PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CITE OR CIRCULATE WITHOUT AUTHORS' PERMISSION.

Projection Bias and the Demand for Insurance

Tom Y. Chang, Wei Huang, and Yongxiang Wang*

January 2015

ABSTRACT

Using data on insurance contracts from one of China's largest insurance companies, we find that air pollution has a significant effect on the decision to purchase or cancel health insurance. A one standard deviation increase in daily air pollution leads to a 9 percent 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 4.1% These results are strongly consistent with projection bias and suggest the importance of projection bias in understanding the demand for insurance.

*Chang and Wang are with the University of Southern California, and Huang is with the University of International Business and Economics.

1 Introduction

By definition the important decisions people make have lasting consequences, and as such require them to predict their utility in future states. While standard economic theory assumes that individuals can accurately do so, evidence from psychology and behavioral economics suggests that people exhibit systematic biases in predicting future utility (see DellaVigna (2009) for a review). One such bias, captured in such cliches as "sleep on it," or "never go grocery shopping on an empty stomach," is the idea that current conditions have an oversized influence on intertemporal decision making. Empathy gaps (Lowenstein (2005); Ariely and Loewensteing (2006), projection bias (Loewenstein, O'Donoghue, and Rabin (2003)), salience (Bordalo, Gennaioli, and Shleifer (2013); Koszegi and Szeidl (2013)), and present bias (Laibson (1997); O'Donoghue and Rabin (1999)) are examples of mechanisms for why such might be the case.

In this paper, we use transaction-level data from one of the largest insurance companies in China to examine the role of air pollution in determining an individual's decision to purchase and cancel health insurance . The insurance policies do not cover pre-existing conditions and have a 180 waiting period before coverage begins, so the value of the policy is a function of the premiums and the probability of illness in future periods. Given the high variability in day-to-day air pollution levels, daily air pollution levels should essentially be a non-factor in a fully rational person's decision to purchase or cancel health insurance.

We find that both the purchase and cancellation of insurance contracts are significantly influenced by idiosyncratic variation in the daily levels of particulate matter (PM2.5). Specifically we find that when air pollution is high individuals are more likely to purchase insurance contracts, and that insurance contracts purchased on high pollution days are also more likely to be canceled during the government mandated 10-day "regret period" during which individuals can costlessly cancel their insurance contracts. This cancellation effect is negatively related to air pollution during the cooling-off period, 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 improve during the cooling-off period relative to the day 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 as measured by Air Quality Index (AQI) leads to a 7% increase in the number of insurance contracts sold that day. This effect of pollution on sales is non-linear, with measurable effects occurring at AQI levels associated with adverse acute health effects. We also find that a one standard deviation decrease in AQI during the cooling-off period *relative* to the order-date leads to a 4.1% increase in the rate of insurance contracts that are canceled. In contrast, AQI levels have no impact on either sales or cancellations of other insurance products the company sells. In addition, the results of a distributed lag model show that pollution affects the aggregate level of insurance contracts sold as opposed to causing temporal substitution across days.

We hypothesize that these results are driven by projection bias as formalized in Loewenstein, O'Donoghue and Rabin (2003). Projection bias posits that individuals exaggerate the degree to which their future tastes will resemble their current tastes. In this context, the poor health brought on by high levels of air pollution causes individuals to overestimate the probability of poor health in the future. If such is the case, projection bias would predict that when air pollution is high individuals will be more likely to purchase insurance contracts. In addition, since projection bias affects individuals both when the purchase and cancellation decisions are made, lower air pollution levels during the cooling-off period should lead to higher cancellation rates. Our results are consistent with both predictions.

Our results are important for several reasons. First, insurance is one of the world's largest industries, eclipsed only by real estate, finance and the public sector. In 2014 insurance premiums in the U.S. exceeded $2 trillion, with health insurance premiums in the private insurance market alone amounting to $839 billion. Health care spending is also a huge part of the economy, accounting for 5.7% of China's GDP and 9% of GDP for all O.E.C.D countries (OECD 2015). And 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 of the importance role of projection bias in the market for health insurance in China, and potentially insurance markets more widely.

Second, while projection bias has received significant attention in both the economics and psychology literature, there is only recently some evidence that projection bias influences demand for real goods and services. The most convincing prior evidence of projection bias in a real-world market remains the first paper to document projection bias in a real-world market: Colin, O'Donoghue and Volgelsang (2007). Their paper convincingly shows that catalog 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 finds only mixed evidence regarding the impact of return-date temperature on returns. Simonsohn (2009) and Busse, Pope, Pope and Silva-Risso (2014) show that weather also affects college enrollment and the type of automobile purchased respectively. Busse et al. (2014) conclude that their results are incompatible with standard, rational agents but consistent with both projection bias and salience. To our knowledge, these are the only other papers to show evidence that projection bias affects the demand for real goods and services.

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: labor productivity (Graff Zivin and Neidell (2012), Chang, Graff Zivin, Gross, and Neidell (2014, 2016), Li, Liu and Salvo (2015)), student test scores (Lavy, Ebenstein, and Roth (2014)), and crime (Herrnstadt and Muehlegger (2015)).

