Fire Technology, 47, 255–273, 2011

DOI: 10.1007/s10694-010-0145-1

Enticing Arsonists with Broken Windows and Social Disorder

Douglas S. Thomas and David T. Butry, Office of Applied Economics, Building and Fire Research Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, USA*

Jeffrey P. Prestemon, Forestry Sciences Laboratory, Southern Research Station, USDA Forest Service, PO Box 12254, Research Triangle Park, NC 27709, USA

Received: 18 November 2008/**Accepted:** 2 February 2010

Abstract. In criminology, it is well understood that indicators of urban decay, such as abandoned buildings littered with broken windows, provide criminals with signals identifying neighborhoods with lower crime detection and apprehension rates than better maintained neighborhoods. Whether it is the resident population's sense of apathy, lack of civic pride, or fear of confrontation that causes criminals to perceive an easy mark, it nevertheless emboldens them to strike. Previous research of wildland arson hints that broken windows (e.g., areas of criminal activity) are partly responsible for arson outbreaks within the wildland–urban interface. We model the incidence of wildland and non-wildland arson ignitions in Michigan from 2001 to 2005 as a function of constructed Broken Windows indices. Our results suggest that crime prevention and urban revitalization programs may be as valuable as fire suppression, fuels management, and law enforcement in limiting incidence and the damage from both wildland and non-wildland arson.

Keywords: Broken Windows, Arson, Social disorder, Wildland–urban interface

1. Introduction

Along with all crimes, arson wildfires have declined considerably over the past several years in the United States. Data from USDA Forest Service and state agencies indicate reductions in both the total number of such fires and the number of fires per person. These declines are similar for non-wildland arson fires, as well, which fell by more than two-thirds between 1980 and 2006 [1]. In spite of the reduced trends in numbers of fires and fires per person, arson wildfire area per person has not declined, at least in national forests of the United States, consistent with a trend of larger fires on average [2]. From 1971 to 2005, for example, in U.S. national forests, the 3-year moving average arson fire size quintupled from 11 ha to 57 ha while the 3-year moving average area burned by arson fires tripled, from 15,000 ha to 46,000 ha. As a result, arson wildfires continue to exact large costs on society. Their proximity to where people live and work means that, in

* Correspondence should be addressed to: David T. Butry, E-mail: david.butry@nist.gov

spite of the trend in numbers of fires, values at risk are likely to have increased substantially over the past few decades. A single arson wildfire can cause losses of over $100 million [3]. Data from the National Fire Protection Association show that arson fires in the United States cost the public an estimated $1.1 billion annually in the period 2000–2005 [4]. In addition, hundreds of deaths are associated with the approximate 315,000 intentional fires set annually in the nation [5]. Wildland managers and law enforcement therefore continue to seek ways to further reduce arson occurrence.

Evidence suggests, however, reducing arson through law enforcement remains a challenge; arson (wildland and non-wildland) and other major property crimes are cleared at a low rate. Arson was cleared by arrest in a total of 22.5% of cases for reporting agencies in the United States in 2006 [6]. From 1995 to 2006, an average of 22% of structure, 7.6% of vehicle based, and 19.4% of other arson offenses were cleared by arrest.[1] In 2006, non-negligent murders were cleared by arrest in 61% of cases, forcible sexual offenses in 41% of reported cases, aggravated assault in 54%, and robbery in 25% [7]. Although similar to the clearance rate for other property crimes, the low rate for arson highlights both the investigative challenges of arson and the limited law enforcement resources available.

Achieving further reductions, however, may be possible with new research and technologies focused on building software that can predict crime occurrences, and initial evidence indicates some successes [8]. Arson wildfires may be particularly amenable to such new approaches, because arson tends to cluster in space and time [9]. The challenge is specifying statistical models that can reliably predict arson occurrence. Although recent statistical models use weather and economic conditions to help explain arson wildfires at fine spatial and fine and large temporal scales, we found no research that has identified correlations of arson with other kinds of crimes, an approach that could be useful for prediction. Evidence suggests, however, that crime prediction may be aided by using information on recent less serious crimes than arson or other major felonies. The so-called Broken Windows theory of crime, a motivator of some of these analyses, we contend, can serve as a framework for hypothesis testing and model specification.

The objective of this research is to specify and report statistical models that test the Broken Windows theory for arson on both wildland and non-wildland fires (i.e., to test whether measures of physical and social disorder are correlated with arson). To assist in prediction, we use information on petty crimes. These crimes, like arson, are typically committed by young people and may index the kinds of deteriorating and deteriorated social conditions that can lead to intentional firesetting. The primary contribution of our research is that it is the first known test of the Broken Windows theory for arson. A second contribution is that we are the first to statistically evaluate how wildland and non-wildland arson may be related and what drives their relationship. Finally, we demonstrate how to conduct a rigorous test of predictive power of such models based on hold-out samples.

[1] As reported in Table 27 (*Percent of Offenses Cleared by Arrest or Exceptional Means*) of the Crime in the United States, U.S. Federal Bureau of Investigations published annually from 1996 to 2006.

In the following pages, we describe a theoretical structure for understanding both wildland and non-wildland arson fires. This theoretical structure, which includes the Broken Windows theory, leads to an empirical model that explains the observed spatial and temporal variation in arson incidence. Section 2 presents the Broken Windows theory, how it relates to arson, the factors that might contribute to it, and its relation to arson behavior. Section 3 discusses our zero-inflated Poisson (ZIP) model used to test the Broken Windows theory. Section 4 includes a presentation of the results, and Sects. 5 and 6 provide a discussion and offers some final concluding thoughts.

