Environ Ecol Stat (2009) 16:291–319
DOI 10.1007/s10651-007-0083-3

Fighting fire with fire: estimating the efficacy of wildfire mitigation programs using propensity scores

David T. Butry

Received: 1 March 2006 / Revised: 1 August 2006 / Published online: 1 March 2008

Abstract This paper examines the effect wildfire mitigation has on broad-scale wildfire behavior. Each year, hundreds of million of dollars are spent on fire suppression and fuels management applications, yet little is known, quantitatively, of the returns to these programs in terms of their impact on wildfire extent and intensity. This is especially true when considering that wildfire management influences and reacts to several, often times confounding factors, including socioeconomic characteristics, values at risk, heterogeneous landscapes, and climate. Due to the endogenous nature of suppression effort and fuels management intensity and placement with wildfire behavior, traditional regression models may prove inadequate. Instead, I examine the applicability of propensity score matching (PSM) techniques in modeling wildfire. This research makes several significant contributions including: (1) applying techniques developed in labor economics and in epidemiology to evaluate the effects of natural resource policies on landscapes, rather than on individuals; (2) providing a better understanding of the relationship between wildfire mitigation strategies and their influence on broad-scale wildfire patterns; (3) quantifying the returns to suppression and fuels management on wildfire behavior.

Keywords Endogeneity · Prescribed fire · Propensity score · Treatment effects · Wildfire production functions

D. T. Butry (✉)
Office of Applied Economics, Building and Fire Research Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Mailstop 8603, Gaithersburg, MD 20899-8603, USA
e-mail: david.butry@nist.gov

1 Introduction

Nationwide, hundreds of millions of dollars are spent fighting wildfires and still millions of acres of forestland burn each year. Additionally, hundreds of millions more are spent on fuels management programs (programs that mitigate wildfire risk by physically reducing the amount of available flammable vegetation within the forest). What is not clear is whether society is optimally funding these wildfire management (suppression or fuels management) programs. One reason for this uncertainty is that little information exists on the damages and benefits associated with wildland fire, although Butry (2001) estimates the economic damage of the 1998 Florida wildfires. Another reason is little research exists quantifying the effects of wildfire mitigation programs on fire behavior and damage.

The discipline of economics is well suited to examine issues concerning tradeoffs, optimality, and allocations of resources. Tradeoffs are inherent in wildland management. Tradeoffs exist between how much wildfire is allowed, how much is controlled, and how to determine the best mix of suppression and fuels management to achieve these goals. The idea of an "optimal wildfire size" was introduced by Sparhawk in 1925. He explains that the optimal size is one that yields the minimum 'cost plus loss' (cost of suppression plus loss due to wildfire damage). This is an important finding as it may justify, in some instances, "let burn" or limited-action strategies—when the wildfire damages pale in comparison to the cost of treatment. Rideout and Omi (1990) formalize innovations instituted into the least-cost model by Gorte and Gorte (1979) and Davis (1965), who contend that optimal wildfire size is that which corresponds with the minimum cost plus *net value change* (NVC—damages net of benefits), as a profit maximization type problem. The goal here is to maximize the value of wildfire avoidance given the costs of wildfire management.

However, even before one can determine the optimal allocation of fire suppression and fuels management resources, it is critical to understand the physical relationship between management strategies and wildfire behavior. For example, society must understand the effectiveness of fire management before it can decide how much is needed. Unfortunately, there is a real lack of research modeling wildfire behavior at broad, policy relevant spatial scales. While experimental research has focused on the relationship between fuel loads and fire risk, these are based on small-forested plots. Although fine scale analysis is informative, perhaps it is not appropriate for policy-making. Further, it often ignores spatial and spatio-temporal relationships between wildfire and physical intervention strategies. It also cannot account for many of the complexities of landscapes and larger scale weather and climate phenomena.

There is a vulnerability of fine scale and experimental laboratory analysis to ignoring the influence society and culture, institutions, and environmental phenomena have on ignition patterns and fire propagation. The wildland–urban interface (WUI) is the area where communities abut forested areas. In 2000, the WUI comprised 9.4% of the land in the coterminous U.S. and contained 38.5% of all housing units (Stewart et al. 2005). These numbers are up substantially from 1990 levels, as the decade experienced a 19.2% and 22.3% growth in WUI lands and homes within the WUI (Stewart et al. 2005). Figure 1 presents a stylized depiction of wildfire in the wildland–urban interface and highlights its interconnectedness with communities and the environment.

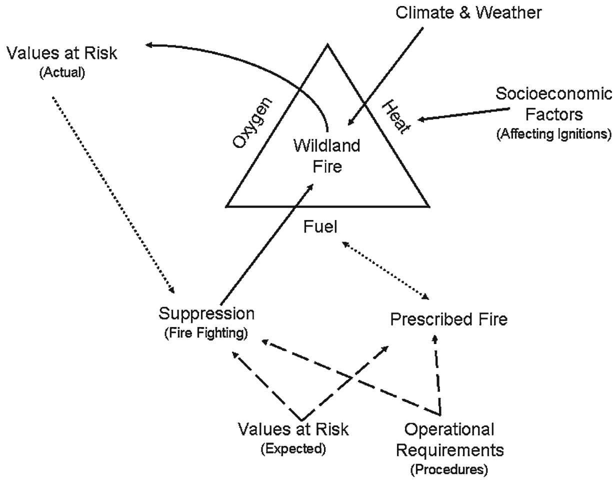

Fig. 1 The fire triangle in the wildland–urban interface

In the center of Fig. 1 is the fire triangle, which is made up of three sides—heat, fuel, and oxygen. Remove anyone of these sides and fire cannot exist. Figure 1 illustrates that wildfire is influenced by several factors, including climate and weather, socioeconomic characteristics (population, housing density, roads, etc.), and wildland management (fire suppression and fuels management). Each one of these factors is related to one or more of the sides of the fire triangle (heat, fuel, oxygen).

