temperature	19.4	9.4	−6	39	2,577
Rain	0.049	0.215	0	1	2,577
Snow	0.014	0.119	0	1	2,577
Sales per day	224.8	509.1	0	9,313	2,577
Contract characteristics					
Contract length (Years)	54.7	31.6	1	85.0	579,303
Purchased for oneself	0.47	0.50	0	1	579,303
Age (Years)	25.4	15.5	0.79	66.1	579,303
Female	0.55	0.50	0	1	579,303
canceled	0.028	0.17	0	1	414,064

Notes: The demographic variables associated with the health insurance contracts are for the insured, and not the purchaser of insurance.

4. EMPIRICAL RESULTS

4.1. *Effect of air pollution on purchases*

Our base specification for estimating the impact of air pollution on the sales of insurance contracts is then given by

$$\text{Log}(\text{Insurance}_{jt}) = \beta \text{AQI}_{jt} + X_{jt}\gamma + D_{jt} + \epsilon_{jt}, \tag{1}$$

where Insurance_{jt} is the number of insurance purchased from the firm by the residents of city j on date t, AQI_{jt} is the high hourly AQI in city j over a two-day window that consists of date t and $t-1$. This allows for the purchase decision to have been made the day before purchase, which is possible because pollution tends to peak in the evening, when the firm is closed and unable to take customer orders.[13] The vector X_{jt} consists of a quadratic function of high temperature and dummy variables for precipitation and snowfall. D_{jt} are day-of-week, month-of-year by city, and year-by-city fixed effects, included to account for trends within the week and over time, respectively. Standard errors are clustered on city*date.

The main coefficient of interest is β, which captures the effect of air pollution on the demand for health insurance. The coefficient can be interpreted as the percentage change in the total number of insurance contract sold on a given day caused by a one-unit increase in the AQI.

The results of estimating equation (1) are presented in Table 2. Column 1 indicates that a one-unit increase in the daily AQI generates approximately a 0.072% increase in daily sales, or that a one-standard deviation increase in the daily AQI leads to a 7.2% increase in daily sales. For column 2, we allow AQI to have a non-linear effect on sales by re-estimating equation (1) with indicator variables that correspond to the different EPA categories for pollution levels in place of a linear measure of AQI (Figure 1). The withheld category is an AQI of between 0 and 50, which corresponds to "Good" air quality. The results indicate that the effect of the AQI on sales become significant only when the AQI is higher than 150, corresponding to the level deemed "Unhealthy" by the EPA; the coefficient for "Moderate" levels of PM2.5 is small and statistically insignificant, while the coefficient for the "Unhealthy for Sensitive Groups" level of PM2.5 is approximately two-thirds as large as the coefficient for "Unhealthy" but not statistically significant at the conventional levels. AQIs of 150 to 200 (Unhealthy), 200 to 300 (Very Unhealthy), and greater than 300 (Hazardous) are associated with statistically significant increases in daily sales

13. Using either the one-day *AQI* for date t or $t-1$ produces similar results.

TABLE 2
The effect of pollution on insurance sales

Dependent variable: Log(number of contracts sold)				
Insurance type		Health		Other
$AQI_{PM2.5}$	0.00072***		0.00067***	−0.00018
	(0.00018)		(0.00019)	(0.00022)
$AQI_{PM2.5}$ 50–100		0.0116		
		(0.0715)		
$AQI_{PM2.5}$ 100–150		0.1147		
		(0.0759)		
$AQI_{PM2.5}$ 150–200		0.1681**		
		(0.0800)		
$AQI_{PM2.5}$ 200–300		0.1680*		
		(0.0861)		
$AQI_{PM2.5}$ 300+		0.2340**		
		(0.0936)		
Other City $AQI_{PM2.5}$			0.00010	
			(0.00022)	
Temperature	−0.0191*	−0.0189*	−0.0283**	0.0221**
	(0.0111)	(0.0111)	(0.0116)	(0.0097)
Temperature^2	0.0008***	0.0008***	0.0010***	−0.0006**
	(0.0003)	(0.0003)	(0.0003)	(0.0002)
Rain	−0.0391	−0.0355	−0.0406	0.0980
	(0.0755)	(0.0756)	(0.0769)	(0.0660)
Snow	−0.2059	−0.1943	−0.2576	−0.2837
	(0.1700)	(0.1707)	(0.1852)	(0.1656)
Adjusted R^2	0.481	0.481	0.478	0.427
Observations	2,573	2,573	2,489	2,573

Notes: All columns present the results from ordinary least square regressions. For city j, "Other City $AQI_{PM2.5}$" is the $AQI_{PM2.5}$ of its nearest neighbour. Insurance type "Other" consists of non-health insurance policies. All regressions included dummy variables for day of week, city*month and city*year. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

of 16.8%, 16.8%, and 23.4%, respectively, compared to days with an AQI of less than or equal to 50. Taken together, these results indicate that the air pollution in one's immediate vicinity increases demand for health insurance, and that this increase in demand occurs only when air pollution has reached levels associated with noticeable and immediate health effects.

4.2. *Robustness*

A concern is that our results are affected by an unobservable that is correlated with both pollution and demand for insurance. While the highly localized, high-frequency nature of our observations make this less likely, we provide additional checks to address such concerns. First to the extent that air pollution is correlated across our cities, air pollution could be proxying for a unobserved regional or national factor. To address this concern, we first rerun the regression in column with an additional term that captures the pollution in the other cities in our sample. To do this, we first match each city to its closest neighbour, then regress that city's daily sales against both that city's pollution and the pollution of the matched city. The results of this regression are shown in column 3. We see that controlling for the AQI of the nearest city slightly reduces the size of the coefficient from 0.00072 to 0.00066 and that the coefficient for other city's AQI is both small and statistically insignificant. This results indicates that demand for health insurance is affected by only local, and not regional idiosyncratic variation in air pollution levels.

Next we rerun our main regression with the number of non-health insurance contracts sold by the company as the dependent variable. This category consists primarily of "endowment"

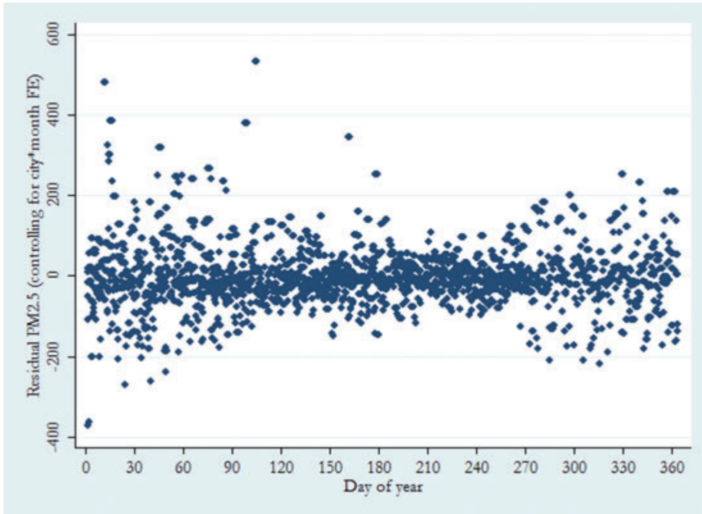

FIGURE 3
Daily residual PM2.5 levels after controlling for city*month fixed effects

insurance and personal accident insurance. Endowment insurance combines a term-life policy with an annuity, and constitutes the majority of the non-health insurance policies in our sample.

