Do Homeowners Care About Air Quality? Estimating the Effect of Poor Air Quality on Home Prices Using Wildfire Smoke Data

Literature Review

Tristan Misko

20 October 2021

1 Research Design

My paper uses the variation introduced by wildfire smoke to instrument for air quality in order to measure the causal effect of air quality on housing prices. Using satellite imagery data to measure wildfire smoke (the unit being days of smoke coverage for a particular county), I can instrument for air quality to obtain exogenous variation. Because the wildfire smoke distribution has been changing heterogeneously both spatially and temporally over the last two decades, I can compare counties which have received increases in air pollution due to wildfire smoke plumes with those which have not, as well as comparing time series trends within counties. Comparing housing price changes across these groups will yield causal estimates.

2 Relationship to the Literature

A 2005 paper by Chay and Greenstone uses a spatial hedonic approach based on a United States Clean Air Act policy which implemented stricter regulations on counties which failed to meet pre-specified particulate pollution targets. They estimate that a 1% increase in particulate matter pollution concentration decreases home values by approximately 0.2% to 0.35% (Chay and Greenstone, 2005). The fundamental idea of Chay and Greenstone's study is to use an instrument for air quality, namely non-attainment status under the Clean Air Act, to get plausibly exogenous variation in air quality; my strategy is similar in that wildfire smoke introduces plausibly exogenous variation in air quality that I will use in my

estimation strategy. My wildfire smoke instrument is an improvement on the non-attainment status instrument for a few reasons. Non-attainment status is a dummy, while wildfire smoke is a continuously valued variable, which offers more variability to exploit when running a first stage regression. Furthermore, one must treat selection bias arguments very seriously in the non-attainment case: perhaps there are systematic unobserved factors which simultaneously affect whether a county is a non-attainment county and which also affect housing quality, potentially introducing omitted variables bias. Wildfire smoke, treatment by which is largely determined by wind patterns, admits no such arguments.

Kim et al. use a spacial hedonic approach at a very local level to estimate the effects of air pollution on housing prices in Seoul, South Korea, associating a 4% air quality increase with a 1.4% housing price increase (Kim et al., 2003). They use a relatively small sample of households with detailed housing price data across all of the major districts of Seoul, controlling for neighborhood income and housing characteristic factors, along with a spatially interpolated local air pollution dataset. Their estimates use spatially lagged variables as instruments, which is an econometric technique to overcome the lack of exogenous variation present in the cross-sectional price data but which is subject to many technical challenges (*ibid*). Zabel and Kiel employ a similar strategy in four U.S. cities, gathering data about individual housing units' prices and characteristics, controlling for neighborhood factors, and estimating a hedonic model for housing prices in Chicago, Denver, Philadelphia, and Washington D.C. They find a small, significant negative relationship between air pollution and housing prices in two of them (Zabel and Kiel, 2000). Instead of relying on sophisticated econometric techniques to overcome endogeneity in a cross-sectional sample as in both of these papers, my paper uses plausibly exogenous variation in air quality measured over multiple periods, which I believe is a stronger design. My design applies to a more general setting than the estimates of these papers, which are city specific and may therefore lack the external validity that estimates generated from county-level data from across the U.S. would carry.

Borgschulte et al. use wildfire smoke to instrument for air quality in their 2018 working paper estimating the causal effect of air quality on the labor market, particularly on employment and adaptation costs. Although their outcome variable of interest is unrelated to housing prices, the machinery of their research design, which they claim is novel in their paper, is quite similar to the design I propose to use. They use daily air quality data from the EPA, a wildfire smoke dataset based on satellite imagery to instrument for air quality, and identify their observations at the county level, all of which I plan to use in my research design as well (Borgschulte et al., 2018). Their paper is focused on estimating the effect of short term shocks of bad air quality on the labor market, which is different from my medium to long-run focus on housing prices as effected by changing trends in wildfire smoke.

In summary, my paper can be understood as applying a research design to instrument for air quality using wildfire data similar to that deployed by Borgschulte et al. to the domain of housing prices. Previous estimates of the causal effect of air quality housing prices have often been local, identified at the housing-unit level within a city or collection of cities, as in the cases of Kim et al. and Zabel and Kiel. These estimates are clearly useful in the context of these cities, but they are open to external validity and generalizability critiques that my paper attempts to address by identifying observations across the United States at the county level. The county-level identification of air quality and housing prices follows in the footsteps of Chay and Greenstone, but I intend to use a different instrument which I believe offers more exogenous variability to exploit in my estimation.

Table 1:

Statistic	N	Mean	St. Dev.	Min
April	0.09956	0.12162	0.12112	0.12279
	(0.24814)	(0.36143)	(0.28791)	(0.29743)
May	0.10841	0.08784	0.07764	0.09735
	(0.23574)	(0.20595)	(0.18165)	(0.25913)
June	0.07854	0.12838	0.07764	0.06969
	(0.19606)	(0.28639)	(0.21512)	(0.18667)
July (1-17)	0.03927	0.02027	0.05901	0.03097
	(0.14506)	(0.06918)	(0.21013)	(0.14194)
July (18-31)	0.03927	0.02703	0.07298	0.06858
	(0.14601)	(0.0787)	(0.21766)	(0.23863)
August	0.11836	0.0473	0.06677	0.12721
	(0.28707)	(0.17518)	(0.21965)	(0.30481)
September	0.10177	0.01351	0.09006	0.09845
	(0.25659)	(0.05731)	(0.27607)	(0.2483)
October	0.12113	0.06081	0.09783	0.0885
	(0.26712)	(0.21575)	(0.26097)	(0.23669)
November	0.09624	0.02703	0.11025	0.10177
	(0.2404)	(0.0787)	(0.28892)	(0.21766)
December	0.10177	0.02703	0.11646	0.10619
	(0.26821)	(0.0787)	(0.27815)	(0.2256)