Finally our results also provides direct evidence in support of the hypothesis put forth in Loewenstein, O'Donoghue and Rabin (2003) that "cooling-off laws" might be effective "as devices for combating the effects of projection bias." Moreover, they suggest that the efficacy of cooling-off periods in combating projection bias is determined in part by the correlation in the driving state variable (i.e., projection bias due to slowly moving state variables could potentially benefit from longer cooling-off periods.)

The paper proceeds as follows. The subsequent section presents a simple model of insurance orders and cancellations in the presence of projection bias. In Section 3, we present some basic information on the impact of air pollution on health. Section 4 describes the data used in the paper along with our empirical strategy. Section 5 presents our main results on the effect of PM2.5 on the decision to purchase and cancel insurance contracts. Section 6 concludes.

2 Projection Bias

Projection bias is the tendency for individuals to exaggerate the degree to which their future tastes will resemble their current tastes. Lowenstein, O'Donoghue and Rabin formalize this idea with a model in which an agent's utility is given by

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

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 an agent has $\alpha = 0$, she accurately predicts her future utility. In contrast if $\alpha > 0$, then she mis-predicts her future utility as a convex combination of her true future utility from c and the utility she would receive from c given her current state s'.

As a simple, illustrative case of the influence s_0 can have on the demand for health insurance, we assume a particularly simple utility function of the form

$$\tilde{u}(c_t, s_t|s_t) = (1-\alpha)B(s_t) + \alpha B(s_0) - p, \tag{2}$$

where s_t is a measure of how sick an individual is at time t, B is non-zero, increasing function that represent the per period benefit provided by the insurance policy, and p is the per period insurance premium.

Mapping this utility function to the case of health insurance choice is straightforward. Consider an individual who purchases a health insurance I at time t with a policy period equal to T, and let s_τ represent her expected future health. Conditional on purchase, she can costlessly cancel her policy at time t=1 (e.g., the cooling off period), with premiums and coverage to begin at time t=2 (i.e. a one period "waiting period"). Her perceived utility from purchasing health insurance is then given by

$$\tilde{U}^t(\tilde{u}_{t+2}, ..., \tilde{u}_{t+T}|s_t) = \sum_{\tau=2+t}^{T+t} \delta^{(\tau-t)}[(1-\alpha)B(s_\tau) + \alpha B(s_t) - p]. \tag{3}$$

This simple framework illustrates the influence s_t can have on the demand for insurance. While 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 today. So for a given price p, demand will be higher when individuals are more unwell.

Next, consider the behavior of an individual who has purchased insurance during the cooling off period ($t = 1$). Again, if she is rational, her predictions are not affected by her current health s_1, but if she is affected by projection bias, her predictions regarding the utility from insurance will be biased by her current health. She will then choose to cancel her insurance if $\delta\tilde{U}^0(c_{t+2}, ..., c_{t+T}|s_1) < 0$. So for a high enough p, there will be an $\underline{s} < s_0$ such that if $s_1 < \underline{s}$, she will cancel her insurance in period 1. That is an individual with projection bias will cancel her insurance if her health level during the time she is making the decision to cancel her policy is sufficiently high *relative* to her purchase day health.

These results generate a pair of testable predictions:

(1) Order-date health shocks (s_0) will increase sales of health insurance.

(2) If individuals feel healthier during the cooling off period *relative* to the order-date (i.e., $s_1 < s_0$), they are (weakly) more likely to cancel their insurance policy.

3 Air pollution and You

Particulate matter (PM) consists of solid and liquid particles in the air that can range considerably in size. In response to a growing body of toxicological and epidemiological evidence suggests that exposure to PM2.5 significantly harms health (see EPA (2004) for a review) public awareness and regulation has evolved to focus on this form of air pollution. PM2.5 consists of particles less than 2.5 micrometers across, small enough to ether the bloodstream through the lungs. In addition, its diminutive size allows it to easily enter buildings, with penetrating rates of over 70% (Tahtcher and Layton (1995)). Thus unlike most other forms of air pollution, which either remain outdoors or rapidly break down indoors, PM2.5 cannot be avoided by remaining indoors.

A large body of toxicological and epidemiological evidence suggests that exposure to PM2.5 harms health (see EPA, 2004 for a comprehensive review). The health risks effects arising from exposure to PM2.5 arise primarily from changes in pulmonary and cardiovascular functioning (Seaton et al., 1995). They may manifest themselves in respiratory episodes, such as asthma attacks, and cardiovascular events, such as heart attacks, that lead to hospitalizations and mortality (Dockery and Pope, 1994; Pope, 2000). They also lead to more subtle effects, such as changes in blood pressure, irritation in the ear, nose, throat, and lungs, and mild headaches (Pope, 2000; Ghio et al., 2000; Auchincloss et al, 2008). Importantly for our empirical design, some response to high level of PM2.5 is immediate (e.g., watery eyes, scratchy throat, shortness of breath) and more symptoms can arise in as little as a few hours after exposure. Figure 1 describes the PM2.5 levels as expressed in Air Quality Index (AQI) levels and the relevant heath effects as per the U.S. EPA.