The result of this analysis is presented in column 4. Here, in contrast to column 1, the coefficient of interest is small and statistically insignificant, indicating both that air pollution is not a significant driver of demand for other insurance products sold by the firm and that air pollution is not serving as a proxy for an unobserved variable that drives demand for insurance products, more generally. While it seems plausible that idiosyncratic variation in current health could affect the demand for other insurance products, specifically life insurance products, the fact that we find no effect of contemporaneous pollution on demand suggests this is not the case. Given a 95% confidence interval of $[-0.00058, 0.00032]$ though, we cannot entirely rule out pollution having a meaningful effect on the demand for other insurance products, only that the effect of air pollution on other forms of insurance is substantially less that what we find for health insurance. This indicates that for an unobservable factor to be driving our results it must not only be geographically and temporally localized, but also very limited in scope affecting only the demand for health insurance contracts.

4.2.1. Seasonal patterns. Another potential concern is that the included time controls do not adequately account for seasonal variation.[14] We attempt to address this concern in two ways. First we examine the residual air pollution level after controlling for city by month-of-year fixed effects. As shown in Figure 3, while the dispersion in pollution appears higher in the winter versus summer months, the pollution level does not exhibit any clear visual pattern indicating an uncontrolled for seasonal pattern.

Second, we repeat our main analysis seasonal controls both finer and coarser than our main specification, with and without weather controls. Specifically we rerun the regression with month, week-of-year, and day-of-year fixed effects with and without city level interactions. While we would ideally like to also rerun the main regression controlling for the average air pollution level

14. See http://datacolada.org/46 for a detailed discussion of this issue.

for each day of the year in each city, the lack of historic pollution data makes this approach infeasible. The results of these regression are presented in Appendix Table A1. Across all specifications, we find a strong and statistically significant relationship between pollution and the demand for health insurance. Together these results suggest that our results are not driven by a failure to adequately control for seasonal patterns.

4.3. *Effect of PM2.5 on cancellations*

We next examine the effect of air pollution on insurance cancellations. For this analysis, we start with the base regression specification

$$\text{Cancel}_{ijt} = f(\text{AQI}_{ijt},...,\text{AQI}_{ij,t+11})\beta + C_i b + X_{jt}\gamma + D_{jt} + \epsilon_{jt}, \tag{2}$$

where Cancel_{ijt} is a dummy variable that equals 1 if individual i in city j cancels an insurance contract purchased on date t within 11 days of purchase.[15] Cancellation information was not provided for the most recent year of data (2015), so those contracts are not included in the cancellation regressions. In addition, we drop the 212 observations in which the policy was canceled on the day of purchase because these individuals are not exposed to the CoP AQI.[16] AQI_{ijt} is the previously used measure of air pollution on the date of purchase, and $(\text{AQI}_{ij,t+1,...,ij,t+11})$ are the eleven daily leads of the pollution variable. C_i includes controls for policy characteristics: the age and gender of the policyholder, whether the insurance was purchased for oneself or another family member, and the length of the insurance contract period in years. As before, X_{jt} is a vector of weather variables, and D_{jt} are day-of-week, and city specific month and year fixed effects designed to capture trends both within a week and over time. Standard errors are clustered on city*date.

We use four different specifications to capture the effect of pollution during the CoP on cancellation rates. Our first specification directly tests if cancellations are affected by differences in the AQI during the times when the purchase and cancellation decisions are made. Specifically, we replace AQI with a measure of the *change* in AQI during the CoP relative to order-date AQI (Relative AQI). That is we run the regression

$$\text{Cancel}_{ijt} = \beta(\text{Relative AQI}_{ijt}) + C_i b + X_{jt}\gamma + D_{jt} + \epsilon_{jt}, \tag{3}$$

where

$$\text{Relative AQI}_{ijt} = \left(\sum_{\tau=1}^{11} \frac{1}{11}\text{AQI}_{ij,t+\tau} - \text{AQI}_{ijt}\right). \tag{4}$$

That is we measure the effect of the average AQI during the CoP normalizing the order-date AQI to zero.

The second specification includes separate controls for both the level of the order-date AQI and the average AQI during the CoP ($\text{AQI}_{\text{CoP}} = \sum_{\tau=1}^{11} \frac{1}{11}\text{AQI}_{ij,t+\tau}$). This specification is essentially identical to that used in Conlin *et al.* (2007) as their test of projection bias. In cases for which

15. Although the legally mandated CoP is ten days, the firm does not appear to strictly enforce the ten-day rule. Consequently, a significant number of cancellations occur eleven days after purchase. Limiting the analysis to a ten-day post-purchase period generates similar results.

16. Including these observations does not materially affect the regression results (see Appendix Tables A2 and A3).

one or more of the daily pollution measures during the CoP were not available, the $\mathrm{AQI}_{\mathrm{CoP}}$ was calculated excluding the missing values.

The third specification is a variant on the second specification but replaces the CoP AQI with the eleven leads of pollution as separate regressors and then sums the eleven resulting coefficients. That is, we replace $\mathrm{AQI}_{\mathrm{CoP}}$ with $\sum_{\tau=1}^{11} \beta_\tau \mathrm{AQI}_{ij,t+\tau}$ and report $\sum_{\tau=1}^{11} \beta_\tau$ as the effect of pollution during the CoP on insurance cancellations. We drop from this regression the 9% of contracts for which one of the lead pollution measures was missing.[17] Subject to the linear functional form assumption, $\sum_{\tau=1}^{11} \beta_\tau$ provides us with a measure of the cumulative effect of daily pollution during the CoP on cancellations.

For our final specification, we utilize a dummy variable to indicate whether air pollution during the CoP is *lower* than on the purchase date. In this case, $f(\mathrm{AQI}_{ijt+1,...,ijt+11})$ is an indicator variable equal to 1 if $\mathrm{AQI}_{\mathrm{CoP}} < \mathrm{AQI}_{\text{order-date}}$.

Table 4 presents the estimated marginal effects at the sample mean associated with the probit regressions of our four variants of equation (2). Column 1 presents the results of regressing Relative AQI on cancellations so that the coefficient of interest represents the effect of AQI during the CoP normalized, such that order-date AQI = 0. We find a negative and statistically significant relationship between Relative AQI and cancellations, indicating that *decreases* in AQI relative to order-date AQI leads to *increases* in the probability of cancellation. Specifically, for every one-unit (standard deviation) decrease in the AQI relative to order-date AQI, the probability of cancellation increases by 0.001% (0.10%). Given the baseline cancellation rate of 2.52%, this corresponds to a 0.040% (3.97%) increase in the cancellation rate.

When we include both order-date AQI and the average AQI during the CoP as regressors (Column 2), we find that higher order-date AQI leads to a positive and statistically significant increase in cancellations, whereas CoP AQI has the opposite effect. Specifically, we find that a one-unit (standard deviation) increase in order-date AQI leads to a 0.0087% (0.087%) increase in the probability of cancellation. In contrast, the coefficient for our measure of air pollution levels during the CoP is negative and statistically significant, with a one-unit increase in CoP AQI decreasing the probability of cancellation by 0.024%. These results indicate that individuals are more likely to cancel their insurance policy if they purchased on a high-pollution day or experienced low pollution during their CoP.

Column 3 presents a repeat of this analysis, but the average CoP AQI is replaced by a disaggregated daily measure of daily AQI, shows essentially the same pattern of results as seen in Column 2: higher order-date AQI leads to a positive and statistically significant increase in cancellations, whereas the aggregate effect of daily air pollution levels during the CoP is negative and statistically significant.[18]

Finally, in Column 4, the analysis shown in Column 2 is repeated but with a dummy variable for whether the average of daily AQI is lower during the CoP relative to purchase-date AQI. Unlike what is seen in Columns 2 and 3, here we find that order-date AQI no longer predicts an increased probability of cancellation. Instead, we find that the effect of air pollution on cancellations depends solely on whether air pollution is lower during the period in which the purchaser can decide to cancel his or her policy relative to the order-date. Specifically, if $\mathrm{AQI}_{\mathrm{CoP}} < \mathrm{AQI}_{\text{order-date}}$, the probability that a contract is canceled increases by 0.19%, representing a 7.25% increase in the

17. Replacing missing observations with a value interpolated from the nearest two non-missing observations produces essentially identical regression results.