Focusing on the wildfire management relationship, it is apparent that fire suppression targets all three sides of the fire triangle. Fire fighters use water to cool the fire and reduce its heat content; fire fighters use fire retardants to smother the flames and deprive fire of oxygen; fire fighters also build fire lines (fire barriers devoid of flammable materials) to separate fire from available fuels. Fuels management, such as prescribed fire fuels treatments, targets only one side of the triangle—fuels—by eliminating dangerous fuel loads before the start of the fire season. Not surprisingly, factors that affect wildfire also affect suppression and fuel management. In addition, wildfire conditions influence suppression and fuels management decisions through a series of feedback mechanisms. While fire fighting mitigates fire size and intensity, the initial fire conditions, including proximity to values at risk (e.g., populations, housing development, infrastructure) influence the size and swiftness of response, especially during times of multiple fires (requiring prioritization of wildfires). Fuels treatments are performed far in the advance of the fire season, but their locations are not distributed randomly across the landscape, nor are their size and intensity randomly selected. Instead, fuels treatments are strategically located to areas of higher wildfire risk, vulnerable to damage from fire.

In economics, it is rare to use experimentally controlled data due to expense and inability to replicate real-world conditions. Instead, the majority of applied economic studies rely on observational data—one observes individuals behaving a certain way. Evaluating the effect of programs using observational data is challenging because often there are underlying factors that influence whether or not an individual enters a program. In the case of wildfire, wildfires are not randomly selected to receive wildfire management. The decision depends on a number of factors, some of which also influence wildfire behavior. For instance, because prescribed fire is not randomly applied across the landscape, it is not unusual to observe areas with higher levels of prescribed fire to have also higher levels of wildfire. Yet, this does not necessarily imply that prescribed fire is ineffective or counterproductive, but it highlights the complex relationship between prescribed fire and wildfire. The true performance of wildfire management can be determined if one compares the same group of wildfires having been managed with their untreated selves. Of course, this latter group, called the counterfactual, is not observable.

Because factors influencing wildfire behavior also influence wildland management decisions, these factors must be accounted for in statistical models. Failing to acknowledge this complex relationship renders estimates from standard statistical models biased as it invalidates some of the required modeling assumptions. Much of the previous empirical wildfire research has focused on modeling either the probability of a wildfire occurrence (ignition) or wildfire size or intensity. The former research primarily examines the factors involved in the probability of fire ignition (fire risk modeling), while the latter examines the factors contributing to a fire's final size or degree of intensity. In general, previous research finds wildfire behavior (however defined—whether meaning frequency, occurrence, size, or intensity) to be related to four sets of factors: wildfire specific characteristics, climate and weather conditions, wildfire management and mitigation (including prescribed fire and suppression effort), and landscape attributes (including dominant landuse and landcover characteristics and socioeconomic characteristics). While the dynamic relationship between wildfire and management has been explored using simulation models (e.g., Reed 1987; Yoder 2004; Amacher et al. 2005), it is rare to find empirical examples that explicitly account for this potential endogeneity.

In Sect. 2, I present program evaluation concepts and econometric techniques, focusing on propensity score matching and blocking. Section 3 presents the data and study site information used in the empirical modeling explained in Sect. 4. Finally, Sects. 5 and 6 present the model findings and research conclusions.

2 The evaluation of programs

In general terms, programs (or their treatments) are designed to produce benefits for the participants, perhaps in the form of higher wages (job training program), higher educational achievement (education program), health enhancement (medicinal drug study), or in the case of wildfire, damage averted from wildfire (wildfire management program). In the program evaluation terminology, a program induces a treatment that affects outcome. For instance, wages, educational attainment, health attribute, and damage averted are all outcomes affected by the program treatment.

In much of the program evaluation literature the effect of a program or treatment is conceptualized as (for example, see Heckman and Hotz 1989; Dehejia and Wahba 1999; Smith and Todd 2005):

$$y_i = \tilde{y}_i + \Delta t_i \tag{1}$$

where y_i is the observed outcome for observation i, $\tilde{y}_i$ is the outcome if i is not selected into treatment (one can also think of $\tilde{y}_i$ as potential outcome before treatment), t_i is an indicator variable coded 1 for treatment and 0 for no treatment, and Δ is the treatment effect. Equation (1) formalizes the simple notion that if an observation undergoes treatment, the outcome value is changed by the treatment effect (Δ), whereas if an observation does not undergo treatment, then the outcome value is equal to the potential outcome ($\tilde{y}$).

Equation (1) is not a structural equation because the outcome is not represented as a function based on economic theory. Essentially the term $\tilde{y}$ embodies all the variables that affect outcome, except for treatment. From a structural model viewpoint, one might envision that outcome is a function of treatment, a set of variables Z that directly affect only outcome, and a set of variables X that directly affect both outcome and treatment (this distinction will prove important later), and e a random error component. The sets X and Z are uncorrelated with each other, hence $\mathrm{Cov}(X, t) \neq 0$ and $\mathrm{Cov}(X, Z) = 0$. The variables represented by X are called confounders because they are correlated with the treatment; i.e., the effects of X and the treatment are not independent of each other, but are confounded with each other. In a structural model of program outcome, these components comprise $\tilde{y}_i$ ($\tilde{y}_i = \pi_z Z_i + \pi_x X_i + e_i$) from (1). Writing (1) as a structural equation yields:

$$y_i = \pi_z Z_i + \pi_x X_i + \Delta t_i + e_i. \tag{2}$$

The reason for distinguishing between (1) and (2) is because some program evaluation econometric techniques assume a structure similar to (1), for instance propensity score matching, while others, such as ordinary least squares, assume a structure similar to (2). I present Eqs. (1) and (2) to demonstrate that they can represent the same concept.

It is the inability of either observing (or measuring) the confounders or specifying the relationship between confounders and outcome that complicates estimation of the true relationship between outcome and treatment. Re-writing (2) to illustrate the situation where X is a set of unobservable confounders yields:

$$y_i = \pi_z Z_i + \Delta t_i + \varepsilon_i \tag{3}$$

where,

$$\varepsilon_i = \pi_x X_i + e_i \tag{4}$$

and $\mathrm{E}[\varepsilon_i | t_i] \neq 0$ and $E[\varepsilon_i] = 0$.