Our empirical strategy for testing the predictions of projection bias described in the previous section is to use the daily level of PM2.5 in a city as a health shock to the city's population. Specifically we assume that PM2.5 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 PM2.5 level serves as a proxy for the average contemporaneous health it's citizens (i.e., $AQI \sim s_t$).

While day-to-day variation in PM2.5 levels is quite high,[1] PM2.5 levels generally follow a cyclical pattern correlated with other environmental factors more generally (e.g., weather). In Beijing, for example, PM2.5 tends to be lower during the rainy season when precipitation serves to wash away airborn pollutants and higher in winter months when people burn more fossil fuels for warmth (see Figure 2). Similar to temperature, while the current AQI level provides some additional information regarding the AQI levels one can expect in the near future, it provides essentially zero additional information about the AQI levels one should expect 180 days from now. For a rational agent, this would mean that the current AQI will not effect how much she values of health insurance. In contrast an agent with projection bias will, all else equal, value insurance more when AQI is high leading to a positive relationship between AQI and demand as described in the previous section.

4 Data

The data were obtained from three sources: a large Chinese company that sells a variety of insurance products, U.S. State Department, and 15 Tianqi. From the insurance company we have detailed information on over one million insurance contracts. These contracts represent the universe of health insurance policies, along with 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. Due to the sensitive nature of the sales data, and to ensure the anonymity of the firm providing the data, we cannot reveal the identities of the cities in our sample or sales patterns for the various insurance products in our sample. For each insurance policy sold, the firm provided us with the date of purchase, city of residence, contract length, whether the insurance is both for oneself or for someone else (e.g., a family member), and some basic demographic information for the person covered by the insurance policy. The firm also provided cancellation information for these policies through the end of 2014.

[1] The within-city day-to-day correlation in AQI levels is less than 0.5.

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, around 90 percent of the population has health insurance through the government. This coverage is accomplished through three insurance programs, 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 programs is though 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 a rapidly growing market in China, especially among China's growing middle class. This form of insurance is considered especially important to cover expenses due to significant adverse health events like cancer. The insurance contracts in our data consist of this types of 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 prevent the covered individual from receiving benefits if their 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 as collected by air quality monitors located on U.S. Embassy compounds in the relevant cities. The PM2.5 level is expressed in terms of an air quality index (AQI) following the U.S. EPA formula (EPA (2006)). The AQI values are designed to help inform health-related decisions by mapping pollution levels to round number breakpoints corresponding to categories of health impact (see Figure 1).[2]. While we cannot provide full details regarding the pollution levels in our sample as they could be used to determine the identity of the cities in our sample, the mean daily AQI in our composite sample is 125.6 with a standard deviation of 98.4. While this level of air pollution is typical for a large Chinese city, it would be considered

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

quite high in the U.S.[3]

Finally we retrieved weather information from each city from 15 Tianqi, a Chinese weather website. This data included daily low and high temperatures, precipitation, and a dummy 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.[4] As shown in Table I, this left us with a sample of 579,303 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.[5] Approximately half the time, an individual is purchasing insurance for him or herself. In the other half, an individual is purchasing 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 cooling-off period is 2.8%.

5 Empirical Results

5.1 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

$$Log(Insurance_{jt}) = \beta AQI_{jt} + X_{jt}\gamma + D_{jt} + \epsilon_{jt}, \tag{4}$$

where $Insurance_{jt}$ is the number of insurance contracts sold by the firm to residents of city j on date t, AQI_{jt} is the high hourly AQI in city j on the over a two-day window

[3]While not exactly comparable since different technologies are used to measure air pollution at different temporal resolutions, as an illustrative example the EPA reports that the median AQI in Cambridge, MA and Los Angeles, CA in 2015 was 46 and 77 respectively.

[4]One date observation with an AQI of zero and two date observations with AQI > 800 were dropped from the sample.

[5]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.

consisting of date t and $t-1$. This allows for the purchase decision to have been made the day before purchase, which is possible since pollution tends to peak in the evening when the firm is closed unable to take customer orders.[6] The vector X_{jt} consists of a quadratic function of high temperature and dummies 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.

Our 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 percent change in total number of insurance contract sold on a given day by a one unit increase in AQI.

The results of estimating Equation 4 is presented in Table II. Column 1 indicates that an one unit increase in daily AQI generates with a 0.072% increase in daily sales, or that a one standard deviation increase in daily AQI leads to a 9.0% increase in daily sales. For column 2 we allow AQI to have a non-linear effect on sales by re-estimating Equation 4 with indicator variables corresponding to the different EPA categories for pollution levels in place of a linear measure of AQI (see Figure 1). The withheld category is AQI of between 0 and 50, corresponding to "Good" air quality. The results indicate that the effect of AQI on sales only become significant once 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, while the coefficient for the "Unhealthy for Sensitive Groups" level of PM2.5 is around 2/3rds as large as the coefficient for "Unhealthy" but not statistically significant at conventional levels. AQIs of between 150-200 ("Unhealthy"), 200-250 ("Very Unhealty"), and greater than 300 ("Hazardous") are associated with increases in daily sales of 16.8%, 16.8% and 23.4% respectively compared to days with ab AQI of less than 50.