18. We reject at a $p < 0.01$ that the sum of the individual lead coefficients is equal to zero.

TABLE 3
The effect of pollution on cancellations

Dependent variable: Indicator equal to 1 if contract is canceled				
% of Contracts canceled	2.90%	2.90%	2.79%	2.90%
Relative AQI	−0.00110***			
	(0.00043)			
Order-date AQI		0.00085*	0.00092*	0.00002
		(0.00045)	(0.00049)	(0.00053)
CoP AQI		−0.00252***		
		(0.00094)		
$\sum_{\tau=1}^{11} \beta_{AQI,\tau}$			−0.00268***	
			(see notes)	
1(CoP AQI<Order-date AQI)				0.2008**
				(0.0881)
Log(Term Length)	−0.507***	−0.570***	−0.563***	−0.571***
	(0.018)	(0.018)	(0.018)	(0.018)
Log(Age)	0.402***	0.402***	0.372***	0.403***
	(0.034)	(0.034)	(0.035)	(0.034)
Self	1.203***	1.201***	1.160***	1.201***
	(0.080)	(0.080)	(0.081)	(0.079)
Female	0.118***	0.119**	0.100*	0.116**
	(0.057)	(0.057)	(0.058)	(0.057)
Adj. R^2	0.059	0.059	0.060	0.059
Observations	411,525	411,525	381,146	411,525

Notes: For each column, the dependent variable is whether an insurance contract is canceled during the CoP. All coefficients represent the marginal effects from a probit regression. *Relative AQI* is the average AQI during the CoP minus the order date AQI. *CoP AQI* is the mean value of AQI $PM_{2.5}$ during the CoP ($\frac{1}{11}\sum_{\tau=1}^{11} \text{AQI}_\tau$). $\sum_{\tau=1}^{11} \beta_{\text{AQI},\tau}$ is the sum of the coefficients for the eleven daily leads of the pollution variable AQI_0 $PM_{2.5}$. We can reject at $p=0.003$ that the sum of these coefficients is greater than zero. For legibility, all coefficients and standard errors have been multiplied by 100. All regressions included controls for temperature, temperature squared, rain, snow, and dummy variables for day of week, city*month and city*year. Column 3 includes additional controls for the eleven daily leads of temperature, temperature squared, rain, and snow. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

cancellation rate. This result suggests that, as predicted by our model, the impact of air pollution on cancellation rates is driven by *relative* differences and not absolute levels. That is, the AQI during the time the decision to cancel is made matters only in how it differs from the AQI that the decision maker faced when making the decision to purchase insurance in the first place.

4.4. *Effect of PM2.5 on cancellations of non-health insurance contracts*

We next examine whether pollution has an effect on the cancellation of non-health insurance policies. To the extent that health shocks do not have a direct effect on the valuation of other forms of insurance, our model would predict that air pollution should not influence whether an individual cancels other types of insurance policies sold by the firm. We test for such a differential effect by re-estimating the values in Column 1 of Table 3 for all insurance contracts, interacting *Relative AQI* with a dummy variable for non-health insurance policies. The results of this regression are presented in Table 4.

Column 1 includes the same controls as does the regressions in Table 3, while Column 2 includes interaction terms between the weather controls, contract characteristics, and a dummy variable for non-health insurance policies to allow those characteristics to have differential effects for health versus other insurance policies. This second specification allows the various control variables to have differential effects on the different insurance products. For both specifications, the main effect of the difference in AQI between the order-date and the CoP remains negative

TABLE 4
Cancellations including non-health insurance

	Dependent variable: indicator equal to 1 if contract is canceled	
% of Contracts canceled	5.36%	5.36%
Relative AQI	−0.00273***	−0.00250***
	(0.00083)	(0.00081)
(Relative AQI)(Other)*	0.00237***	0.00207***
	(0.00080)	(0.00078)
Other	0.02573***	−0.00758***
	(0.00078)	(0.00032)
Log(Term Length)	−0.840***	−1.083***
	(0.035)	(0.031)
Log(Term Length)*Other		0.423***
		(0.057)
Log(Age)	0.542***	0.638***
	(0.042)	(0.059)
Log(Age)*Other		−0.144**
		(0.060)
Self	2.358***	2.045***
	(0.102)	(0.122)
Self*Other		0.626***
		(0.165)
Female	0.288***	0.182*
	(0.065)	(0.095)
Female*Other		0.150
		(0.101)
Adj. R^2	0.062	0.062
Observations	898,064	898,064

Notes: For each column, the dependent variable is whether an insurance contract is canceled during the CoP. All coefficients represent the marginal effects from a probit regression. *Relative AQI* is the average AQI during the CoP minus the order date AQI. For legibility, all coefficients and standard errors have been multiplied by 100. All regressions included controls for temperature, temperature squared, rain, snow, and dummy variables for day of week, city*month, and city*year. Column 2 includes interactions of the *Other* dummy with the controls for temperature, temperature squared, rain, and snow. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

and statistically significant. The interaction term, however, is positive, statistically significant, and only slightly smaller in magnitude than the main effect. Thus, the marginal effect for other insurance types has a magnitude close to zero and is statistically insignificant with p-values of 0.54 and 0.46 for Columns 1 and 2, respectively.[19]

4.5. *Effect of air pollution on insurance contract characteristics*

We next examine the effect on pollution on the characteristics of the insurance contracts purchased. It is important to note that the price of an insurance contract is not individually negotiated. Instead pricing is set at the company level, and is adjusted infrequently. As such the price is unrelated to either idiosyncratic variation in daily air pollution levels or the demand for insurance. Thus any relationship between insurance contract characteristics and the AQI would indicate that air pollution affects either the composition of who purchases insurance or what kinds of insurance features are valued more by individuals due to pollution.

19. As with effect of pollution on demand for other insurance products (Table 2, column 4), the confidence intervals do not allow us to explicitly rule out a potentially meaningful relationship between pollution and cancellation behaviour non-health insurance contracts.

TABLE 5
Pollution and insurance contract characteristics

	Term length	*Age*	*Self purchase*	*Female*	*Female & self*
AQI $PM_{2.5}$	0.00001	0.00004	0.00003	0.00003**	0.00001
	(0.00006)	(0.00007)	(0.00002)	(0.00001)	(0.00002)
Temperature	−0.0024	−0.0102	−0.0006	−0.0015*	−0.0026**
	(0.0036)	(0.0044)	(0.0014)	(0.0008)	(0.0011)
Temperature^2	0.0002**	0.0002**	−0.0000	0.0000**	0.0001**
	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Rain	−0.0230	−0.0050	0.0101	0.0039	−0.0004
	(0.0214)	(0.0253)	(0.0074)	(0.0048)	(0.0067)
Snow	−0.0785	−0.0210	0.0193	0.0207**	0.0110
	(0.0511)	(0.0594)	(0.0162)	(0.0103)	(0.0141)
Adj. R^2	0.052	0.009	0.010	0.000	0.004
Observations	579,303	579,303	579,303	579,303	274,102