If there are no unobservables or if treatment is uncorrelated with the unobservables, then ordinary least squares estimation of (3) will yield an unbiased treatment effect estimate (Δ). If the confounders X are unobserved and correlated with treatment, least squares estimation of (3) will yield a biased treatment effect estimate because $E[\varepsilon_i | t_i] \neq 0$ (Greene 2000).

2.1 Propensity score matching models

Propensity score matching allows for consistent estimation of the treatment effect between comparison groups that are not experimentally controlled (see Rosenbaum and Rubin 1983; Rosenbaum and Rubin 1985a; Rosenbaum and Rubin 1985b; Dehejia and Wahba 1999; Dehejia and Wahba 2002). Propensity score matching eliminates selection bias by comparing matched participant and non-participant outcomes that have been paired based on their estimated probability of treatment. The difference between propensity score matching and matching on covariates (confounders) is that matching is performed on a single variable (the propensity score), hence the curse of dimensionality (Bellman 1961) is eliminated. The overwhelming bulk of the propensity score matching (PSM) literature involves a binary treatment and is applied in labor economics and epidemiological studies (see Rosenbaum and Rubin 1983; Heckman and Hotz 1989; Dehejia andWahba 1999; Smith and Todd 2005 for a more detailed introduction). Butry (2006) summarizes and compares other program evaluation econometric approaches that may be used to measure the effect wildfire management has on fire behavior, but which involve different assumptions than propensity score matching. Propensity score matching requires a few conditions to be met to produce unbiased treatment effect estimates—strong ignorability (or "selection on observables" per Heckman and Robb (1985) or no unobserved confounders) and that the outcome after a treatment is applied is the same as the potential outcome to treatment (the "Stable–Unit–Treatment–Value" assumption).
Strong ignorability (Rosenbaum and Rubin 1983). If the outcome-treatment relationship is confounded by X, then treatment status is strongly ignorable if the potential outcome (before treatment) is independent of treatment assignment given X:

$$\tilde{y}, \tilde{y} + \Delta \perp t | X \tag{5}$$

where "$\perp$" denotes independence. This means that all confounders X must be known and measurable. Strong ignorability means that conditional on the confounders, outcome before treatment is independent of (future) treatment status. In context of a linear model (such as Eq. (3)), strong ignorability implies $\varepsilon \perp t | X$.

Propensity score matching requires the Stable–Unit–Treatment–Value–Assumption (SUTVA; Rubin 1990), meaning the outcome of one observation when it receives a treatment is the same as its potential outcome to that treatment before it actually receives the treatment. Thus, the unit does not change between the time of assigning and receiving the treatment and the treatment of other units does not change the response of a unit to treatment. This assumption could be violated if the treatment of a neighboring unit affects its response; e.g., in treatment of infectious diseases vaccinating other children can affect the probability of contracting a disease for all children in the community, even those who did not receive the vaccine. Another important requirement that must be satisfied to implement the method is $0 < pr(t = 1|X) < 1$. Every unit in the study must have non-zero probability of receiving either treatment; i.e., every unit must be a candidate for either treatment. The probability of treatment must be strictly greater than zero and strictly less than unity (this condition proves important when matching on the propensity score—it ensures the possibility of matches between the treated and untreated observations). Matching on the confounders will eliminate the selection bias by pairing (matching) treated and non-treated observations based on the set of covariates that influence selection—pairs are created between observations with like covariate values. If X is observable and if $\tilde{y}\perp t|X$ holds, then:

$$\Delta = E_X\left[E[y|t=1,X] - E[\tilde{y}|t=0,X]\right] \tag{6}$$

where E_X is the expectation operator taken with regard to the distribution of X. Because the interest here is on the treatment effect on the treated, only a weaker form of strong ignorability is needed, $\tilde{y}\perp t|X$ rather than $\tilde{y}, \tilde{y}+\Delta\perp t|X$ (Dehejia and Wahba 2002). Note that as the number of covariates increases, it becomes more difficult to match on all the confounders (the curse of dimensionality).

Balancing score (Rosenbaum and Rubin 1983). A balancing score $b(X)$ is defined as one which when conditioned on, results in independence between the treatment assignment and the confounders X, i.e.,

$$X\perp t|b(X). \tag{7}$$

The balancing score eliminates the curse of dimensionality because it is a scalar function of the multivariate X. The balancing score is not simply any function of X; rather, the balancing function b is specified so that, conditioned on the score, treatment status is independent of the confounders X. The notion of balance comes from matching treated and untreated observations based on the balancing score. If a score is a balancing score then the distribution of X in the matched treated observations should be similar to the distribution in the matched untreated observations. They are then said to be balanced.

The propensity score $pr(t=1|X) = p(X)$ is a balancing score, i.e.,

$$X\perp t|p(X). \tag{8}$$

The propensity score is the probability of an observation being treated. In randomized experiments the propensity score is known because the experimental design specifies the probability of assignment to each treatment group for each experimental unit. However, in observational studies it must be estimated. The estimated propensity score is only useful so long as it is a balancing score, meaning that conditional on the estimated propensity score, the distribution of confounders between treatment status groups is the same. Probability models, such as the probit specification, often are used to estimate the propensity score.