In column 3, we re-run 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 neighbor, then regress that city's daily sales against both that city's pollution, and the pollution of the matched city. We see that controlling for the AQI of the nearest city

[6]Using either the one day AQI for date t or $t-1$ produces similar results.

slightly reduces the size of the coefficient from 0.00072 to 0.00066, and that the coefficient for other city AQI is both small and statistically insignificant. Together the results in these first three columns indicate that the air pollution in ones immediate vicinity increases demand for health insurance, and that this increased demand effect occurs at the pollution levels associated with noticeable health effects.

Finally in column 4 we repeat the regression shown in column 1, but with the dependent variable the number of other insurance contracts sold by the company. This category consist primarily of term life and personal accident insurance. Here the coefficient of interest is small and statistically insignificant, indicating that air pollution is not a significant driver of demand for other insurance products sold by the firm.

5.2 Air Pollution and Temporal Substitution

While the results in the previous section demonstrate that pollution affects the demand for insurance, it does not answer the question of whether the increase represents a true increase in aggregate demand or reflects pollution's effect on intertemporal substitutions. That is pollution may not generate additional demand for insurance, but simply shift when a person purchases insurance.

To assess whether our results are driven by intertemporal substitution with respect to daily pollution, we estimate a distributed lag model. Specifically we re-run Equation 4 with N daily lags of AQI and weather added to the estimating equation:

$$Log(Insurance_{jt}) = \beta AQI_{jt} + \sum_{\tau=1}^{N} \beta_\tau AQI_{j,t-\tau} + \sum_{\tau=0}^{N} X_{j,t-\tau}\gamma_\tau + D_{j,} + \epsilon_{jt}. \tag{5}$$

Including lagged pollution variables in our regression allows us to test whether pollution in the days leading up (or following) to the day of purchase affects the impact of contemporaneous pollution on purchase decisions. For example, a negative coefficient on the 5th day lagged pollution measure would indicate both that high pollution 5 days ago leads to lower sales today and that high pollution *today* leads to lower sales 5 days in the future.

Thus the sum of the lagged coefficients are 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. [7] To test whether the increase in demand we measured in the previous section is due to displacement, we can test the null hypothesis that the sum of $k \leq N$ coefficients for the lagged pollution variable is equal to the negative of the current period coefficient β.

Figure 2 present the results of this analysis by plotting the estimated coefficients and 95% confidence intervals from estimating Equation 5 for a period of 6 weeks ($N = 42$). As shown in the figure, while current period pollution has a large, positive and statistically significant impact on the demand for health insurance contracts, the coefficients for the lagged pollution are smaller and never statistically significant. Moreover the fact that most coefficients tend to be positive, even if not statistically significantly so, suggest that high pollution in the recent past leads to higher insurance sales today. The current day pollution coefficient β in this regression equals 0.00081 with a standard error of 0.00024, a value which is slightly larger than the coefficient of 0.00072 from Table II. Testing the null hypothesis that the sum the coefficients for first k lags is equal to the negative of the current day coefficient β, we find we can reject the null hypothesis with a p-value < 0.001 for k equal to 7,14,21,28,35 or 42 days. These results indicate that the increase in daily sales generate by air pollution can be interpreted as an increase in the aggregate demand for insurance, and due to changes in the timing of insurance purchases.

5.3 Effect of PM2.5 on Cancellations

In addition to total demand for insurance, we also examine insurance cancellation rates. For this analysis, we start with the base regression specification

$$Cancel_{ijt} = \beta_0 AQI_{p,ijt} + f(AQI_{ij,1}, ..., AQI_{ij,11}) + C_i b + X_{jt}\gamma + D_{jt} + \epsilon_{jt}, \qquad (6)$$

[7] See Jacob, Lefgren and Moretti (2007), Deschenes and Moretti (2009) and Busse, Pope, Pope and Silva-Risso (2014) for more detailed discussion on the methdology used here.

where $Cancel_{ijt}$ is a dummy that equals 1 if the contract is canceled by the purchaser within 11 days of purchase.[8] We drop observations if the policy was canceled within 24 hours of purchase (212 same day cancellations, 466 next day cancellations) as in such cases there is considerable overlap in the AQI level during when the purchase and cancellation decisions are made.[9] AQI_{ijt} is the previously used measure air pollution on the date of purchase and $f(AQI_{ijt+1,\ldots,ijt+11})$ is a function of the 11 daily leads of the pollution variable. C_i includes controls for policy characteristics: the age and gender of the policy holder, 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 fixed effects designed to capture trends both within week and over time.

We use four different specifications to capture the effect of pollution during the cooling off period (CoP) on cancellation rates. Our first specification directly tests the prediction that projection bias's effect on cancellations operates via differences in 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 cooling off period relative to order-date AQI (Relative AQI). That is we run the regression

$$Cancel_{ijt} = \beta(Relative\ AQI_{ijt}) + C_i b + X_{jt}\gamma + D_{jt} + \epsilon_{jt}, \tag{7}$$

where

$$Relative\ AQI_{ijt} = (\sum_{\tau=1}^{11} \frac{1}{11} AQI_{ij,t+\tau} - AQI_{p,ijt}). \tag{8}$$

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

The second specification includes controls for both the level of the order-date AQI and

[8] Although the legally mandated cooling-off period is 10 days, the firm is not strict in enforcing the 10 day rule. Consequently a significant number of cancellations 11 days after purchase. Limiting the analysis to a 10 day post-purchase period generates similar results.