Notes: Columns 1 and 2 present the results from ordinary least square regressions, and columns 3 through 5 present marginal effects based on a probit model. The dependent variable for columns 1 and 2 are the log of the contract term and the log of the age of the person covered by the health insurance contract. For columns 3 through 5, the dependent variable is a dummy equal to 1 if (3) the insurance was purchased for oneself, (4) the insurance was purchased for a female, and (5) the insurance was purchase by a female. The sample size is smaller for column (5) because the sample was limited to insurance purchased for oneself as those are the only cases for which we can identify the gender of the purchaser. All regressions included dummy variables for day of week, city*month, and city*year. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

To determine whether pollution affects the characteristics of insurance policies sold, we estimate the following equation:

$$C_{ijt} = \beta \text{AQI}_{jt} + X_{jt}\gamma + D_{jt} + \epsilon_{jt}. \tag{5}$$

Here, the dependent variable C_{ijt} is a characteristic of insurance plan i sold in city j on date t. As in equation (1), AQI_{jt} is a measure of the high AQI in city j on date t, X_{jt} consists of a quadratic function of high temperature and dummy variables for precipitation and snowfall, and D_{jt} are day-of-week, month-of-year*city, and year*city fixed effects.[20] All standard errors are clustered at the city*date level.

The results of estimating equation (5) are presented in Table 5. Columns 1 and 2 present the results from an OLS regression for which the dependent variable is the log of the term length of the insurance contract or the log of the age of the covered individual, respectively. Columns 3–5 present the estimated marginal effects at the sample means from a probit regression, where C_{ijt} is an indicator variable equal to 1 if the purchaser and covered individuals are the same (3), if the covered individual is female (4), or if the purchaser is female (5). For Column 5, the sample is limited to those insurance contracts for which the purchaser is the same as the covered individual, as those are the only cases for which we can determine the gender of the purchaser.

In all cases, β is small and, with the exception of Column 4 (the covered individual's gender), statistically insignificant. For Columns 1, 2, 3, and 5, given that the standard errors are at least an order or magnitude smaller than the effect sizes shown in Table 2, we can rule out the AQI having an economically meaningful effect on these contract characteristics. Further, although the coefficient for Column 4 is statistically significant, the effect size itself is quite small, with a one-unit (standard deviation) increase in AQI causing a 0.00001% (0.01%) increase in the share

20. Adding additional controls for contract characteristics other than the dependent variable generates effectively identical results.

of contracts that insure females off a baseline of 55%. Overall, these results suggest that, although air pollution significantly increases the demand for insurance, it does not appear to meaningfully change either the characteristics of the insurance contracts nor the composition of who buys insurance.

5. POTENTIAL MECHANISMS

Our two main empirical findings are that (1) higher air pollution leads to greater demand for health insurance, and (2) cancellation rates are higher if air pollution levels during the CoP are lower than that on the order-date. In this section, we discuss several potential explanations for our findings.

5.1. *Learning*

Perhaps the most obvious explanation for the increase in the sales of health insurance policies on high pollution days involve learning, broadly defined. The biggest obstacle to such explanations is the fact that, as discussed in Section 3 and illustrated by Figure 1, today's pollution contains essentially no additional information about pollution six months in the future (when the insurance policy takes effect).

One type of learning that addresses this critique is learning by inattentive individuals (see Schwartzstein (2014) for a recent example). For example, if many individuals are unaware of the relationship between air pollution and health, and high pollution days cause individuals to learn about the deleterious effects of air pollution on their future heath, such learning could increase their demand for insurance. While such a mechanism is consistent with the increase in sales on high pollution days, it is harder to reconcile with the increase in cancellation if pollution drops post-purchase. That is while an increase in awareness about the connection between air pollution and health will lead to an increase in demand for health insurance, unless low air pollution levels causes people to "unlearn" this connection, it cannot lead to increases in cancellations.

Alternatively high pollution may cause an individual to learn something about their own, or family member's health. Since in the policies we study pre-existing conditions are not covered, and the policy itself does not take effect until six months after purchase, high levels of air pollution can generate increased demand if it induces learning about an undiagnosed long-run or chronic condition. Then as with learning about pollution or the pollution-health connection more generally, while such a mechanism is compatible with the increase in demand on high pollution days, it is much harder to reconcile with the cancellation results.

A key prediction of many models of inattentive learning is that while high levels of pollution would be correlated with higher sales, such sales would simply represent shifts in the timing of insurance purchases, and not true increases in demand (*e.g.* harvesting). That is everyone who buys insurance would have purchased it eventually, and high pollution day simply moves forward the timing of the purchase.

To assess whether the increase in demand associated with daily pollution is driven by intertemporal substitution, we estimate a distributed lag model. Specifically we rerun equation (1) with N daily lags of AQI and weather added to the estimating equation:

$$\text{Log}(\text{Insurance}_{jt}) = \beta \text{AQI}_{jt} + \sum_{\tau=1}^{N} \beta_\tau \text{AQI}_{j,t-\tau} + \sum_{\tau=0}^{N} X_{j,t-\tau}\gamma_\tau + D_{j,} + \epsilon_{jt}. \qquad (6)$$

As before Insurance_{jt} is the number of insurance contracts sold by the firm to residents of city j on date t, while $\text{AQI}_{j,t-\tau}$ and $X_{j,t-\tau}$ are lagged measures of AQI and weather in city j relative to

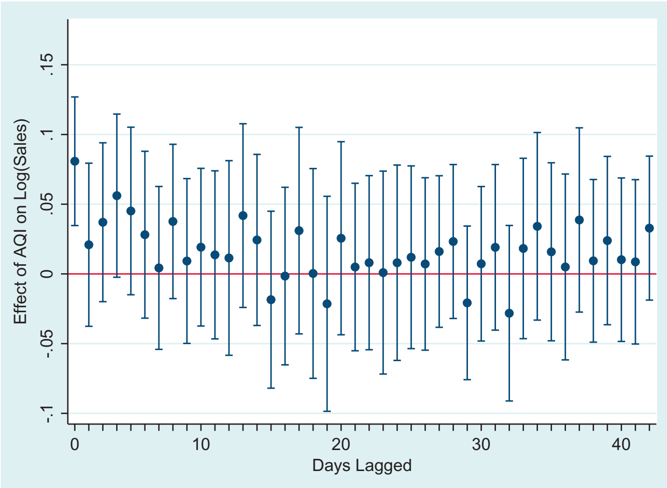

FIGURE 4

Coefficient values and 95% confidence intervals for the effect of contemporaneous and lagged AQI/100 on daily insurance sales.

date t. Including lagged pollution variables in our regression allows us to test whether pollution in the days leading up to (or following) the day of purchase affects the impact of contemporaneous pollution on purchase decisions. For example, a negative coefficient on the fifth-day-lagged pollution measure would indicate both that high pollution five days prior leads to lower sales for the current day and that high pollution *today* leads to lower sales five days in the future. The sum of the lagged coefficients are then a measure of the extent to which the current period effect is due to intertemporal substitution and how much is an increase in aggregate total demand for insurance. Thus, if the increase in demand that we measured in the previous section is due to displacement, the sum of $k \leq N$ coefficients for the lagged pollution variable will be equal in magnitude to the current period coefficient β and of opposite sign.[21]

Figure 4 presents the results of this analysis through a plot of the estimated coefficients on current and lagged API along with 95% confidence intervals from estimating equation (6) for a period of six weeks ($N=42$). As shown in the figure, the current period pollution has a large, positive, and statistically significant impact on the demand for health insurance contracts. The current-day pollution coefficient β in this regression equals 0.00081, with a standard error of 0.00024, a value that is slightly larger than the coefficient of 0.00072 in Table 2.