2. Theory

Because arson crimes are often difficult to solve, evidence is fragile, and witnesses are rare, the ability to prevent arson by identifying those areas at highest risk and mitigating those risk factors is vital for arson prevention [9]. Scientific practice compels the analyst to begin the task of model specification by appealing to theories. In arson, this means appealing to both crime and fire modeling theoretical constructs. In Becker's economic model of crime, he suggests that the "cost" of committing a crime influences the likelihood a would-be criminal engages in illegal behavior [10]. The implication is that, as arrest and conviction rates fall, so do the expected costs of committing a crime (i.e., there exists a lower likelihood of arrest, conviction, and jail time or a fine). Prestemon and Butry [11] found that wildland arson occurrence in Florida is consistent with Becker's theory. The Broken Windows theory posits that dilapidated buildings and infrastructure along with disorderly behavior influence criminal behavior [12]. Broken windows become signals to others that there is little or no concern about destructive behavior directed toward the building and there are no consequences for breaking another window [13] (i.e., a lower likelihood of arrest). It suggests that indicators of urban decay and social disorder encourage crime, or that observed disorder fosters greater perceived disorder [14]. Comparing illegal firesetting in the urban environment with that found within the wildland–urban interface allows us to identify those factors that limit arson in both settings, while identifying those conditions which distinguish success in the wildland versus the urban (e.g., climate and weather). We test whether a wildland analog to the Broken Window theory exists; for instance, are rural measures of disorder (e.g., illegal dumping) correlated with increased arson activity? Based on Becker's economic model of crime, the expected costs of crime are perceived to be lower in broken window areas. If this is correct, then crime prevention and urban revitalization programs may be just as valuable as fire suppression, fuels management, and law enforcement in limiting incidence and the damage from arson.

3. Empirical Model

The analysis estimates a two-equation model; one equation modeling the count of non-wildland arson ignitions (e.g., structure and vehicle) and a second modeling the count of wildland arson ignitions. In this analysis, we explore the number of arson counts by county by year over the years 2001 to 2005—415 observations in total. We use a ZIP model [15] to model the counts of arson ignitions occurring in each county/year combination while allowing for a "hurdle" process to exist—i.e., the count of arson is either (1) always zero or (2) zero or sometimes a positive integer. In state 1 (always zero), arson does not occur there (for example, we will not observe wildland arson ignitions in highly urbanized counties) or is not reported; in state 2 arson may or may not occur there. The ZIP modeling structure acknowledges that a zero count of arson could occur from either one of these states and allows for a different set of covariates to explain states 1 and 2. Statistically it is important to differentiate between the two.

The primary equations for the model are:

$$\Pr(A_{ki}=0)=\Pr(s_{ki}=0)+(1-\Pr(s_{ki}=0))\mathrm{e}^{-\lambda_{ki}} \tag{1a}$$

$$\Pr(A_{ki}=a_{ki})=\frac{(1-\Pr(s_{ki}=0))e^{-\lambda_{ki}}\lambda_{ki}^{a_{ki}}}{a_{ki}!},\quad a_{ki}=0,1,2,\ldots \tag{1b}$$

where A is the count of arson; k indexes the type of arson (i.e., non-wildland, wildland); i indexes the observations (again, we use county-year combinations); s is an indicator (binary) variable identifying the state (=0 if arson never occurs [state 1]; =1 otherwise); we assume the probability of state 1 can be estimated as a function of covariates $\mathbf{z}$ ("inflation factors") and parameters $\boldsymbol{\gamma}$, such that $\Pr(s_{ki}=0)=F(\mathbf{z}_i,\boldsymbol{\gamma}_k)$; the expected number of arson events per period, given state 2, is: $\lambda_{ki}=e^{\boldsymbol{\beta}_k'\mathbf{x}_i}$, where the number of arson events are a function of covariates $\mathbf{x}$ ("arson count factors") and parameters $\boldsymbol{\beta}$. The expected number of arson fires per period is:

$$E[a_i|\mathbf{x}_i]=(1-F(\mathbf{z}_i,\boldsymbol{\gamma}_k))\mathrm{e}^{\boldsymbol{\beta}_k'\mathbf{x}_i} \tag{2}$$

The hypotheses tested are (1) whether Broken Windows variables influence the number of arson ignitions and (2) whether the effects of the explanatory variables differ between non-wildland and wildland arson (e.g., do police have the same deterrent effect on non-wildland arsonists as they do on wildland arsonists?). Specifically, we are interested in the elasticities of the model (e.g., how a 1% change in a Broken Windows variables affects the number of arson fires in percentage terms). The elasticities for the ZIP model are:

$$\frac{\partial E[a_{ki}|\mathbf{x}_i]}{\partial \mathbf{x}_i}\frac{\mathbf{x}_i}{a_{ki}}=(1-F(\mathbf{z}_i,\boldsymbol{\gamma}_k))\mathrm{e}^{\boldsymbol{\beta}_k'\mathbf{x}_i}\boldsymbol{\beta}_k'\mathbf{x}_i a_{ki}^{-1} \tag{3}$$

In this analysis the probability of arson state 1 ($\Pr(s_{ki}=0)$) is estimated using the logit specification, so that $F(\mathbf{z}_i,\boldsymbol{\gamma}_k)=\left(1+e^{\boldsymbol{\gamma}_k'\mathbf{z}_i}\right)^{-1}$. Using Stata (Stata Corporation), we maximize the following log-likelihood functions (one for each k) to estimate the ZIP model's parameters,

$$\ln L_k = \sum_{i=1}^{N} (1 - s_{ki}) \left[\ln\left\{ \left(1 + e^{\gamma_k' \mathbf{z}_i}\right)^{-1} \right\} + \ln\left\{ 1 - \left(1 + e^{\gamma_k' \mathbf{z}_i}\right)^{-1} \right\} - e^{\beta_k' \mathbf{x}_i} \right] + s_{ki} \left[\ln\left\{ 1 - \left(1 + e^{\gamma_k' \mathbf{z}_i}\right)^{-1} \right\} - e^{\beta_k' \mathbf{x}_i} + a_{ki} \beta_k' \mathbf{x}_i - \ln(a_{ki}!) \right] \quad (4)$$

3.1. Study Site and Data

Arson can be defined a number of ways. For the purpose of this paper, we defined our arson measure as an intentional fire that "includes deliberate misuse of a heat source or a fire of an incendiary nature" [1]. This definition is used by the National Fire Incident Reporting System (NFIRS), the source for arson data in this paper, to identify the cause of ignition for a fire [1]. We define wildland arson fires as any intentional fire set in natural vegetation; the remaining intentional fires are non-wildland fires. The primary arson data for this analysis was taken from the NFIRS 5.0 for 2000 through 2005 in the state of Michigan and was aggregated by county. (The analysis runs from 2001 through 2005 due to the inclusion of 1-year lags of multiple model regressors and because 2000 was the first year Michigan used NFIRS Version 5.0.) The NFIRS data for 2002 through 2005 was augmented with data from the Michigan State Fire Marshal due to under reporting in Detroit.