Strong ignorability given propensity score (Rosenbaum and Rubin 1983). If treatment status is strongly ignorable given the confounders, then the treatment status is strongly ignorable given the propensity score; that is,

$$\begin{aligned}&\text{If } \tilde{y}, \tilde{y}+\Delta\perp t|X \text{ and if } 0 < pr(t=1|X) < 1 \text{ for all } X,\\ &\text{then } \tilde{y}, \tilde{y}+\Delta\perp t|p(X) \text{ and } 0 < pr(t=1|p(X)) < 1 \text{ for all } p(X).\end{aligned}$$

Propensity score matching assumes that an observation's assignment to the program can be expressed as a function of observed variables X, where $0 < pr(t=1|X) < 1$

(the overlap assumption, per Imbens 2004). Note that the overlap assumption does not include the endpoints (0, 1). In a randomized experiment, the researcher chooses the probability of treatment then randomly selects observations, based on this probability of selection, into treatment. If the researcher were to choose 0 or 1 then they would have to either admit no observations into treatment or they would have to admit all observations into treatment, respectively. Choosing either 0 or 1 would not produce a comparison group. With the propensity score method, the researcher does not choose the probability of treatment. Rather, the researcher matches paired observations (treated with control) with the same empirically estimated probability. In theory, a researcher should not find a match between a treated and control pair with a treatment probability of 0 or 1, hence matching is confined to where overlap is possible.
Propensity score matching as an unbiased treatment effect estimator (Rosenbaum and Rubin 1983). Given strong ignorability, the propensity score matched mean treatment effect is unbiased:

$$\begin{aligned}\tilde{\Delta} &= E_X\left[E[y|t=1,X]-E[\tilde{y}|t=0,X]\right]\\ &= E_{p(X)}\left[E[y|t=1,p(X)]-E[\tilde{y}|t=0,p(X)]\right]\\ &= E_{p(X)}\left[E[y|p(X)]-E[\tilde{y}|p(X)]\right]\\ &= \Delta\end{aligned} \tag{9}$$

where $E_{p(X)}$ is the expectation operator taken with regard to the distribution of $p(X)$. Thus, given strong ignorability (observable confounders) and the propensity score as a balancing score, the expected treatment effect is equal to the mean difference between the matched participants' and non-participants' outcomes, where observations are matched by their estimated probability of treatment. The importance of score matching versus covariate matching should again be duly noted. Score matching allows for pairing based on a scalar (probability to participate), while covariate matching pairs observations based on vectors of X; thus the greater the number of confounders, the more difficult it becomes to match observations (Dehejia and Wahba 2002).

Because the treatment effect is estimated using a matching procedure, the relationship between outcome and the treatment and confounders need not be known. In fact, PSM has been found to be robust when the outcome is a nonlinear function (LaLonde 1986; Winship and Mare 1992; Joffe and Rosenbaum 1999; Rubin and Thomas 2000; Imai and van Dyk 2004).

2.2 Propensity score models with continuous treatment

The majority of the PSM literature focuses on a binary treatment decision or multi-valued, ordinal treatment (for instance, see Behrman et al. 2004), but very recent work explores the use of propensity scores with treatment as a continuous variable (Imbens 2000; Lu et al. 2001; Hirano and Imbens 2004; Imai and van Dyk 2004). In the binary treatment case, treated observations are matched with non-treated observations based on their like propensity score (probability of treatment). When treatment is a continuous variable, rather than binary, all of the observations may have received some treatment and differ only with respect to their level of treatment. Unlike the binary case, there is no longer a control (untreated) group to compare with the treated group. However, some authors have taken the approach of discretizing the continuous treatment variable into several ordinal treatment groups (see Behrman et al. 2004). Of course the boundaries for the groups may be arbitrary and meaningless. Also with a continuous treatment variable, it may be more interesting (and meaningful) to examine the marginal effect of treatment.

When treatment is a continuous variable, matching gives way to modeling the outcome as a function of treatment while conditioning on the propensity score (Imai and van Dyk 2004). The requirements that the propensity score achieve strong ignorability when treatment is a continuous variable necessitates only a slight modification to that shown for matching.

Strong ignorability given propensity score when treatment is a continuous variable (Rosenbaum and Rubin 1983; Imai and van Dyk 2004). If treatment status is strongly ignorable given the confounders, then the treatment status is strongly ignorable given the propensity score; that is,

$$\begin{aligned}&\text{If } \tilde{y}, \tilde{y} + \Delta \perp t|X \text{ and if } 0 < E[t|X] \text{ for all } X,\\&\text{then } \tilde{y}, \tilde{y} + \Delta \perp t|E[t|X] \text{ and } 0 < E[t|E[t|X]] \text{ for all } E[t|X].\end{aligned}$$

In practical terms, the distribution of outcome is modeled using a parametric form while conditioning on the propensity score. The propensity score is defined as $E[t|X]$. The propensity score is also estimated parametrically, but no functional form is assumed for the relationship between the propensity score and the outcome. The two-step method proposed by Imai and van Dyk (2004) applied to (3) requires X to be observable (strong ignorability), so treatment can be estimated as a function of X, such as

$$E[t|X] = \delta_0 + \delta_x X \tag{10}$$

where the δ's are estimated using OLS. The linear form of (10) is not necessitated by the technique, only that the propensity score fitted from (10) is a balancing score. Blocking on the propensity score means that the observations are divided into P blocks (groups or sub-samples) and OLS estimation of y on Z and t is performed within each block (propensity score blocking ensures that the regressions are conditioned on the propensity score). The second stage estimates Eq. (3) while blocking on the estimated propensity score (the fitted values from (10)):

$$y_{pj} = \pi_p Z_{pj} + \Delta_p t_{pj} + \varepsilon_{pj} \tag{11}$$

where j references observations contained in the propensity score block p, where $p = 1, \ldots, P$. The average treatment effect is a weighted average (weighted by the number of observations within each block), calculated as:

$$\hat{\Delta} = \frac{1}{N} \sum_{p=1}^{P} \hat{\Delta}_p n_p \tag{12}$$

where n_p is the number of observations in block p.

When treatment is a continuous variable, the balancing condition $X \perp t | p(X)$ is difficult to test in the manner used when treatment is a binary variable. This is because treatment may take on any number of values instead of only two. Imai and van Dyk (2004) explain that the balancing condition may be tested by regressing a function of each of the continuous confounders on the treatment variable and the propensity score. Similar to the binary treatment case, the balancing test is a test of the propensity score model specification. A function of each confounder is used to test the propensity score specification, because by construction, regressing the untransformed covariate (the form used in the propensity score model) on treatment and the propensity score will yield an insignificant parameter estimate on treatment (if using simple linear regression). If the estimated parameter for treatment effect is significant, then treatment and the covariate are related and the specified propensity score is not a balancing score (the propensity score model is misspecified). However, any insignificance does not rule out the possibility that higher order transformations of the covariate are related to treatment (or a Type II error occurred in the test).