In contrast, the coefficients for the lagged pollution are smaller and never statistically significant. Moreover, that most coefficients tend to be positive, even if not statistically significantly so, suggests that high pollution in the recent past leads to higher insurance sales for the current day. Indeed testing the null hypothesis that the sum of the coefficients for first k lags is equal to the negative of the current-day coefficient β, we find that we can reject the null hypothesis with a $p<0.001$ for k equal to $7, 14, 21, 28, 35$, or 42 days. These results indicate that the increase in daily sales generated by air pollution can be interpreted as an increase in the aggregate demand for insurance rather than a change in intertemporal substitution across days.

21. See Jacob *et al.* (2007), Deschenes and Moretti (2009) and Busse *et al.* (2015) for a more detailed discussion of the methodology used here.

To be clear, while the cancellation results are incompatible with some models of inattentive learning, we cannot definitely say that inattentive learning does not have a a role in driving the increase in demand on high pollution days. Rather the empirical results rules out the possibility that learning by inattentive individuals is the *only* mechanism at work.

5.2. *Projection bias*

Projection bias is the tendency for individuals to exaggerate the degree to which their future tastes will resemble their current tastes. Although projection bias has received significant attention in both the economics and psychology literature, there is only some recent evidence that projection bias influences demand for real goods and services. The lack of empirical evidence from the field is largely due to the fact that detailed data on refunds or cancellations is required to distinguish projection bias from alternative mechanisms.[22]

Perhaps the most convincing prior evidence of projection bias in a real-world market comes from the first paper to document such bias in a real-world context: Conlin *et al.* (2007). Conlin *et al.* convincingly show that catalogue orders for weather-related clothing items are overinfluenced by the weather. They find that lower order-date temperature leads to an increase in the return probability for cold-weather items, but find only mixed evidence regarding the impact of return-date temperature on returns. Simonsohn (2010) and Busse *et al.* (2015) show, respectively, that weather also affects college enrollment and the type of automobile purchased. Busse *et al*. conclude that their results are incompatible with the behaviour of standard, rational agents but consistent with both projection bias and salience.

Loewenstein *et al.* formalize this idea with a model in which an agent's utility is given by

$$\tilde{u}(c,s|s')=(1-\alpha)u(c,s)+\alpha u(c,s'), \tag{7}$$

where s is a state variable that affects the utility of good c, s' is a person's current state, and $\alpha \in [0,1]$ is a measure of the projection bias exhibited by the agent. In this case, if agents have $\alpha=0$, they accurately predict their utility from good c in state s. In contrast, if $\alpha>0$, then they mis-predict their utility in state s as a convex combination of their true utility from c in state s and the utility that they would receive from c given their current state s'.

As illustrative example of the influence s' can have on the demand for health insurance, we can assume a simple utility function of the form

$$\tilde{u}(c_t,s|s')=(1-\alpha)B(s)+\alpha B(s')-p, \tag{8}$$

where s' is a measure of how sick an individual is now, s is a measure of how sick an individual will be at some point in the future, $B(s)$ is non-zero, increasing function that represents the per-period benefit provided by the insurance policy, and p is the insurance premium.

Mapping this utility function to the case of health insurance choice is straightforward. Consider individuals who purchase health insurance I at time t with a policy period equal to T, and let s_τ represent their expected future health.[23] Conditional on purchase, they can, without cost, cancel

22. For this reason nearly all of existing evidence for projection bias comes from either surveys or experiments. See for example Nisbett and Kanouse (1969), Loewenstein and Frederick (1997), Read and van Leeuwen (1998), Badger *et al.* (2007), Acland and Levy (2014), and Augenblick and Rabin (2016).

23. Note that the individuals expectations about their future health need not be correct, rather the model requires that their expectations about the future states of the world are unaffected by current conditions. We discuss the possible role of current conditions affecting peoples expectations (*e.g.* the mistaken beliefs channel) in Section 5.2.1.

their policy at time $t+1$ (*e.g.* the CoP), with coverage to begin at time $t+2$ (*i.e.* after the "waiting period"). Their perceived utility from purchasing health insurance in period t is then given by

$$\tilde{U}^t(\tilde{u}_{t+2},...,\tilde{u}_{t+T}|s_t)=\sum_{\tau=2+t}^{T+t}\delta^{(\tau-t)}[(1-\alpha)B(s_\tau)+\alpha B(s_t)]-p. \tag{9}$$

This simple framework illustrates the influence that s_t can have on the demand for insurance. Although the current state s_t will have no effect on the perceived utility of rational agents ($\alpha=0$), individuals who suffer from projection bias will value insurance more the sicker they are at present. Thus, for a given price p, demand will be higher on days when individuals feel unwell.

Next, consider the behaviour of individuals during the CoP ($t+1$) conditional on having purchased insurance at t. Again, if they are rational, their predicted utility will not be affected by their current health s_{t+1}; however, if they are affected by projection bias, then their predictions regarding the utility from insurance will be biased by their current health. They will then choose to cancel their insurance if $\delta\tilde{U}^t(c_{t+2},...,c_{t+T}|s_{t+1})<0$. Thus there will then be an $\underline{s}<s_t$ such that, if $s_{t+1}<\underline{s}$, they will cancel their insurance in period $t+1$. That is, individuals with projection bias will cancel their insurance if their sickness level during the time that they are making the decision to cancel their policy is sufficiently low *relative* to their purchase-day level.

These results generate a pair of testable predictions that exactly match our results:

(1) Negative transitory health shocks (s_t) will increase contemporaneous demand for health insurance.
(2) If individuals feel healthier during the CoP *relative* to the order-date (*i.e.*, $s_{t+1}<s_t$), they are (weakly) more likely to cancel their insurance policy.

5.2.1. Mistaken beliefs. Another closely related explanation for our results is that high pollution today causes individuals to *mistakenly* believe that air pollution will be higher in the future. The standard model of projection bias assumes individuals correctly predict future states, and instead errs only in predicting the utility they will receive in those states. But since utility is state dependent, if current conditions causes individuals to err in predicting futures, individuals may also fail to predict future utility in similar ways. Indeed, Busse *et al.* (2015) refer to the former as "projection bias of utility" and the second as "projection bias of states".

To the extent that beliefs are overly sensitive to current conditions, it could generate both the increase in demand for health insurance on high pollution days and the increase in cancellations conditional on (relatively) low pollution days. Indeed recent studies have shown that belief in global warming is affected by recent outdoor temperatures (*e.g.* Li *et al.*, 2011)

Although such a mechanism would also predict current conditions having an overly large impact on the demand, they operate through a fundamentally different channels. For example, under the mistaken belief channel, a cold day causes individuals to mistakenly believe that there will be more cold days in the future. In contrast, according to projection bias, a cold day causes individuals to mistakenly believe they will receive higher utility from a cold weather coat on non-cold days.

To try and differentiate between these two channels, we ran an online survey designed to elicit individual's beliefs regarding pollution in their city of residence in the future. Since this pattern of mistaken beliefs (*i.e.* current conditions have an outsized influence) has been attributed to several different psychological mechanism (*e.g.* recency bias, availability bias, limited recall, myopic learning), such a test cannot distinguish between these different mechanisms. Rather the results can be interpreted as evidence for or against any mechanism that operates through affecting the beliefs about future pollution.