The population of Michigan is concentrated toward the southern section of the state where the major urban centers are located; therefore, one would expect more arson fires in those areas (Figure 1) and that they would tend to follow major

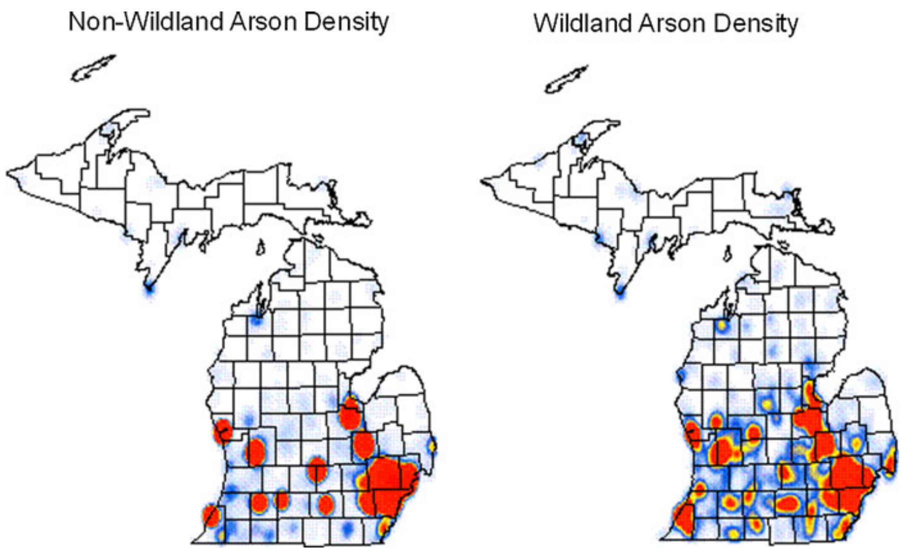

Figure 1. Arson density map of Michigan using NFIRS data (2000–2005). ***Shaded areas*** **denote high levels of arson.**

Table 1
Arson in Michigan Recorded by Augmented NFIRS and UCR

	NFIRS data			UCR data	
Year	Wildland	Non-wildland	Total	Arrests	Reported
2000	907	1815	2722	597	4540
2001	671	2003	2674	600	4522
2002	659	5089*	5748	537	4263
2003	1158	5193*	6351	441	4909
2004	553	4569*	5122	443	4112
2005	930	4373*	5303	350	3729
Average	813	3840	4653	495	4346

* This data has been augmented with data from the Michigan State Fire Marshal

Table 2
Average Proportion of Months of Fire Department Reporting to NFIRS by County

Year	Mean	Median	Minimum	Maximum
2000	0.72	0.75	0.25	0.95
2001	0.74	0.79	0.29	1.00
2002	0.77	0.83	0.27	1.00
2003	0.79	0.83	0.29	0.99
2004	0.70	0.75	0.19	1.00
2005	0.68	0.70	0.18	0.89
Entire period	0.73	0.77	0.18	1.00

thoroughfares through the state. The northern part of the state has a high concentration of U.S. national forests, which might be attractive targets to wildland arsonists, but there are far fewer people living and fewer roads in those areas. Notice that a number of the counties with high rates of non-wildland arson occur near urban areas; in contrast, high rates of wildland arson are more spread out. It appears that the counties surrounding high population areas might experience a higher rate of wildland arson, suggesting that they bear a disproportionate amount of the cost of wildland arson for their population size.

Based on augmented NFIRS, Michigan averaged 4653 reported arson fires, including 3840 reported non-wildland arson fires and another 813 reported wildland arson fires between 2000 and 2005 (Table 1). NFIRS is a voluntary system, making underreporting an issue. Michigan state law requires reporting, however, some fire departments are behind in their reporting. Therefore, we investigated the level of underreporting in the reported NFIRS data. Most fire departments consistently reported fire incidents over the study period; however, some department reporting appeared spotty. Because it was not known whether these gaps were reporting gaps or accurately reflect periods of no fire activity, we developed a metric to measure fire department reporting by county (Table 2).

In our manual examination, we typically found fire departments to report multiple incidents per month if any were reported at all. Meaning it was rare to find a department reporting only a single incident within a month. Gaps tended to begin 1 month and end another. We computed the proportion of months per year that a fire department reported fire incidents. Our reporting metric was to average the individual fire department statistics over each county/year. Over the study period, the median proportion of months of fire department reporting, averaged over each county, was 0.77. Thus, the average fire department reported to NFIRS between 9 months and 10 months out of the year. We interpret this finding as evidence that, for three fire departments that reported year-around, a fourth did not.

Underreporting is also a problem for standard econometric models. Fortunately, the ZIP specification is perfectly designed to handle such an issue. As described in the previous section, the ZIP model allows for the possibility of zero inflation, which in our case may be caused partly from the underreporting of arson incidents. By introducing the reporting metric into the first stage of the ZIP model estimation, it limits any biases or inconsistencies that could result in the second stage.

3.2. Arson Count Factors (x)

We model the count of arson ignitions as a function of the "costs" of crime commission (i.e., the likelihood of being apprehended), opportunity costs (i.e., other non-criminal economic opportunities), and measures of urban/rural decay (disorder). Year dummy variables are included to account for slow moving, state-wide conditions related to arson starts. For instance, criminal penalty levels are believed to deter criminal activity by raising the cost of commissioning a crime; the effect of criminal sanctions, applied to all counties, are therefore captured by the year dummy variables.