[9] Including these observations does not materially affect the regression results.

the average AQI during the cooling off period (*CoP AQI*). This specification is essentially identical to that used in Conlin, O'Donoghue and Vogelsang (2007) as a direct test of projection bias. In cases where one or more of the daily pollution measures during the CoP was not available, the CoP AQI was calculated excluding the missing value.

The third specification is a variant on the second specification, but replaces the *CoP AQI* with the 11 leads of pollution as separate regressors and then sum the 11 resulting coefficients. That is we set $f(AQI_{ijt+1,...,ijt+11}) = \sum_{\tau=1}^{11} \beta_\tau AQI_{ij,t+\tau}$, and report $\sum_{\tau=1}^{11} \beta_\tau$ as the effect of pollution during the cooling off period on insurance cancellations. In approximately 30 thousand cases, one of the lead pollution measures were missing and the insurance contract was not included in the regression. [10] Subject to the linear functional form assumption, this provides us with a measure of the cumulative effect of daily pollution during the cooling off period on cancellations. For our second specification, we first calculate the average pollution levels for the 11 days post purchase. We then include this average pollution measure as our control for pollution during the cooling off period: $f(AQI_{ijt+1,...,ijt+11}) = \beta^{CoP} \sum_{\tau=1}^{11} \frac{1}{11} AQI_{ij,t+\tau}$.

For our final specification, we utilize a dummy variable to indicate whether air pollution during the time when an individual makes her decision to cancel an insurance policy is *lower* then the air pollution level during the time when she makers her decision to purchase said insurance policy. Specifically $f(AQI_{ijt+1,...,ijt+11})$ is an indicator variable equal to one if $Cop\ AQI_{ij,0} < Order - dateAQI$. As with our measure of average AQI, in cases when one of the lead pollution measures was not available, the comparison was done to the exclusion of the the missing value.

Table IV reports the marginal effects at the sample mean associated with estimating Equation 6. Column 1 presents the results of regressing Relative AQI on cancellations so that the coefficient of interest represents the effect AQI during the CoP normalized such that order-date AQI=0. Here we find a negative and statistically significant relationship between Relative AQI and cancellations, indicating that decreases in AQI relative to order-

[10] Replacing missing observations with a value interpolated from the nearest two observations leads to essentially identical results.

date AQI leads to increases in the probability of cancellation. Specifically for ever one unit (standard deviation) *decrease* in AQI relative to order-date AQI, the probability of cancellation increases by 0.001% (0.10%).

In column 2, we include both order-date AQI and the average AQI during the cooling off period (CoP AQI) as regressors. We find that higher order-date AQI lead to a positive and statistically significant increases cancellations. 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 cooling off period is negative and statistically significant, with a one unit CoP AQI decreasing the probability of cancellation by 0.024%, indicating that individuals are less likely to cancel if pollution is high in the window during which they can choose to cancel their policy.

Column 3 repeats this analysis, but replaces the average CoP AQI with disaggregated daily measure of daily AQI, and finds essentially the same pattern of results as column 2: higher order-date AQI lead to a positive and statistically significant increases cancellations, while the aggregate effect of daily air pollution levels during the cooling off period is negative and statistically significant.[11]

Finally in column 4, we repeat the analysis shown in column 2 but with a dummy for whether the average of daily AQI is lower during the cooling off period relative to purchase-date AQI. Unlike the columns 2 and 3, here we find that order-date AQI no longer predicts increased probability of cancellation. Instead we find that the effect of air pollution on cancellations depends solely on whether or not air pollution is lower during the period in which the purchaser can decide to cancel her policy relative purchase day air pollution. Specifically if $Cop\ AQI_{ij,0} < Order - dateAQI$, the probability that a contract is canceled increases by 0.19%, or a 7.25% increase in the cancellation rate. This suggest that the impact of air pollution on cancellation rates is driven by *relative* differences. That is the AQI during the time the decision to cancel is made matters only in how it differs from the

[11] We reject at a p-value < 0.01 that the sum of the individual lead coefficients are equal to zero.

AQI the decision maker faced when decision to purchase insurance in the first place.

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

Column 1 includes the same controls as the regressions in Table III, while column 2 includes interaction terms between the weather controls, contract characteristics and a dummy for non-health insurance policies to allow those characteristics to have differential effects for the health vs. other insurance policies. For both specifications, the main effect of the difference in AQI between the order-date and the CoP remains negative and statistically significant. The interaction term though is positive, statistically significant, and only slightly smaller in magnitude than the main effect. Thus the marginal effect of the for other insurance types has a magnitude close to zero and statistically insignificant with p-values of 0.54 and 0.46 for columns 1 and 2 respectively.