TABLE 6
The effect of pollution on beliefs

Dependent variable: pollution is getting Better = 1, Same = 2, Worse = 3				
	Ordered probit		Probit (Worse = 1)	
AQI	−0.00074	−0.00066	−0.00021	−0.00008
	(0.00046)	(0.00055)	(0.00050)	(0.00062)
Age		0.0011		−0.0038
		(0.0026)		(0.0017)
Female		−0.0452		0.0892**
		(0.0345)		(0.0439)
Temperature (high)		0.0048		−0.0107**
		(0.0053)		(0.0055)
Temperature (low)		−0.0029		0.0048
		(0.0074)		(0.0081)
Years in Current City		0.0028*		0.0038**
		(0.0015)		(0.0017)
Pseudo R^2	0.003	0.011	0.000	0.029
Observations	461	461	461	461

Notes: All coefficients represent marginal effects. The results are from a multi-city online survey in China. The dependent variable is in response to a question about pollution "a year from now" while the key dependent variable is the AQI for the city of the respondent on the day the survey was completed. Columns 3 and 4 show the results of a Probit regression where responses of "Better" or "Same" were coded as 0, while responses of "Worse" were coded as 1. All regressions use robust standard errors.
* significant at 10%, ** significant at 5%, *** significant at 1%.

The survey asked for basic demographic information (age, gender, years in current city) as well as the following question: "Do you think that a year from now air pollution where you currently live will be 1) better, 2) the same, 3) worse." Of the 461 respondents, 98 thought that air pollution would be better in a year, 289 though that it would be the same, and 77 thought it would get worse. The survey data was then merged with air pollution and weather data for the city of the respondent on the day the survey was completed.

The results of the survey analysis are presented in Table 6. Columns 1 and 2 present the result of a ordered probit analysis (better, the same, worse) with and without controls respectively. Columns 3 and 4 present the result of a probit analysis where the dependent variable equals one if the respondent said that she felt pollution would be worse in a year. Across all four specifications, we find no relationship between AQI and one's beliefs about air pollution in a year's time. For the probit analysis, gender, temperature, and tenure in one's current city of residence are all significantly correlated with the probability that the survey respondent feels pollution will be worse in a year's time.

While these results do not provide empirical support the mistaken belief hypothesis, these results also do not provide particularly strong evidence against it. First, there are likely considerable differences in sample composition between the survey respondents and the insurance company's potential customers. For example, our survey participants were from a much wider range of cities than in our sample, and nearly all of our survey participants had private health insurance. In addition, we may simply be asking the wrong question. That is while current pollution might not affect ones belief about future pollution, perhaps the ill health caused by current pollution affects ones belief about the probability of being sick in the future. Finally, while we find no relationship between current pollution levels and belief about future pollution, the confidence intervals prevent us from rejecting a meaningful relationship between the two variables. Indeed the best argument against the mistaken beliefs channel may be the same offered in previous work: that the ubiquity of state information (air pollution in large Chinese cities) relative to utility information (how much utility one will receive from supplemental health insurance in different states of the world), suggests that individuals are more able to accurately predict future pollution than future utility.

5.3. *Salience*

Salience refers to the tendency for an individual's attention to be directed towards certain "salient" features, and that such features will disproportionately affect decision making.[24] As with projection bias, salience has received significant attention in both the economics and psychology literature, but unlike projection bias there is a substantial body of empirical evidence that finds evidence of salience in the field.[25]

In our context, this might suggest that when air pollution is high, the risks of contracting a pollution-related disease is more salient or top of mind, thus increasing the perceived value of health insurance.[26] Then when pollution levels drop, the risks of such diseases are less top of mind, decreasing the perceived value of health insurance. As such, salience would generate both the prediction that higher air pollution levels increases demand for insurance and that a subsequent, relative decreases in air pollution levels would lead to a increase in the cancellation rate.

5.3.1. Salience versus projection bias. Since projection bias and salience both generate directionally similar predictions regarding individual response to changes in air pollution, empirically distinguishing between them is difficult in this domain. One possibility is to exploit the fact that, roughly speaking, projection bias is about the absolute level of a state variable, while salience (as formalized in Bordalo *et al.* (2013)) is about differences. That is while projection bias would predict that demand responds to the absolute air pollution level, salience would predict demand responds to "surprises" relative to some benchmark. That is under salience, the relationship between daily air pollution and the demand for health insurance is dependent on one or more benchmarks, while the effects of projection bias are not. Unfortunately, one difficulty with exploiting this difference is that it is not clear what the appropriate benchmarks are for measuring surprise in a our context. Following Busse *et al.* (2015), we explore expected conditions for a particular time of year and recent experience as two potential benchmarks for measuring surprise for differentiating between salience and projection bias.

The first potential benchmark discussed in Busse *et al.* (2015) is the expected weather for a particular time of year. Under this "surprise relative to expectations" interpretation, pollution would be more salient when it is differs from the seasonal average. In this case, salience would predict that high pollution has a larger impact at times of year when pollution is typically low compared to when it is typically high. In contrast, projection bias would predict that the impact of high pollution would be the same regardless of the time of year. Unfortunately, as in Busse *et al.* (2015), we cannot differentiate between projection bias and salience under this benchmark because air pollution follow a seasonal trend. As such, we cannot rule out a discounted utility explanation for a correlation between air pollution and demand in general, rather only in the case of high frequency, idiosyncratic variations in pollution levels (*i.e.* daily air pollution levels should have effectively zero effect on the demand for health insurance of a rational agent). And since both projection bias and salience would predict a positive relationship between idiosyncratic

24. See Bordalo *et al.* (2013, 2014); Koszegi and Szeidl (2013); Bushong *et al.* (2016) for recent formalizations.

25. Recent examples include, Gabaix and Laibson (2006), Chetty *et al.* (2009), Malmendier and Lee (2011), Lacetera *et al.* (2012), Hastings and Shapiro (2015), and Dalton *et al.* (2017).

26. Indeed, Zwane *et al.* (2011) find that surveying households about health increased take-up of water treatment products and medical insurance.

pollution and demand, we cannot differentiate between them when seasonally adjusted pollution expectations are the benchmark for surprise.

The second potential benchmark examined in Busse *et al.* (2015) is the idea of a "surprise relative to recent" experience. In that case, high levels of air pollution are more salient when pollution was previously low. Under this interpretation of surprise, salience would predict that high levels of pollution will lead to a larger increase in the demand for health insurance if it was preceded by a period of low pollution days compared to a period of high pollution days. In contrast, under projection bias demand for health insurance will be unaffected by recent past pollution levels.

The results from the distributed lag model in Section 5.1 allows us to test whether recent air pollution levels affects the demand for health insurance. As shown in Figure 4, consistent with projection bias, but not salience where recent pollution is the benchmark for surprises, recent pollution is not associated with demand for health insurance. Indeed, if anything, recent pollution is positively, and not negatively, correlated with demand. But since there is no evidence to suggest that recent pollution is the appropriate choice context benchmark, this is at best weak evidence against salience in general as the driver of our results.[27] Arguably, the main take-away from the preceding analysis is to illustrate the difficultly in differentiating between projection bias and salience.

6. CONCLUSION

Our main empirical findings are that (1) transiently higher levels of air pollution leads to greater demand for health insurance, and (2) cancellation rates are higher if air pollution levels during the CoP are lower than that on the order-date. These effects are limited to health insurance contracts, with air pollution levels having no analogous effects on other insurance products sold by the firm. We also find that the increase in daily demand for health insurance engendered by daily air pollution levels represents an increase in total demand for insurance, and not the result of temporal displacement of purchases. Finally, we have some evidence that current pollution levels do not affect individual expectations regarding pollution in the future. We explore a range of potential mechanisms and find that the results are most consistent with projection bias and salience.

These results show that transitory conditions can have an oversized, and significant impact on real-world product markets in a way that is difficult to reconcile with rational choice theory. To the extent that this result can be generalized to other settings, our results adds to the evidence that intertemporal behavioural biases can be an important driver of demand for real goods and services.