3.2.1. Crime Commission Costs.

Police: Total number of officers by year. Source: Inter-University Consortium for Political and Social Research [16].
Arson: Arrests and number of offenses reported for arson (all types). Source: Uniform Crime Reports [17].
Precipitation: Annual accumulation of precipitation. Source: National Oceanic and Atmospheric Administration [18].
Temperature: Mean annual high temperature. Source: National Oceanic and Atmospheric Administration [18].

3.2.2. Opportunity Costs.

Unemployment: Annual unemployment rate. Source: Bureau of Labor Statistics [19].
Population: Annual population. Source: Census Bureau [20].
Youth Population: Annual number of youth aged 15–21. (Studies have shown that youth participate in arson behavior more frequently than adults; since 1992

they have accounted for half of all arson arrests [4].) Source: Census Bureau and used in the model [21].

3.2.3. Urban (Rural) Decay.

Disorder: Arrests for prostitution, vandalism, vagrancy, curfew violation, public drunkenness, drug possession and sale, and runaways. (Reported street crime data does not exist.) Source: Uniform Crime Reports [17].

Vacancy Rate: Number of units in the county that the post office has deemed vacant; this number is then divided by the total number of units in the county. Source: United States Post Office [22] and Census Bureau [23].

Illegal Dump Sites: Number of illegal dump sites on park lands. Source: Michigan Coalition for Clean Forests [24].

The set of regressors include (**x**): disorder (DIS), disorder normalized by numbers of police (DIS_POL), vacancy rate (VAC), vacancy rate normalized by population (VAC_POP), vacancy rate normalized by youth population (VAC_YPOP), found illegal dump sites normalized by federal land area (DUMP_FED), total number of officers by year normalized by population (POL_POP), arson arrest rate (ARSON_AR) (arrests normalized by reports), annual deviation in unemployment over the study period (DU), mean annual high temperature (TMAX), annual accumulation of precipitation (PRECIP), a trend variable (TREND), and year dummy variables (Y2002, Y2003, Y2004, Y2005). Disorder and its interaction, vacancy rate and its interactions, police, and arson arrest rate are all lagged 1 year to avoid possible simultaneity bias with the dependent variable. Disorder is normalized by police to account (proxy) for police success. The expectation is that disorder will be positively related to arson, holding police success constant. Police success is expected to be negatively related to arson rates. Vacancy rate is normalized by population to account for the net of two possible effects: a surveillance effect (arsonist may not target areas with "lots of eyes") or an ignition source effect (larger populations include more potential arsonists). Vacancy rate is also normalized by youth population to evaluate whether vacant homes are targets of juveniles or older arsonists with different motivations (e.g., arson for profit). Summary statistics are presented in Table 3.

3.3. Inflation Factors (z)

To account for the possibility of zero inflation (i.e., if we observe more county/year combinations of zero arson events than would be expected from a standard Poisson process), we model the probability of no arson occurring as a function of the NFIRS reporting metric (described previously), the square miles of six different land types (representing unoccupied areas, areas with little surveillance, or areas with a low (no) likelihood of ignition success). These variables are to be thought of as influencing cost of crime commission. For instance, some fuel types are easier to burn than others, thus reducing the effort (i.e., cost) of arson.

Table 3
Summary of Data Set

	Mean	Minimum	Maximum
DUMP_FED	0.0004	0.0000	0.0635
ARSON_AR	0.1728	0.0000	4.0000
VAC	0.0298	0.0000	0.1033
DIS	551.8434	1.0000	16,190.0000
DIS_POL	2.4506	0.1667	10.8333
TMAX	56.9996	48.5565	63.2616
PRECIP	32.7439	18.0900	46.3200
POL_POP	0.0018	0.0008	0.0105
DU	0.4432	−1.4833	2.4167
VAC_YPOP	0.0000	0.0000	0.0001
VAC_POP	0.0000	0.0000	0.0000
TREND	3.0000	1.0000	5.0000
LAND_GH	0.0512	0.0053	0.1600
LAND_SL	0.0176	0.0004	0.4301
LAND_MF	0.0361	0.0007	0.1943
LAND_DF	0.2371	0.0544	0.5132
LAND_OT	0.5979	−0.2105	0.9296
REPORT	0.7347	0.1833	1.0000
CRIME	636.1735	0.0000	21,075.0000
CRIME_POL	2.2986	0.0000	8.1667

3.3.1. NFIRS Reporting.

Reporting: Proportion of fire department months, over a year, with active NFIRS reporting, average by county. Source: NFIRS 5.0.

3.3.2. Crime Activity.

Crime: Arrests for index crimes (homicide and manslaughter, robbery, rape, aggravated assault, burglary, larceny, motor vehicle theft, and arson). Source: Uniform Crime Reports [17].

3.3.3. Fuel Type.

Deciduous forest: The proportion of land within the specified county that is "dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover; additionally more than 75% of the tree species shed foliage simultaneously in response to seasonal change" [25]. Source: Multi-Resolution Land Characteristics Consortium 2001 National Land Cover Database.
Mixed forest: The proportion of land within the specified county that is "dominated by trees generally greater than 5 m tall, and greater than 20% of

total vegetation cover; additionally, neither deciduous nor evergreen species are greater than 75% of total tree cover" [25]. Source: Multi-Resolution Land Characteristics Consortium 2001 National Land Cover Database.

Shrub- and Scrub-Land: The proportion of land within the specified county that is "dominated by shrubs; less than 5 m tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage, or trees stunted from environmental conditions" [25]. Source: Multi-Resolution Land Characteristics Consortium 2001 National Land Cover Database.

Grassland and Herbaceous: The proportion of land within the specified county that is "dominated by grammanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing" [25]. Source: Multi-Resolution Land Characteristics Consortium 2001 National Land Cover Database.

Others: The remaining proportion of land within the specified county that is not grassland and herbaceous, shrub- and scrub-land, mixed forest, or deciduous forest. Source: Multi-Resolution Land Characteristics Consortium 2001 National Land Cover Database [25].