3 Data and study site

I examine the effectiveness of wildfire management on wildfires occurring in the St. Johns River Water Management District (SJRWMD) in northeast Florida, from 1996 to 2001. The SJRWMD area comprises portions of the 18 northeast counties in Florida. The SJRWMD is an ideal study area given its abundance of wildfire (in size and occurrence), use of prescribed fire, values at risk, and quality of available data. Florida, unlike many areas of the western United States, is heavily populated (it is fourth in state population and ninth in population density; United States Census Bureau 2004) and averages 218,638 wildfire acres a year and 763,205 acres of silvicultural-based prescribed fire a year (476,590 is for hazard reduction). Between wildfire and silvicultural prescribed fire, approximately 3% of Florida burns annually (not including agricultural burns). The SJRWMD averages over 48,596 acres a year in wildfire and 133,833 acres of prescribed burning (73,099 for hazard reduction) resulting in about 2.3% of the SJRWMD burning annually.

3.1 Wildfire records

Data on individual wildfire occurrences were obtained from the Florida Division of Forestry (FDOF). FDOF's wildfire data contains detailed information on fires found on private and state-owned lands including, but not limited to, the date and time of ignition, location (township, range, and cadastral section), size (acres), associated weather conditions (e.g., wind speed, wind direction, and humidity), rate of spread, flame length, and cause (arson, campfires, cigarettes, children, debris burning, equipment, lightning, miscellaneous, railroad, and unknown) from 1981 to 2001. Both the wildfire data and the prescribed burning data are geo-located to a Public Land Survey section (township, range, cadastral section), which is approximately a one-square mile rectangle.

From 1981 to 2001 there were 31,603 wildfires in the SJRWMD, with almost half of them (48%) coming from non-incendiary human-caused sources (accidental ignitions), followed by arson (29%), and lightning (23%). The SJRWMD (and Florida in general) fire season appears to begin late-winter to early-spring and last until the middle of the summer, with burned area peaking around May and June.

3.2 Wildfire management

The FDOF provided a second dataset that details all prescribed fire activities within the state (in order to conduct a prescribed burn in Florida, a permit must be obtained from the FDOF). Permit data include information on the location (located by the township, range, and cadastral section), reason (hazard reduction, prior to seeding, site preparation, disease control, wildlife, ecological, or other), and total size (in acres). The dataset includes permits issued between 1989 and 2001, although full state-wide reporting did not occur until 1993.

In addition, wildfire start time, wildfire report time (time between ignition and report to fire department), and fire crew arrival time, is provided in the FDOF database, allowing for the creation of a measure of initial attack (fire crew response time—the time between report and arrival).

3.3 Climate and weather

The El Niño Southern Oscillation (ENSO) measure used in this analysis is the Niño-3 sea surface temperature (SST) anomaly, which was obtained from the National Oceanic and Atmospheric Administration (National Oceanic and Atmospheric Administration 2002). The Niño-3 SST anomaly is measured as the positive (El Niño) or negative (La Niña) deviation, in degrees centigrade, of the Pacific sea surface temperature (at a specific location). The Keetch–Byram Drought Index (KBDI) was calculated for two weather stations in the SJRWMD region using daily data collected by the National Climate Data Center and provided by EarthInfo (2002).

3.4 Landscape characteristics

Section-level road and census data (population, income, and education) were created from U.S. Census Bureau TIGER/Line GIS data. The National Land Cover Data, based on the Multi-Resolution Land Characteristics Consortium's land cover map (30-m resolution grid) was used to determine landcover composition within and surrounding each section. Five landcover classes were assembled—grass (grassland/

herbaceous), upland forest (deciduous, evergreen, and mixed forest), urban (low intensity residential, high intensity residential, and commercial/industrial/transportation), water (open water), and wetland (woody wetland).

4 Empirical model

I examine the effectiveness of wildfire management on wildfires over the period 1996–2001. In this study, wildfire management is composed of two components—fire suppression and fuels management. The fire suppression (fire fighting) variable is defined as the amount of time (in hours) it takes fire crews to respond to a fire call (the time from when the fire is reported until the fire crew arrives at the fire site). It is commonly acknowledged that a swift first-response is an important step in controlling wildfire (Parks 1964; Gorte and Gorte 1979; Hirsch et al. 2004). The fuels management variable is defined as the number of acres of prescribed fire (for hazard reduction), over the last three years, permitted to the section, and those contiguous surrounding sections, in which the fire occurred. Full coverage of the prescribed fire data begins in 1993, therefore the analysis begins with wildfire occurring in 1996 to include for the three year lag. Because suppression and fuels management may affect both wildfire size (extent) and intensity, I use an intensity-weighted acres measure of wildfire behavior (as described in Mercer et al. 2007), measured in kW-acre/meters, for the wildfire variable. Wildfire intensity-weighted acres are used as a measure of fire behavior rather than acres or intensity alone because large acre fires may not be damaging if they are of low intensity. Similarly, very intense fires may not be damaging if they are very small. The intensity-weighted acre is policy-relevant because it is a better metric of fire damage and is defined as:

$$w_i = a_i^{*}259.833^{*}l_i^{2.174}$$

where w_{i} is wildfire intensity-weighted acres (kW-acre/meter), a_{i} is acres burned, l_{i} is fire flame length (meters). Wildfire intensity is approximated as a function of flame length ($= 259.833^{*}l_{\mathrm{i}}^{2.174}$; see Kennard (2004)), which is reported by FDOF.