From a policy perspective, our results provide evidence in support of the idea that CoPs attenuate the effects of intertemporal behavioural bias. Furthermore, they suggest that the autocorrelation in the state variable driving the biased should be taken into consideration when determining the length of such periods. More directly, our results suggests that policy makers need to take into account the fact that an individual's current health can have an oversized influence on their demand for health insurance when designing and regulating health care insurance markets.

27. Another differential prediction discussed in Busse *et al.* (2015) is that "When the weather is extremely cold", small changes in temperature would be associated with a change in demand under salience but not projection bias. They find that "car sales are positively correlated with temperature even at very low temperature levels", a finding that "could be seen as evidence for salience rather than projection bias". In our case, we find the opposite result (demand for insurance is positively correlated with pollution only at higher pollution levels). But this should not necessarily be seen as evidence against salience as, unlike changes in temperature, individuals are generally not be able to detect changes in air pollution when levels are low.

APPENDIX

TABLE A1
The effect of pollution on insurance sales (robustness)

	Dependent variable: log(number of contracts sold)									
$AQI_{PM2.5}$	0.00094*** (0.00018)	0.00095*** (0.00018)	0.00085*** (0.00019)	0.00084*** (0.00018)	0.00072*** (0.00018)	0.00073*** (0.00018)	0.00056*** (0.00018)	0.00054*** (0.00019)	0.00056*** (0.00019)	0.00084*** (0.00032)
City	Y	Y	N	N	Y	Y	N	N	Y	N
Year	Y	Y	N	N	Y	Y	N	N	Y	N
DOW	Y	Y	N	N	Y	Y	N	N	Y	N
Month	Y	Y	N	N	N	N	N	N	N	N
Week	N	N	N	N	Y	Y	N	N	N	N
Day-of-year	N	N	N	N	N	N	N	N	Y	N
City*Year	N	N	Y	Y	N	N	Y	Y	N	Y
City*DOW	N	N	Y	Y	N	N	Y	Y	N	Y
City*Month	N	N	Y	Y	N	N	N	N	N	N
City*Week	N	N	N	N	N	N	Y	Y	N	N
City*Day-of-year	N	N	N	N	N	N	N	N	N	Y
Weather	N	Y	N	Y	N	Y	N	Y	N	N
Adjusted R^2	0.447	0.452	0.524	0.528	0.481	0.484	0.572	0.575	0.489	0.567
Observations	2,573	2,573	2,573	2,573	2,573	2,573	2,573	2,573	2,573	2,573

Notes: All columns present the results from ordinary least square regressions. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE A2
The Effect of pollution on cancellations (with same day cancellations)

Dependent ariable: indicator equal to 1 if contract is canceled				
% of Contracts canceled	2.91%	2.91%	2.84%	2.91%
Relative AQI	−0.00100**			
	(0.00043)			
Order-date AQI		0.00078*	0.00086*	0.00004
		(0.00045)	(0.00050)	(0.00054)
CoP AQI		−0.00224*		
		(0.00095)		
$\sum_{\tau=1}^{11} \beta_{AQI,\tau}$			−0.00251***	
			(see notes)	
1(CoP AQI<Order-date AQI)				0.1963**
				(0.0885)
Log(Term Length)	−0.588***	−0.587***	−0.580***	−0.587***
	(0.018)	(0.018)	(0.018)	(0.018)
Log(Age)	0.413***	0.414***	0.384***	0.414***
	(0.034)	(0.034)	(0.035)	(0.034)
Self	1.227***	1.226***	1.118***	1.226***
	(0.080)	(0.080)	(0.083)	(0.080)
Female	0.120**	0.120**	0.101*	0.118**
	(0.057)	(0.057)	(0.058)	(0.057)
Adj. R^2	0.061	0.061	0.063	0.061
Observations	411,737	411,737	381,347	411,737

Notes: For each column, the dependent variable is whether an insurance contract is canceled during the CoP. All coefficients represent the marginal effects from a probit regression. *Relative AQI* is the average AQI during the CoP minus the order date AQI. *CoP AQI* is the mean value of *AQI* $PM_{2.5}$ during the CoP ($\frac{1}{11}\sum_{\tau=1}^{11} \text{AQI}_\tau$). $\sum_{\tau=1}^{11} \beta_{\text{AQI},\tau}$ is the sum of the coefficients for the eleven daily leads of the pollution variable AQI_0 $PM_{2.5}$. We can reject at $p=0.003$ that the sum of these coefficients is greater than zero. For legibility, all coefficients and standard errors have been multiplied by 100. All regressions included controls for temperature, temperature squared, rain, snow, and dummy variables for day of week, city*month, and city*year. Column 3 includes additional controls for the 11 daily leads of temperature, temperature squared, rain, and snow. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE A3
Cancellations including non-health insurance (with same day cancellations)

	Dependent variable: indicator equal to 1 if contract is canceled	
% of Contracts canceled	5.36%	5.36%
Relative AQI	−0.00251***	−0.00229***
	(0.00084)	(0.00082)
(Relative AQI)(Other)*	0.00235***	0.00207***
	(0.00080)	(0.00078)
Other	0.02618***	0.00893***
	(0.00079)	(0.00032)
Log(Term Length)	−0.879***	−1.118***
	(0.035)	(0.032)
Log(Term Length)*Other		0.419***
		(0.058)
Log(Age)	0.555***	0.662***
	(0.043)	(0.060)
Log(Age)*Other		−0.165***
		(0.060)
Self	2.448***	2.099***
	(0.103)	(0.124)
Self*Other		0.697***
		(0.167)
Female	0.292***	0.186*
	(0.065)	(0.095)
Female*Other		0.150
		(0.100)
Adj. R^2	0.064	0.064
Observations	899,358	899,358

Notes: For each column, the dependent variable is whether an insurance contract is canceled during the CoP. All coefficients represent the marginal effects from a probit regression. *Relative AQI* is the average AQI during the CoP minus the order date AQI. For legibility, all coefficients and standard errors have been multiplied by 100. All regressions included controls for temperature, temperature squared, rain, snow, and dummy variables for day of week, city*month, and city*year. Column 2 includes interactions of the *Other* dummy with the controls for temperature, temperature squared, rain, and snow. Standard errors are clustered on city*date.
* significant at 10%, ** significant at 5%, *** significant at 1%.

Acknowledgments. We are grateful to Harry DeAngelo, Stefano DellaVigna, Cary Frydman, Uri Gneezy, Ori Heffetz, David Hirshleifer, John Matsusaka, Sendhil Mullainathan, Ted O'Donoghue, Sally Sadoff, Josh Schwartzstein, Charles Sprenger, Avanidhar Subrahmanyam, and participants at the Behavioral Economics Annual Meeting, China International Conference in Finance, Federal Reserve Bank of Cleveland, LA Finance Day, UCSD Rady, and USC for helpful suggestions. We are also grateful for helpful comments from the editor and two anonymous referees. Wang thanks the China National Natural Science Foundation (Grant Number: 71432008) and the Australian Research Council (DP150102339) for financial support. Huang thanks the National Program for Special Support of Eminent Professionals (Top-notch Young Professionals) and the National Natural Science Foundation of China (Grant Number: 71673044) for financial support.

Supplementary Data

A PDF explaining the absence of supplementary data is available at *Review of Economic Studies* online.

REFERENCES

ACLAND, D. and LEVY, M. (2014), "Naivete, Projection Bias, and Habit Formation in Gym Attendance", *Management Science,* **61**, 146–160.