The final set of inflation factors include: reporting metric (REPORT), index crime arrests (CRIME), index crime arrests normalized by police (CRIME_POL), deciduous forest (LAND_DF), mixed forest (LAND_MF), shrub- and scrub-land (LAND_SL), grassland and herbaceous (LAND_GH), and other land types (LAND_OT). CRIME and CRIME_POL are lagged 1 year to avoid possible simultaneity bias with the dependent variable. CRIME is normalized by police to account (proxy) for police success. The inflation procedure estimates the probability of zero arson counts. It explains why arson may not be observed. We expect that the reporting metric will be negatively correlated with the probability of zero arson. We also expect arrests of index crimes to be negatively related to the probability of zero arson, holding police success (CRIME_POL) constant, and for police success to be positively related to the probability of zero arson. We have no prior expectation regarding the influence of the individual land types on reported arson.

4. Results

Both models are highly significant (see Table 4). Of the 415 observations used in estimation of the non-wildland model, 368 (89%) recorded non-zero arson counts. Of the 415 observations used in estimation of the wildland model, 358 (86%) recorded non-zero arson counts. In each of the models, a Vuong test [26] is utilized to test whether the ZIP specification is preferred over the Poisson. For the non-wildland model, the Vuong test statistic is $z = 1.40$ with $p = 0.08$; for the wildland model, the Vuong test statistic is $z = 3.31$ with $p < 0.01$. Thus, the Zero Inflated Poisson (ZIP) specification is preferred for both models (only weakly preferred for the non-wildland model).

Table 4
Results of Zero Inflated Poisson Regression

Variable	Wildland arson Coefficient	Wildland arson Standard error	Non-wildland arson Coefficient	Non-wildland arson Standard error
DUMP_FED	−1.98E+01**	7.80E+00	N/A	N/A
ARSON_AR	−4.98E−02	6.39E−02	−2.36E−01***	6.00E−02
VAC	1.86E+01***	8.66E−01	3.27E+01***	4.59E−01
DIS	−1.22E−05*	7.25E−06	9.57E−05***	3.51E−06
DIS_POL	1.36E−02	1.83E−02	−1.71E−01***	1.20E−02
TMAX	1.19E−01***	8.51E−03	6.80E−02***	4.94E−03
PRECIP	−3.84E−02***	4.41E−03	−3.85E−02***	2.59E−03
POL_POP	−5.25E+01***	1.54E+01	−6.40E+01***	7.68E+00
DU	7.83E−03	4.28E−02	4.78E−02**	2.49E−02
VAC_YPOP	9.15E+04***	2.79E+04	3.41E+05***	3.14E+04
VAC_POP	−1.79E+06***	2.31E+05	−5.55E+06***	2.57E+05
TREND	6.85E−02***	2.15E−02	−2.67E−02**	1.32E−02
Y2002	−2.45E−01***	6.05E−02	−2.70E−01***	3.15E−02
Y2003	5.10E−01***	6.47E−02	2.51E−02	3.36E−02
Y2004	−8.92E−02	6.31E−02	3.19E−01***	2.99E−02
Constant	−3.65E+00***	4.97E−01	1.09E+00***	2.91E−01
Zero inflation factors				
LAND_GH	−7.24E+00	1.10E+01	3.96E+00	1.27E+01
LAND_SL	−1.78E+01*	9.63E+00	−8.16E+00	9.27E+00
LAND_MF	−2.82E+01**	1.39E+01	−2.10E+01	1.72E+01
LAND_DF	−5.69E+00*	3.32E+00	−4.18E+00	3.78E+00
LAND_OT	−9.75E+00**	4.05E+00	−6.28E+00	4.62E+00
REPORT	−1.57E+00	1.45E+00	−2.61E+00	1.66E+00
CRIME	−1.06E−02**	4.96E−03	−1.14E−02**	5.50E−03
CRIME_POL	3.18E−01	2.53E−01	4.53E−01*	2.75E−01
Constant	7.87E+00**	3.85E+00	4.42E+00	4.59E+00

* $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

4.1. Non-Wildland Arson

Disorder (DIS), vacancy rate (VAC), the vacancy rate normalized by youth population (VAC_YPOP), mean annual maximum temperature (TMAX), unemployment (DU), the constant (CONSTANT), and the 2004 year dummy (Y2004) were positively related to the count of non-wildland arson and statistically significant (10%). Disorder normalized by police (DIS_POL), vacancy rate normalized by population (VAC_POP), police normalized by population (POL_POP), precipitation (PRECIP), the trend variable (TREND), arrests normalized by reports (ARSON_AR), and the 2002 year dummy (Y2002) were negatively related to the count of non-wildland arson and statistically significant (10%). For the year dummies, 2003 was insignificant; note that the 2005 year effect is included in the constant.

The only regressors found statistically (10%) related to the probability of zero count of arson were index crime arrests (CRIME) and index crime arrests normalized by police (CRIME_POL). Arrests were negatively related to the probability of zero, whereas arrests per police were positively related. None of the land type variables were significant. It is notable that the reporting metric was only weakly related (12%) to the probability that zero arsons were reported the county in the year, but its effect (negative) was as expected.

4.2. Wildland Arson

Vacancy rates (VAC and VAC_YPOP), mean annual maximum temperature (TMAX), the trend variable (TREND), and the 2003 year dummy were positively related to the count of non-wildland arson and statistically significant (10%). Found illegal dump sites normalized by area of federal lands (DUMP_FED), disorder (DIS), vacancy rate normalized by population (VAC_POP), police normalized by population (POL_POP), precipitation (PRECIP), the constant (CONSTANT), and the 2002 year dummy (Y2002) were negatively related to the count of non-wildland arson and statistically significant (10%). Arson arrest rate, disorder normalized by police (DIS_POL), unemployment, and the 2004 year dummy were not found to be statistically related to the count of reported arson. Several land type control variables (LAND_DF, LAND_MF, LAND_SL, and LAND_OT) were found to be negatively related to the probability of no wildland arson. Finally, as in the case for non-wildland arson, crime arrests (CRIME) were found to be negatively related to the probability of no wildland arson counts.