5.4 Effect of Air Pollution on Insurance Contract Characteristics

We next examine the effect on pollution on the characteristics of the insurance contracts purchased. The price of insurance contracts are not individually negotiated, but instead is set by the firm and change infrequently. Prices therefore should be unaffected by idiosyncratic variation in both daily air pollution levels and the demand for insurance. As any such changes in insurance characteristics would indicate either changes in the composition of who purchases insurance, or what kinds of insurance features pollution causes individuals to value more.

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

$$C_{ijt} = \beta AQI_{jt} + X_{jt}\gamma + D_{jt} + \epsilon_{jt}. \tag{9}$$

Here the dependent variable C_{ijt} is a characteristic of insurance plan i sold in city j on date t. As in Equation 4, 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 dummies for precipitation and snowfall, and D_{jt} are day-of-week, month-of-year*city, and year*city fixed effects.[12] All standard errors are clustered at the city*date level.

The results of estimating Equation 9 is presented in Table III. Columns 1 and 2 present the results from an OLS regression where 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 one if purchaser and covered individuals are the same (3), if the covered individual is female (4), and 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, 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 II, we can rule out AQI having an economically meaning effect on these contract characteristics. And while the coefficient for column 4 is statistically significant, the effect size itself is quite small with a one unit (standard deviation) increase in AQI leads to a 0.007% (0.88%) increase in the share of contracts that insure females off a baseline of 55%. Overall these results suggest that while air pollution significantly increases the demand for insurance, it does not appear to change the type of insurance product purchased.

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

6 Conclusion

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 cooling off period are lower than that on the order-date. 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. These results not only provide strong empirical evidence for projection bias, but suggests that projection bias may be an important factor in understanding the demand for insurance more generally.

References

Ariely, D. and G. Loewenstein, (2006). "The heat of the moment: The effect of sexual arousal on sexual decision making," *Journal of Behavioral Decision Making,* 19(2), 87-98.

Bordalo, P., N. Gennaioli, and A. Shleifer, (2013). "Salience and Consumer Choice," *Journal of Political Economy,* 121(5), 803-843.

Busse, M., J. Pope, D. Pope, and J. Silva-Risso (2015). "The Psychological Effects of Weather on Car Purchases," *Quarterly Journal of Economics*, 130(1): 371-414.

Conlin, M., T. O'Donoghue, and T. Vogelsang, (2007). "Projection Bias in Catalog Orders," *American Economic Review*, 97(4): 1217-1249.

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

Deschenes, O. and E. Moretti, (2009). "Extreme Weather Events, Mortality, and Migration," *Review of Economics and Statistics*, 91(4), 659-681.

Dockery D. W., and C. A. Pope, (1994). "Acute respiratory effects of particulate air pollution." *Annual Review of Public Health,* 15, 107-13.

Environmental Protection Agency (2004). Air Quality Criteria for Particulate Matter. National Center for Environmental Assessment. Research Triangle Park, NC

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

Jacob, B., L. Lefgren, and E. Moretti, (2007). "The Dynamics of Criminal Behavior: Evidence from Weather Shocks," *The Journal of Human Resources*, 42(3), 489-527.

Koszegi, B. and A. Szeidl, (2013). "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics,* 128(1), 53-107.

Laibson, D., (1997). "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics,* 112(2), 443-77.

Lavy, V., A. Ebenstein, and S. Roth, (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(4S): S49-S56.

Loewenstein, G., T. O'Donoghue, and M. Rabin, (2003). "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics,* 118: 1209-1248.

O?Donoghue, T. and M. Rabin, (1999). "Doing It Now or Later," *American Economic Review,* 89(1): 103-24.

Seaton, A., W. MacNee, K. Donaldson, and D. Godden, (1995). "Particulate air pollution and acute health effects." *Lancet* 8943: 176-178.

Simonsohn, U (2010). "Weather to Go to College," *Economic Journal*, 120(543), 270-280.

Thatcher, T. L., and D. W. Layton, (1995). "Deposition, resuspension, and penetration of particles within a residence." *Atmospheric Environment,* 29(13), 1487-1497.