Estimating the effectiveness of wildfire management using wildfire as the unit of observation (as oppose to a location) is appropriate in the case of suppression because suppression affects wildfire behavior, rather than the ignition process. Suppression does not prevent wildfire ignition, so sections without wildfire would also be without suppression (at least as suppression is defined here). If fuels management has a significant impact on ignition success then the current modeling framework introduces its own selection bias. However, the benefits of fuels management are often couched in terms of reduction in fire intensity and size rather than in terms of ignition (Wade et al. 2000; Brose and Wade 2002; Outcalt and Wade 2004). There is evidence that arson ignitions are correlated with fuels management activities (Prestemon and Butry 2005), but this phenomenon may have less to do with the physical relation between fuel levels and ignition and more to with the realization by the arsonist that the wildfire has a lower probability of achieving a devastating size or intensity in actively managed areas. In addition, actively managed forests may pose a greater threat of capture or detection to the arsonist. Therefore, the focus of this study is the effectiveness of wildfire management in wildfire prone areas.

4.1 Modeling wildfire using propensity scores

The PSM wildfire model is estimated using OLS while blocking on the propensity scores as:

$$\ln(w_b) = \beta_{0b} + \mathbf{X}_{Mb}\boldsymbol{\beta}_{Mb} + \mathbf{X}_{Cb}\boldsymbol{\beta}_{Cb} + \mathbf{X}_{Fb}\boldsymbol{\beta}_{Fb} + \mathbf{X}_{Ib}\boldsymbol{\beta}_{Ib} + \mathbf{X}_{Ob}\boldsymbol{\beta}_{Ob} + \xi \quad (13)$$

where w is wildfire-intensity acres (kW-acre/meter), $\mathbf{X}_{\mathrm{M}}$ (management variables) includes the natural log of the suppression (log of fire crew response time in hours) and prescribed fire (log of hazard reducing prescribed fire acres) variables, described above, along with an indicator variable denoting let burn fires (fires that are allowed to burn themselves out), $\mathbf{X}_{\mathrm{C}}$ (climate and weather variables) includes KBDI (0–800 drought index), whether the fire occurred in an El Niño or La Niña phase, and its magnitude (Pacific sea-surface temperature deviation from normal in centigrade), percent humidity, fire spread index (0–100 index; a function of wind, fuel moisture, fuel condition, and precipitation), actual fire spread in miles per hour (recorded as an ordinal variable), the natural log of wind speed (log of miles per hour), and wind direction, $\mathbf{X}_{\mathrm{F}}$ (fuel variables) includes indicator variables for fuel type (pine, swamp, blowy leaf, grass, muck, palmetto–gallberry, and other), build-up index (0–250 index; a measure of available fuels, temperature, and precipitation), latitude (decimal degrees re-projected into the Albers equal-area map projection coordinates), longitude (decimal degrees re-projected into the Albers Equal-area map projection coordinates), elevation (feet), slope (degrees), landscape composition (percent of landscape in upland forest, agriculture, rangeland, residential, wetland, and water), forest density (density of acres of forest in the section), previous 12-year fire history (number of ignitions), previous 12-year fire history in neighboring sections (number of ignitions), the natural log of previous 12-year fire intensity-weighted acres (a measure of previous fire behavior that may be related to available fuels), the natural log of previous 12-year fire intensity-weighted acres in neighboring sections, natural log of road density (density of kilometers of road in the section), $\mathbf{X}_{\mathrm{I}}$ (ignition variables) includes ignition cause (lightning, arson, campfire, cigarette, debris burning, equipment, railroad, children, miscellaneous, and unknown), natural log of population (log of count of population), natural log of family income (log of income in thousands of dollars; income is found to be correlated with wildfire in Florida; Butry et al. (2002)), and natural log of fire report time (log of time from fire ignition to call to fire department in hours), $\mathbf{X}_{\mathrm{O}}$ (other variables) includes indicator variables denoting county, fire year (begins September of the preceding calendar year and ends in October of the current calendar year), month, fire district, and ownership type (public or private), β's are parameters, b denotes the propensity score block, and ξ is an error term. Table 1 provides descriptive statistics for all variables in the models.

Because there are two arrays of propensity scores (one for response time, another for prescribed fire), placement into block groups is directed by the relative rank of the

Table 1 Descriptive statistics

Variables	Units	Mean or Sum	SD	Variables	Units	Mean or Sum	SD
ln(Wildfire intensity-weighted acres)		5.8749	(3.3285)	*Ignition*			
Management				Ignition cause (base = lightning)			
ln(Response time)	Hours	−5814.8091		Campfire	Count	111	
ln(Prescribed fire)	Acres	−31814.1645		Cigarette	Count	71	
Let burn	Count	230		Debris Burning	Count	900	
Climate & Weather				Arson	Count	1908	
KBDI	0–800, Index	422.3254	(174.0372)	Equipment	Count	248	
La Nina	Degrees centigrade	−0.2980	(0.2873)	Railroad	Count	76	
El Nino	Degrees centigrade	0.2545	(0.7401)	Children	Count	472	
Humidity	Percent	48.7856	(13.8323)	Unknown	Count	818	
Spread index	0–100, Index	20.3448	(13.3531)	Misc	Count	545	
Spread (base = not observed)	Ordinal variable			ln(Population)	Count	8.7548	(0.4648)
0 to 1 mph	Count	3883		ln(Income)	$ (000s)	10.8791	(0.2301)
2 mph	Count	880		ln(Report Time)	Hours	−0.5074	(1.5463)
3 mph	Count	311		Private ownership	Count	6851	
4 mph	Count	78		Federal ownership	Count	81	
^{3}5 mph	Count	117		*County (base = Alachua)*			
ln(Wind Speed)	Miles per hour	1.8924	(1.3735)	Baker	Count	411	
Wind Direction (base = calm)				Bradford	Count	8	
East	Count	1071		Brevard	Count	506	
North	Count	541		Clay	Count	526	
Northeast	Count	878		Columbia	Count	1	
Northwest	Count	924		Duval	Count	556	
South	Count	522		Flagler	Count	455	
Southeast	Count	757		Indian River	Count	235	
Southwest	Count	975		Lake	Count	507	
Variable	Count	714		Marion	Count	499	
West	Count	961		Nassau	Count	297	