4.3. Predictions Out of Sample

Using a jackknife procedure, we assessed the predictive ability of the models using a hold-one-back design. The jackknife procedure estimates each model (non-wildland and wildland), minus one observation, which is held out to assess the accuracy of the model (based on $n - 1$ observations) to estimate the out-of-sample (i.e., held-out) observation's arson count. For each model this procedure is replicated 415 times (n). The accuracy of the ZIP models are compared to two "naïve" models. The first naïve model estimates the out-of-sample observation's arson count as that from the prior year ("temporal autoregressive"). This allows us to compare our models' ability to forecast purely in time. The second naïve model estimates the out-of-sample observation's arson count as the mean arson count of all other counties in that year ("spatial autoregressive"). This allows us to compare our models' ability to forecast in space.

The performance of the ZIP statistical prediction compared with the temporal and spatial autoregressive models is shown in Table 5. Using the mean absolute error as the metric of performance, we find that the temporal autoregressive model outperforms the ZIP model in the prediction of arson (both wildland and non-wildland). The ZIP models outperform the spatial autoregressive model in the prediction of non-wildland and wildland arson.

Because arson rates move slowly from year to year, it is not surprising that the temporal autoregressive model predicts well. When forecasting next year's arson

Table 5
Performance of the Out of Sample Prediction

Model	Wildland arson		Non-wildland arson	
	Mean absolute error	Standard deviation	Mean absolute error	Standard deviation
Zero-inflated Poisson	5.9	12.0	24.5	66.6
Naïve 1: temporal autoregressive	5.5	8.8	10.8	33.2
Naïve 2: spatial autoregressive	9.3	13.8	78.8	237.4

for a county, assuming it will be similar to the current year's arson is a reasonable approximation.[2] For all of Michigan, our statistical models cannot improve upon this type of prediction for wildland arson. Our wildland arson model is nearly as accurate, however. The mean absolute error of the temporal autoregressive model is only 7% smaller than that produced by the ZIP model.

Both ZIP models greatly improve upon the prediction provided by the spatial autoregressive mode. This demonstrates that cross-sectional (i.e., county) variation is rather large, and is better explained by the drivers of arson employed in our models. County differences related to crime commission costs, opportunity costs, and urban/rural decay can reasonably explain differences in arson occurrence. In fact, removing the Broken Window variables from the ZIP models erodes the model's predictive performance (increasing the mean absolute error by 174.7%) for the non-wildland model and (by 15.3%) for the wildland model (analysis not shown).

5. Discussion

We are interested in how Broken Windows affect arson ignition rates. Specifically, we seek to better understand how measures of disorder, vacancy, and illegal dumping relate to the number of reported non-wildland and wildland arson fires.

Disorder (DIS), or the previous year's number of arrests for prostitution, vandalism, vagrancy, curfew violation, public drunkenness, drug possession and sale, and runaways, are positively correlated with the number non-wildland arson ignitions, holding police effort constant (DIS_POL), which is as expected. By holding police effort constant, an increase in arrests implies an increase in occurrence. The elasticity of disorder on non-wildland arson ignitions, 0.05 (see Table 6), means that a 1% increase in the incidence of disorder corresponds to an increase in reported non-wildland arson fires by 0.05% (holding all other variables at their means). However, if over the study period the number of arrests for disorder had been 20% higher, we estimate that the number of non-wildland fires would have been 17% higher (4299 more reported arson fires for the state during the 5 year period) (see Table 7). A crackdown on disorder through a 10% increase in police

[2]This finding argues for developing an autoregressive version of a zero-inflated Poisson model. Although autoregressive Poisson models have been developed [27], the corresponding ZIP model has not.

Table 6
Elasticities from Zero Inflated Poisson Regression

Variable	Wildland arson ey/ex	Non-wildland arson ey/ex
DUMP_FED	−7.90E−03	N/A
ARSON_AR*	−8.62E−03	−4.09E−02
VAC*	5.53E−01	9.68E−01
DIS*	−6.74E−03	4.93E−02
DIS_POL*	3.33E−02	−4.20E−01
TMAX*	6.76E+00	3.87E+00
PRECIP	−1.26E+00	−1.26E+00
POL_POP	−9.38E−02	−1.14E−01
DU	3.47E−03	2.14E−02
VAC_YPOP*	5.64E−01	2.11E+00
VAC_POP*	−1.40E+00	−4.36E+00
TREND*	2.05E−01	−8.02E−02
Y2002	−4.91E−02	−5.42E−02
Y2003*	1.02E−01	5.04E−03
Y2004*	−1.77E−02	6.34E−02
Zero inflation factors		
LAND_GH	1.02E−04	−3.75E−05
LAND_SL	8.64E−05	2.66E−05
LAND_MF	2.81E−04	1.40E−04
LAND_DF	3.72E−04	1.83E−04
LAND_OT	1.61E−03	6.92E−04
REPORT	3.17E−04	3.54E−04
CRIME	1.86E−03	1.24E−03
CRIME_POL	−2.01E−04	−1.92E−04

* Denotes statistically different elasticities for wildland arson and non-wildland arson

effort (DIS_POL) would have reduced the number of non-wildland fires by 3%, or 864 fewer reported non-wildland arson fires for the state during the 5 year study period. The correlations between both disorder and police effort and wildland arson are much weaker. In fact, disorder is negatively related to wildland arson, albeit at the 9% significance level. This unexpected relationship might be accounting for the fact that crimes related to prostitution, vandalism, vagrancy, curfew violation, public drunkenness, drug possession and sale, and runaways occur more often in urban areas, which are less likely to experience wildland arson.