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

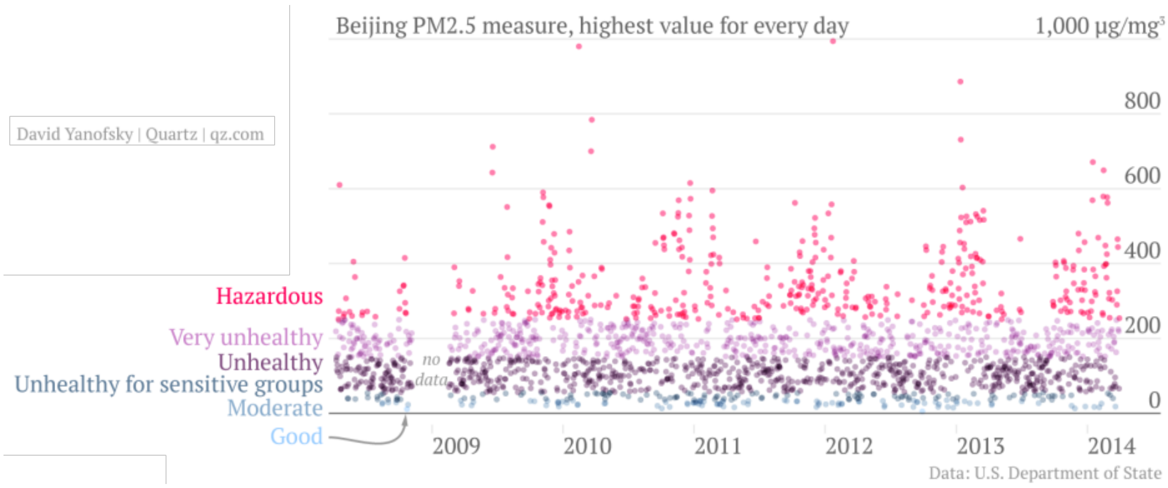

Figure 2. Daily PM2.5 levels as measured by the U.S. Embassy in Beijing.

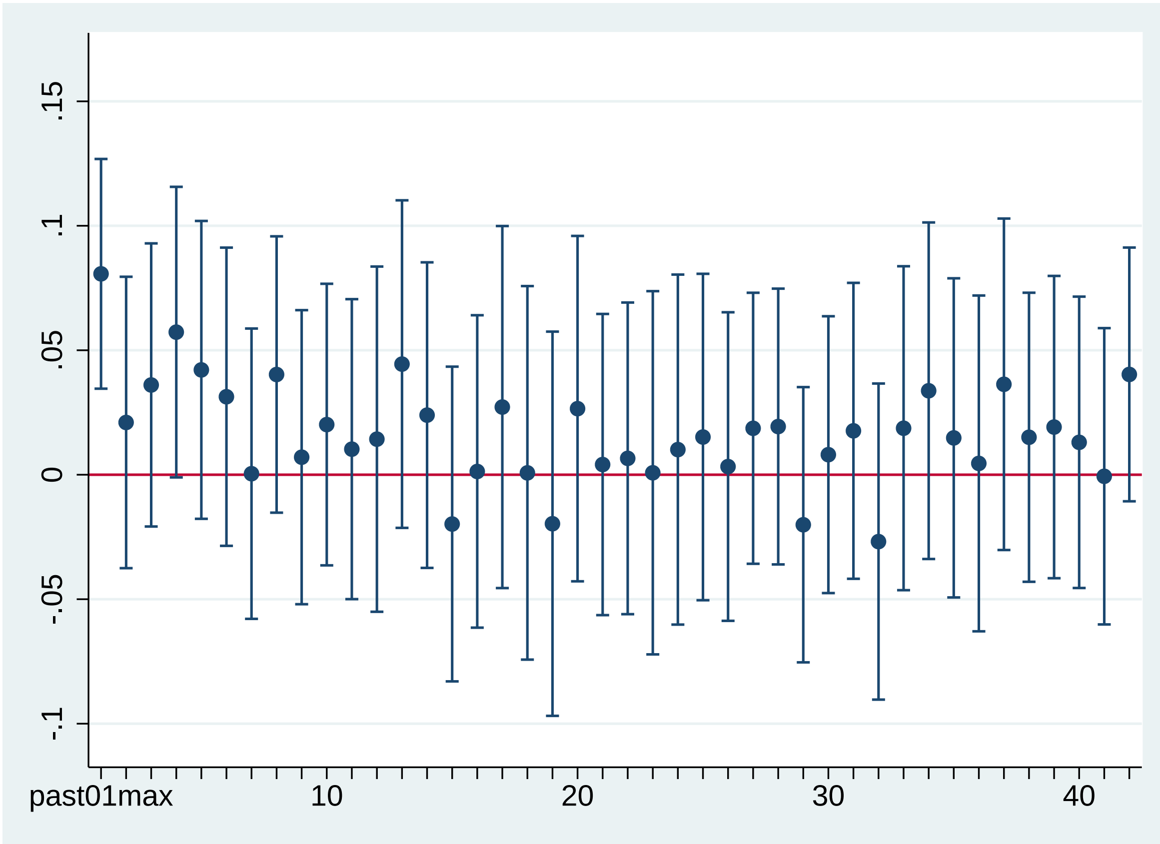

Figure 3. Coefficient values and 95% confidence intervals for the effect of contemporaneous and lagged AQI/100 on daily insurance sales.

Table I
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	9313	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. Information on cancellations was not provided for contracts sold in the most recent year of the data (2015).