Table 1 continued

Variables	Units	Mean or Sum	SD	Variables	Units	Mean or Sum	SD
Fuel				Okeechobee	Count	85	
Fuel Type (base = palmetto-gallberry)				Orange	Count	368	
Dense pine	Count	914		Osceola	Count	182	
Swamp	Count	493		Polk	Count	57	
Blowy leaf	Count	359		Putnam	Count	698	
Grassy fuels	Count	1446		St. Johns	Count	530	
Muck	Count	99		St. Lucie	Count	6	
Other	Count	392		Seminole	Count	143	
Vegetation build-up	0–250, Index	47.0861	(33.7113)	Sumter	Count	18	
Longitude	Decimal degree	642.0426	(47.4262)	Union	Count	1	
Latitude	Decimal degree	591.0584	(85.7503)	Volusia	Count	1118	
Elevation	Feet	17.4051	(13.0145)	*Fire district (base = District 6)*			
Slope	Degrees	0.2034	(0.2016)	Fire district 07	Count	1375	
Upland forest	Proportion	0.3427	(0.2500)	Fire district 08	Count	1453	
Agricultural lands	Proportion	0.1041	(0.1739)	Fire district 10	Count	2033	
Rangelands	Proportion	0.0716	(0.1123)	Fire district 11	Count	549	
Residential area	Proportion	0.1586	(0.2092)	Fire district 12	Count	1219	
Water	Proportion	0.0171	(0.0382)	Fire district 14	Count	67	
Wetland forest	Proportion	0.2178	(0.1827)	Fire district 15	Count	1	
Forest density	Density	48.6891	(21.5516)	Fire district 16	Count	332	
Previous fire ignitions	Count	3.6152	(5.5115)	*Fire year (base = 1996)*			
Previous neighboring fire ignitions	Count	4.0126	(4.0288)	Fire year 1997	Count	869	
ln(Previous fire intensity-acres)	Acres	5.6996	(6.5500)	Fire year 1998	Count	1466	
ln(Previous neighboring fire intensity-acres)	Acres	7.6474	(5.4685)	Fire year 1999	Count	1425	
ln(Road density)	Density	1.2167	(2.0996)	Fire year 2000	Count	1592	
ln(Previous non-hazard prescribed fire)	Acres	−5.8968	(3.1590)	Fire year 2001	Count	1136	

Table 1 continued

Variables	Units	Mean or Sum	SD	Variables	Units	Mean or Sum	SD
ln(Previous neighboring non-hazard Prescribed fire)	Acres	−3.9795	(4.9330)	*Month (base = October)*			
				November	Count	217	
				December	Count	254	
				January	Count	616	
				February	Count	765	
				March	Count	784	
				April	Count	815	
				May	Count	1052	
				June	Count	1386	
				July	Count	752	
				August	Count	529	
				September	Count	187	

Mean and standard deviation shown for continuous variables and sum shown for indicator variables. The units are listed for the untransformed variable. (Latitude and longitude are measured in decimal degrees re-projected into the Albers Equal-Area map projection using a Geographic Information System) two scores (low–low, low–mid, low–high, mid–low, mid–mid, mid–high, high–low, high–mid, high–high). Nine block groups are used. By conditioning on the propensity scores, the estimated relationship between response time and prescribed fire on wildfire intensity-acres represents the actual causal effect of management on fire behavior.

4.2 Propensity score estimators

The fitted values from propensity score estimator models, for both prescribed fire and response time, are the propensity scores. The response time and prescribed fire propensity scores are estimated separately, but are both used in a single model of wildfire behavior. The propensity score estimator models include all the variables in the wildfire equation that are either known at the time of treatment or unaffected by treatment (e.g., climate variables). This is done to ensure that all the covariates that directly affect wildfire behavior and the treatments (response time and prescribed fire) are included in the score estimator models.

The propensity score for fire crew response time is estimated as:

$$\ln(r) = \beta_0^r + \mathbf{X}_M^r \boldsymbol{\beta}_M^r + \mathbf{X}_C^r \boldsymbol{\beta}_C^r + \mathbf{X}_F \boldsymbol{\beta}_F^r + \mathbf{X}_I \boldsymbol{\beta}_I^r + \mathbf{X}_O \boldsymbol{\beta}_O^r + u \tag{14}$$

where r is fire crew response time (hours), the $\mathbf{X}$'s as above, except $\mathbf{X}_M^r$ (management variables) includes only the natural log of prescribed fire and an indicator variable for let burn fires (response time is excluded), $\mathbf{X}_C^r$ (climate and weather variables) includes only KBDI, whether the fire occurred in an El Niño or La Niña phase, and its magnitude, percent humidity, fire spread index, the natural log of wind speed, and wind direction (fire spread is excluded, as it may be affected by suppression), β's are parameters, and u is an error term. Fire crew response time is modeled using OLS.

The propensity score estimator for prescribed fire is estimated as:

$$\ln(p) = \beta_0^p + \mathbf{X}_C^p \boldsymbol{\beta}_C^p + \mathbf{X}_F^p \boldsymbol{\beta}_F^p + \mathbf{X}_I^p \boldsymbol{\beta}_I^p + \mathbf{X}_O \boldsymbol{\beta}_O^p + v \tag{15}$$

where p is prescribed fire acres, the $\mathbf{X}$'s as above, except $\mathbf{X}_C^p$ (climate and weather variables) includes whether the fire occurred in an El Niño or La Niña phase, and its magnitude (excludes KBDI, percent humidity, fire spread index, actual fire spread in miles per hour, the natural log of wind speed, and wind direction), $\mathbf{X}_F^p$ (fuel variables) includes only fuel type, latitude, longitude, elevation, slope, landscape composition, forest density, previous 12-year fire history, previous 12-year fire history in neighboring sections, the natural log of previous 12-year fire intensity-weighted acres, the natural log of previous 12-year fire intensity-weighted acres in neighboring sections, natural log of road density (excludes vegetation build-up), $\mathbf{X}_I^p$ (ignition variables) only includes population, income, and ownership (excludes ignition cause and natural log of fire report time), β's are parameters, and v is an error term.