Vacancy rate (VAC) is positively correlated with both non-wildland and wildland arson ignitions (at the 1% significance level) (holding the vacancy rate to population and vacancy rate to youth population ratio constant). With all variables held at their mean, a 1% increase in the vacancy rate is associated with a 0.97% increase in non-wildland arson fires, and a 0.55% increase in wildland arson fires (again, holding the vacancy rate to population and vacancy rate to youth population ratio constant) (see Table 6). While the vacancy rate affects both types of arson, statistically it has a larger effect on non-wildland arson. As seen in

Table 7
The Number and Percent Change in Arson Incidents for the State During the 5 Year Study Period with a Hypothetical ±20% and ±10% Change in Arson Count Factors (x)

		Wildland				Non-wildland			
		−20%	−10%	10%	20%	−20%	−10%	10%	20%
DUMP_FED	Percent	0.07%	0.04%	−0.03%	−0.07%	n/a	n/a	n/a	n/a
	Incidents	3	1	−1	−3	n/a	n/a	n/a	n/a
ARSON_AR	Percent	0.15%	0.07%	−0.07%	−0.15%	0.51%	0.25%	−0.25%	−0.50%
	Incidents	6	3	−3	−6	126	63	−63	−125
VAC	Percent	−14.44%	−7.64%	8.59%	18.27%	−38.09%	−21.85%	29.54%	69.67%
	Incidents	−575	−304	342	727	−9498	−5448	7366	17373
DIS	Percent	0.45%	0.22%	−0.22%	−0.44%	−13.51%	−7.16%	8.09%	17.24%
	Incidents	18	9	−9	−17	−3369	−1786	2018	4299
DIS_POL	Percent	−0.67%	−0.34%	0.34%	0.68%	7.35%	3.60%	−3.46%	−6.79%
	Incidents	−27	−13	13	27	1833	898	−864	−1694
TMAX	Percent	−75.22%	−19.05%	101.15%	304.94%	−55.74%	−33.48%	50.35%	126.08%
	Incidents	−2995	−758	4027	12140	−13899	−8348	12555	31442
PRECIP	Percent	27.90%	13.08%	−11.54%	−21.72%	26.83%	12.60%	−11.16%	−21.06%
	Incidents	1111	521	−459	−865	6690	3142	−2784	−5251
POL_POP	Percent	2.00%	0.99%	−0.98%	−1.95%	3.68%	1.82%	−1.78%	−3.53%
	Incidents	80	40	−39	−77	919	454	−445	−880
DUMP_FED	Percent	−0.09%	−0.04%	0.04%	0.09%	−0.43%	−0.22%	0.22%	0.44%
	Incidents	−3	−2	2	3	−107	−54	55	110
VAC_YPOP	Percent	−4.16%	−2.13%	2.24%	4.61%	−6.10%	−3.17%	3.46%	7.27%
	Incidents	−165	−85	89	183	−1521	−791	862	1812
VAC_POP	Percent	13.34%	6.21%	−5.49%	−10.39%	18.14%	8.08%	−6.74%	−12.49%
	Incidents	531	247	−218	−414	4523	2015	−1681	−3115

Table 7, if the vacancy rate had been decreased by 10% for non-wildland arson, over the study period, the number of arson incidents would have been 22% lower, or been reduced by 5448 reported fires. For wildland arson, this would have resulted in an 8% decrease, or 304 fewer reported fires.

We controlled for two potential confounders in statistical identification of effects of broken windows and other factors on arson fires, total population (VAC_POP) and youth population (VAC_YPOP). These variables were included to account for a possible net surveillance or ignition effect by population, and to evaluate whether vacant homes are a target of juveniles or older arsonists. Holding the vacancy rate constant, we find that an increase in population is associated with an increase in the number of arson (both non-wildland and wildland), which is consistent with an ignition effect (more fire starters), whereas an increase in youth population is associated with a decrease in the number of arson, which is consistent with the finding that areas with older individuals tend to have more arson (at least in our sample). A 1% increase in total population is associated with a 4.4% increase in non-wildland arson and a 1.4% increase in arson (holding vacancy rate and youth population constant). Areas with larger populations tend to have more arson ignitions, all else equal, and this effect is bigger in urban areas than in rural areas. The potential surveillance effect of a large population appears to be overwhelmed by their fire starting potential. A 1% increase in the youth population is associated with a 2.1% decrease in non-wildland arson and a 0.6% in wildland arson (holding the vacancy rate constant): areas with high vacancy rates and older populations experience more arson than other areas, supporting an "arson for profit" motive in our sample.

An unexpected finding was that the number of illegal dumps sites discovered in the previous year is negatively related to wildland arson. Although the significance of the coefficient is at the 1% level, the elasticity is fairly small (−0.008), so while there is statistical significance, there is little economic significance. Although we have considered an illegal dump site as a "rural broken window," a place where we would expect more arson activity, a discovered dump site may be indexing an opposite effect or fine-scale variations in some variables that are not captured by the spatio-temporal scale of our models. For example, a discovered site may lead to changes in local policing efforts in the vicinity of the dump site.

It would seem that the role of weather would be a more important influence on wildland arson than non-wildland. As measured by maximum daily temperature, this is what we find, with the elasticity of temperature with respect to wildland arson nearly double that for non-wildland arson (6.8 vs. 3.9). However, in terms of total annual precipitation, the elasticities of precipitation with respect to wildland and non-wildland arson are about the same (−1.3 for each), indicating the important role of weather in many kinds of firesetting. This finding has a potential benefit for both tactical and strategic changes in law enforcement efforts focused on arson: weather can be forecasted. For example, the National Oceanic and Atmospheric Administration provides short- and long range forecasts for precipitation, temperature, and drought conditions on land in the U.S. and ocean temperature and pressure conditions that can be linked to continental weather and climatic conditions. Coupling weather forecast with estimated arson models could aid in the development of crime mapping tools for law enforcement.

The significance of the ZIP model inflation factor, as evident by the significance of the Vuong tests, implies that arson is more likely to occur in some areas than others. Fuel types are relevant in the wildland arson model, but not so in the non-wildland arson model. The reporting metric (REPORT) is not significantly correlated with the probability of arson, meaning the under-reporting found in the NFIRS data did not introduce any systematic bias. Taken together, the two crime variables (CRIME and CRIME_POL) suggest that the number of previous year's index crime are correlated with the probability of both arson types.