Table II
The Effect of Pollution on Insurance Sales

Dependent Variable: Log(Number of Contracts Sold)				
Insurance Type		Health		Other
$AQI_{PM2.5}$	0.00072**		0.00066**	-0.00013
	(0.00019)		(0.00020)	(0.00023)
$AQI_{PM2.5}$ 50-100		0.0116		
		(0.0715)		
$AQI_{PM2.5}$ 100-150		0.1147		
		(0.0759)		
$AQI_{PM2.5}$ 150-200		0.1681*		
		(0.0825)		
$AQI_{PM2.5}$ 200-300		0.1680*		
		(0.0849)		
$AQI_{PM2.5}$ 300+		0.2340*		
		(0.0996)		
Other City $AQI_{PM2.5}$			0.00007	
			(0.00023)	
Temperature	-0.0191	-0.0189+	-0.0196+	0.0221*
	(0.0111)	(0.0111)	(0.0119)	(0.0224)
Temperature^2	0.0008**	0.0008**	0.0009**	-0.0006*
	(0.0072)	(0.0003)	(0.0003)	(0.0003)
Rain	-0.0391	-0.0355	-0.0363	0.1006
	(0.0755)	(0.0756)	(0.0763)	(0.0658)
Snow	-0.2059	-0.1943	-0.2078	-0.2639
	(0.1700)	(0.1707)	(0.2242)	(0.1677)
Adjusted R-squared	0.481	0.481	0.478	0.483
Observations	2,573	2,573	2,453	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 neighbor. Insurance type "Other" consists of personal accident and term-life insurance policies. All regressions included dummy variables for day of week, city*month and city*year. Standard errors are clustered on date.
+ significant at 10%, * significant at 5%, ** significant at 1%.

Table III
The Effect of Pollution on Cancellations

Dependent Variable: Indicator equal to 1 if contract is cancelled				
% of Contracts Cancelled	2.62%	2.62%	2.55%	2.62%
Relative AQI	-0.00110**			
	(0.00041)			
Order-date AQI		0.00087*	0.00100*	0.00003
		(0.00044)	(0.00048)	(0.00053)
CoP AQI		-0.00243**		
		(0.0092)		
$\sum_{\tau=1}^{11} \beta_{AQI,\tau}$			-0.00236**	
			(see notes)	
1(CoP AQI<Order-date AQI)				0.1929*
				(0.0866)
Log(Term Length)	-0.487**	-0.486**	-0.482**	-0.487**
	(0.175)	(0.018)	(0.018)	(0.0325)
Log(Age)	0.365**	0.365**	0.337**	0.366**
	(0.032)	(0.033)	(0.033)	(0.032)
Self	1.060**	0.058**	1.026**	1.059**
	(0.076)	(0.076)	(0.078)	(0.076)
Female	0.125**	0.125*	0.106+	0.123*
	(0.057)	(0.057)	(0.058)	(0.056)
Adj. R-squared	0.050	0.050	0.052	0.050
Observations	405,599	405,599	375,841	405,599

Notes: For each column, the dependent variable is whether an insurance contract is canceled during the cooling-off period. All coefficients represent the marginal effects from a probit regression. *Relative AQI* is the average AQI during the cooling off period minus the order date AQI. *CoP AQI* is the mean value of $AQI\ PM_{2.5}$ during the cooling off period ($\frac{1}{11}\sum_{\tau=1}^{11} AQI_\tau$). $\sum_{\tau=1}^{11} \beta_{AQI,\tau}$ is the sum of the coefficients for the 11 daily leads of the pollution variable $AQI_0\ PM_{2.5}$. We can reject at p-value=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 IV
Cancellations Including Non-Health Insurance

Dependent Variable: Indicator equal to 1 if contract is cancelled

% of Contracts Cancelled	5.36%	5.36%
Relative AQI	-0.00286**	-0.00264**
	(0.00082)	(0.00079)
(Relative AQI)(Other)*	0.00257**	0.00229**
	(0.00078)	(0.00076)
Other	0.02618**	-0.00856**
	(0.00078)	(0.00320)
Log(Term Length)	-0.742**	-0.943**
	(0.034)	(0.031)
Log(Term Length)*Other		0.348**
		(0.056)
Log(Age)	0.514**	0.583**
	(0.041)	(0.057)
Log(Age)*Other		-0.102+
		(0.058)
Self	2.148**	1.829**
	(0.098)	(0.119)
Self*Other		0.632**
		(0.161)
Female	0.282**	0.198*
	(0.063)	(0.095)
Female*Other		0.119
		(0.101)
Adj. R-squared	0.059	0.060
Observations	890,247	890,247

Notes: For each column, the dependent variable is whether an insurance contract is canceled during the cooling-off period. All coefficients represent the marginal effects from a probit regression. *Relative AQI* is the average AQI during the cooling off period 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%.

Table V
Pollution and Insurance Contract Characteristics

	Term Length	*Age*	*Self Purchase*	*Female*	*Female & Self*
AQI $PM_{2.5}$	0.00001	0.00004	0.00007	0.00007*	0.00002
	(0.00006)	(0.00007)	(0.00006)	(0.00003)	(0.00005)
Temperature	-0.0024	-0.0102	-0.0015	-0.0038+	-0.0072*
	(0.0036)	(0.0044)	(0.0034)	(0.0020)	(0.0032)
Temperature^2	0.0002*	0.0002*	-0.0000	0.0001*	0.0002*
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)
Rain	-0.0230	-0.0050	0.0253	0.0099	-0.0011
	(0.0214)	(0.0253)	(0.0185)	(0.0122)	(0.0186)
Snow	-0.0785	-0.0210	0.0484	0.0526*	0.0308
	(0.0511)	(0.0594)	(0.0407)	(0.0262)	(0.0400)
Adj. R-squared	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%.