ARIELY, D. and LOEWENSTEIN, G. (2006), "The Heat of the Moment: The Effect of Sexual Arousal on Sexual Decision Making", *Journal of Behavioral Decision Making,* **19**, 87–98.

AUCHINCLOSS A., DIEZ ROUX, A., DVONCH, J., *et al*. (2008), "Associations Between Recent Exposure to Ambient Fine Particulate Matter and Blood Pressure in the Multi-ethnic Study of Atherosclerosis (MESA)", *Environmental Health Perspectives,* **116**, 486–491.

AUGENBLICK, N. and RABIN, M. (2016), "An Experiment on Time Preference and Misprediction in Unpleasant Tasks" (Working Paper).

BADGER, G., BICKL, W., GIORDANO, L., *et al.* (2007), "Altered States: The Impact of Immediate Craving on the Valuation of Current and Future Opioids", *Journal of Health Economics,* **26**, 865–876.

BORDALO, P., GENNAIOLI, N. and SHLEIFER, A. (2013), "Salience and Consumer Choice", *Journal of Political Economy,* **121**, 803–843.

—— (2014), "Memory, Attention and Choice", (Working Paper).

BUSHONG, B., RABIN, M. and SCHWARTZSTEIN, J. (2016), "A Model of Relative Thinking", (Working Paper).

BUSSE, M., POPE, J., POPE, D., *et al*. (2015), "The Psychological Effects of Weather on Car Purchases", *Quarterly Journal of Economics*, **130**, 371–414.

CHANG, T., GRAFF ZIVIN, J., GROSS, T., *et al*. (2014), "Particulate Pollution and the Productivity of Pear Packers", *American Economic Journal: Economic Policy*, **8**, 141–169.

—— (2016), "The Effect of Pollution on Office Workers: Evidence from Call Centers in China" (Working Paper).

CHETTY, R., LOONEY, A. and KROFT, K. (2009), "Salience and Taxation: Theory and Evidence", *American Economic Review*, **99**, 1145–1177.

CONLIN, M., O'DONOGHUE, T. and VOGELSANG, T. (2007), "Projection Bias in Catalog Orders", *American Economic Review*, **97**, 1217–1249.

DALTON, C, GOWRISANKARAN, G. and TOWN, R. (2017), "Salience, Myopia, and Complex Dynaic Incentives: Evidence from Medicare Part D" (Working Paper).

DELLAVIGNA, S. (2009), "Psychology and Economics: Evidence from the Field", *Journal of Economic Literature*, **47**, 315–372.

DESCHENES, O. and MORETTI, E. (2009), "Extreme Weather Events, Mortality, and Migration", *Review of Economics and Statistics*, **91**, 659–681.

DOCKERY, D. W. and POPE, C. A. (1994), "Acute Respiratory Effects of Particulate Air Pollution", *Annual Review of Public Health,* **15**, 107–113.

ENVIRONMENTAL PROTECTION AGENCY (2004), "Air Quality Criteria for Particulate Matter" (Research Triangle Park, NC: National Center for Environmental Assessment).

GABAIX, X. and LAIBSON, D. (2006), "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets", *Quarterly Journal of Economics*, **121**, 505–540.

GHIO, A., KIM, C. and DEVLIN, R. (2000), "Concentrated Ambient Air Particles Induce Mild Pulmonary Inflammation in Healthy Human Volunteers", *American Journal of Respiratory and Critical Care Medicine,* **162**, 981–988.

GRAFF ZIVIN, J. and NEIDELL, M. (2012), "The Impact of Pollution on Worker Productivity", *American Economic Review*, **102**, 3652–3673.

HAGGAG, K. and POPE, D. (2016), "Attribution Bias in Consumer Choice" (Working Paper).

HANNA, R. and OLIVA, P. (2015), "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City", *Journal of Public Economics*, **122**, 65–79.

HASTINGS, J. and SHAPIRO, J. (2015), "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks", *Quarterly Journal of Economics*, **128**, 1449–1498.

HERRNSTADT, E. and MUEHLEGGER, E. (2015), "Air Pollution and Criminal Activity: Evidence from Chicago Microdata" (NBER Working Paper, No. 21787).

JACOB, B., LEFGREN, L. and MORETTI, E. (2007), "The Dynamics of Criminal Behavior: Evidence from Weather Shocks", *The Journal of Human Resources*, **42**, 489–527.

KOSZEGI, B. and SZEIDL, A. (2013), "A Model of Focusing in Economic Choice", *Quarterly Journal of Economics,* **128**, 53–107.

LAIBSON, D. (1997), "Golden Eggs and Hyperbolic Discounting", *Quarterly Journal of Economics,* **112**, 443–477.

LI, Y., JOHNSON, E. and ZAVAL, L. (2011), "Local Warming: Daily Temperature Change Influences Belief in Global Warming", *Psychological Science*, **22**, 454–459.

LAVY, V., EBENSTEIN, A. and ROTH, S. (2014), "The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation" (NBER Working Paper, No. 20468).

LOEWENSTEIN, G. (2005), "Hot-Cold Empathy Gaps and Medical Decision Making", *Health Psychology,* **24**, S49–S56.

LOEWENSTEIN, G. and FREDERICK, S. (1997), "Predicting Reactions to Environmental Change", in Bazerman, M., Messick, D., Tenbrunsel, A. and Wade-Benzoni, K., (eds), *Psychological Perspectives on Ethics and the Environment* (San Francisco: New Lexington Press).

LOEWENSTEIN, G., O'DONOGHUE, T. and RABIN, M. (2003), "Projection Bias in Predicting Future Utility", *Quarterly Journal of Economics,* **118**, 1209–1248.

LACETERA, N., POPE, D. and SYDNOR, J. (2012), "Heuristic Thinking and Limited Attention in the Car Market", *American Economic Review*, **11**, 749–787.

MALMENDIER, U. and LEE, Y. (2011), "The Bidder's Curse", *American Economic Review*, **11**, 749–787.

NISBETT, R. and KANOUSE, D. (1969), "Obesity, Food Deprivation, and Supermarket Shopping Behavior", *Journal of Personality and Social Psychology*, **12**, 289–294.

O'DONOGHUE, T. and RABIN, M. (1999), "Doing It Now or Later", *American Economic Review,* **89**, 103–124.

—— (2015), "Present Bias: Lessons Learned and To Be Learned", *American Economic Review,* **105**, 273–279.

PEW RESEARCH CENTER (2015), Corruption, Pollution, Inequality Are Top Concerns in China. http://www.pewglobal.org/files/2015/09/Pew-Research-Center-China-Report-FINAL-September-24-2015.pdf.

POPE, C. (2000), "Epidemiology of Fine Particulate Air Pollution and Human Health: Biologic Mechanisms and Who's at Risk?", *Environmental Health Perspectives,* **108**, 713–723.

READ, D. and VAN LEEUWEN, B. (1998), "Predicting Hunger: The Effects of Appetite and Delay on Choice", *Organizational Behavior and Human Decision Processes,* **76**, 189–205.

SCHWARTZSTEIN, J. (2014), "Selective Attention and Learning", *Journal of the European Economic Association,* **12**, 1423–1452.

SEATON, A., MACNEE, W., DONALDSON, K. and GODDEN, D. (1995), "Particulate Air Pollution and Acute Health Effects", *Lancet*, **8943**, 176–178.

SIMONSOHN, U., (2010), "Weather to Go to College", *Economic Journal*, **120**, 270–280.

ZWANE, A., ZINMAN, J., DUSEN, E., *et al*. (2011), "Being Surveyed Can Change Later Behavior and Related Parameter Estimates", *PNAS,* **108**, 1821–1826.