6. Conclusions

This analysis demonstrates that proxies of social disorder can be used to identify areas of high arson risk, consistent with the Broken Window theory. We have found that areas marked with physical and social disorder, and/or many vacant buildings are at risk to arson threats. There appears to be a temporal relationship between previous crimes and future arson rates. However, the relationships between our Broken Windows measures and arson rates differ by ignition location (for both wildland and non-wildland fires).

Our models suggest that steps to curb urban decay and social disorder, through the formation of neighborhood crime watch organizations, increased police patrolling, or from community beautification programs, may be effective antidotes to less controllable, drivers of higher arson rates such as weather and economic decline. Sampson and Raudenbush [14] suggest improving the conditions of neighborhoods already beset with disorder may be difficult and have limited effects. If true, this, coupled with our findings, underscores the importance in preventing the deterioration from occurring. What is important then, from the perspective of our estimated models, then, is to identify and regularly monitor the variables that are leading indicators of urban decay.

Finally, our out-of-sample results indicate that simple temporal autoregressive (“naïve”) models do a reasonable job forecasting arson in time, a result of arson being a slow moving temporal process (little yearly variation across time). Such naïve models may be able to forecast arson reasonably well, but they do little to explain trends or, as we have shown, variations across space. Both the non-wildland and the wildland arson models reported in our research are capable of forecasting in space better than a naïve counterpart, the simple spatial autoregressive model. This latter finding suggests our models may be generalizable (i.e., able to forecast) to other regions of the United States.

References

1. U.S. Fire Administration National Fire Data Center (2006) National fire incident reporting system 5.0. FEMA, Washington, DC, pp 4–25

2. USDA Forest Service (2007) National fire management incident database. http://fam.nwcg.gov/fam-web/weatherfirecd/fire_files.htm. Accessed 22 May 2007
3. Kent B, Gebert K, McCaffrey S, Martin W, Calkin D, Schuster I, Martin H, Bender W, Alward G, Kumagai Y, Cohn PJ, Carroll M, Williams D, Ekarius C (2003) Social and economic issues of the Hayman Fire. In Graham RT (tech ed) Hayman Fire case study. Gen. Tech. Rep. RMRS-114 (Revision). U.S. Department of Agriculture, Forest Service, Fort Collins, pp 315–395
4. Hall JR Jr (2007) Intentional fires and arson. National Fire Protection Association, Quincy, p. 1
5. National Fire Protection Association (2008) Intentional fires. http://www.nfpa.org/assets/images//journal/JF08/arson_report_stats_intentional_fires.gif. Accessed 28 Aug 2008
6. U.S. Department of Justice (2007) Crime in the United States, 2006. http://www.fbi.gov/ucr/cius2006/about/variables_affecting_crime.html. Accessed 28 Aug 2008
7. U.S. Department of Justice, Federal Bureau of Investigation (2007) Uniform crime reporting program data (United States): county-level detailed arrest and offense data, 2000–2005. ICPSR. Inter-University Consortium for Political and Social Research [Producer and Distributor], Ann Arbor
8. Gorr W, Harries R (2003) Introduction to crime forecasting. Int J Forecast 19:551–555
9. Icove, DJ (1979) Principles of incendiary crime analysis. Dissertation, University of Tennessee, pp 9–11
10. Becker GS (1968) Crime and punishment: an economic approach. J Polit Econ 76(2):169–217
11. Prestemon JP, Butry DT (2005) Time to burn: modeling wildland arson as a autoregressive crime function. Am J Agric Econ 87(3):756–770
12. Wilson JQ, Kelling GL (1982) The police and neighborhood safety: broken windows. Atlantic 249(3):29–38
13. Lochner L (2007) Individual perceptions of the criminal justice system. Am Econ Rev 97(1):444–460
14. Sampson RJ, Raudenbush SW (2004) Seeing disorder: neighborhood stigma and the social construction of broken windows. Soc Psychol Q 67(4):319–342
15. Greene WH (2008) Econometric analysis, 6th edn. Pearson Education, Inc, Upper Saddle River
16. U.S. Department of Justice, Bureau of Justice Statistics (2003) Census of State and Local Law Enforcement Agencies (CSLLEA), 2000: [United States]. Conducted by U.S. Department of Commerce, Bureau of the Census. 3rd ICPSR ed. Ann Inter-University Consortium for Political and Social Research [Producer and Distributor], Ann Arbor
17. U.S. Department of Justice, Federal Bureau of Investigation (2007) Uniform crime reporting program data (United States): county-level detailed arrest and offense data, 2000–2005. ICPSR. Inter-University Consortium for Political and Social Research [Producer and Distributor], Ann Arbor
18. National Oceanic and Atmospheric Administration (2007) NNDC climate data online. Daily Surface Data. http://cdo.ncdc.noaa.gov/CDO/cdo. Accessed 2007
19. Bureau of Labor Statistics (2007) Local area unemployment statistics. http://www.bls.gov/lau/#data. Accessed 2007
20. Population Division, United States Census Bureau (2006) Table 1: annual estimates of the population for counties of Michigan: April 1, 2000 to July 1, 2006 (CO-EST2006-01-26)

21. Population Division, U.S. Census Bureau (2008) Annual county resident population estimates by age, sex, race, and Hispanic origin. http://www.census.gov/popest/counties/. Accessed 2008
22. Department of Housing and Urban Development and United States Postal Service (2007) HUD aggregated USPS administrative data on address vacancies. http://www.huduser.org/DATASETS/usps.html. Accessed 2007
23. U.S. Census Bureau (2007) Housing estimates. Housing Topics. http://www.census.gov/hhes/www/housing.html. Accessed 2007
24. Michigan Coalition for Clean Forests. http://www.cleanforests.org/
25. Homer C, Huang C, Yang L, Wylie B, Coan M (2004) Development of a 2001 national land-cover database for the United States. Photogramm Eng Remote Sens 70:829–840
26. Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57(2):307–333
27. Brandt PT, Williams JT, Fordham BO, Pollins B (2000) Dynamic modelling for persistent event count time series. Am J Pol Sci 44(4):823–843