4.3 Covariate balancing test

To test the covariate balance in the response time propensity score estimator model (prescribed fire propensity score estimator model), I square each continuous variable and regress it on the natural log of response time (natural log of prescribed fire) and the estimated response time propensity score (estimated prescribed fire propensity score). If the coefficient on the natural log of response time (natural log of prescribed fire) is statistically indistinguishable from zero, then the covariate is deemed balanced, if not, then the squared term is added to the propensity score estimator model. The squared-term of the natural log of previous wildfire intensity-weighted acres, natural log of previous wildfire intensity-weighted acres in neighboring areas, and natural log of fire report time are included in the response time propensity score estimator model due to lack of initial balance.

Figure 2 shows the balancing ability of the propensity scores. If the propensity score balances the covariates and if the covariates include all the variables that confound the treatment-outcome relationship, then the propensity score blocking estimator will be unbiased. The balancing test is described in Sect. 2.2. If the propensity score is a balancing score at a 5% significance level, then the t-statistic associated with treatment (response time and prescribed fire) when regressing a nonlinear transformation of the covariate on treatment and the propensity score should be less than 1.96. Figure 2

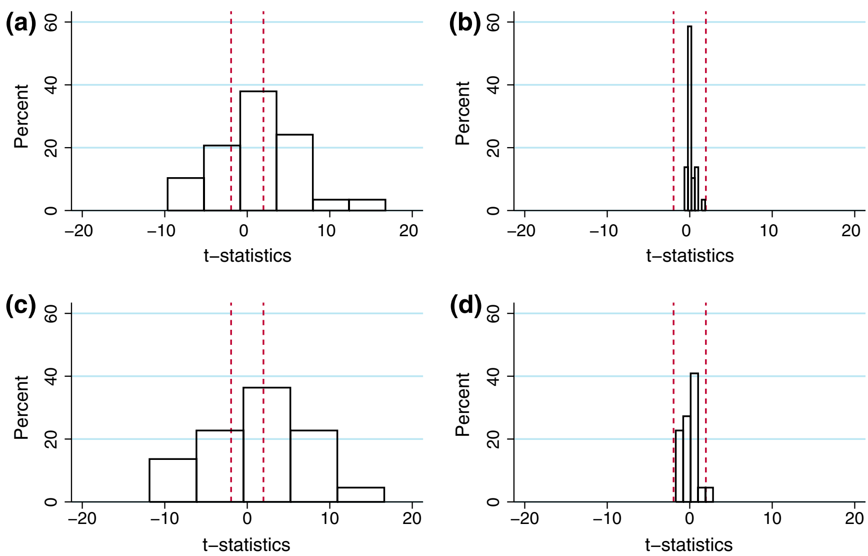

Fig. 2 Histograms of the t-statistics associated with the management variable (ln(response time), ln(prescribed fire)) from regressing the squared value of each of the continuous variables in the ln(response time) propensity score estimator model (**a**) on ln(response time), (**b**) on ln(response time) and the ln(response time) propensity score, and from regressing each of the continuous variables in the ln(prescribed fire) propensity score estimator model (**c**) on ln(prescribed fire), (**d**) on ln(prescribed fire) and the ln(prescribed fire) propensity score. Dashed line denotes 95% conference interval (t-statistic equal to zero)

presents the t-statistics associated with regressing a nonlinear transformation of the covariates on (a) the natural log of response time, (b) the natural log of response time and the propensity score of the natural log of response time, (c) the natural log of prescribed fire, (d) the natural log of prescribed fire and the propensity score of the natural log of prescribed fire. The relevant comparisons to make are between 2a and 2b and between 2c and 2d. Figure 2a shows that some of the covariates are related to response time, whereas figure 2b shows that the covariates are not statistically related (at the 5% significance level) to response time when conditioning on the response time propensity score. This implies that conditional on the propensity score, response time is not related to the covariates—this is the condition that the propensity score must achieve to be a balancing score—so that conditioning on the response time propensity score ensures that differences in wildfire behavior are due to differences in response time and not to any underlying factors. Likewise, Figures 2c and d show that the prescribed fire propensity score is also a balancing score.

5 Results

Tables 2–4 summarize the output of the models. Tables 2 and 3 present the 1st stage propensity score estimator models. Table 4 summarizes the impact of response time and prescribed fire on wildfire intensity-acres, rather than presenting all nine subgroup regressions (available upon request). The 1st stage propensity score model models are significant (based on the F-tests), but explain a limited amount of the variation of the respective management variable (12–18%). Because the wildfire model is estimated using the natural log of wildfire intensity-acres, response time, and prescribed fire variables, the reported management coefficients are elasticities. For interpretation of the dummy variables in semilogarithmic models, see Halvorsen and Palmquist (1980) and Kennedy (1981).

The results of the propensity score model are complex, as they allow for nonlinearity between wildfire management and wildfire intensity-acres. I find that a 1% decrease in response time yields a 0.3175% decrease (0.0420 standard error) in wildfire intensity-acres (based on the average response time elasticity averaged over all subgroups). The average prescribed fire elasticity (again averaged over all subgroups) implies a 1% increase in prescribed fire acres yields a 0.0138% decrease (0.0085 standard error) in wildfire intensity-acres. Upon examination of the subgroup results, there appear some interesting findings. First, the propensity scores in this analysis can be thought of as the expected levels of treatment. The response time propensity scores in the lower third subgroup correspond to those fires with quicker expected fire crew response rates, and those in the upper third correspond with fires with slower expected fire crew response rates. The lower third subgroup prescribed fire propensity scores correspond with fires with less expected prior prescribed burn acres, while those fires in the upper third subgroup correspond with more expected area burned by prescribed fire. Prescribed fire is significant (5% level) in the group with median levels of expected prescribed fire and with lower expected response times. The prescribed fire elasticity is −0.0436. It is weakly significant (11% level) in the group with higher expected levels of prescribed fire and with lower expected response times. Quick fire crew response times

Table 2 1st propensity score model, natural log of response time propensity score estimator

Variables	Coefficient	SE	Variables	Coefficient	SE
Intercept	3.2206	(0.9750)***	*Ignition*		
Management			Ignition cause (base = lightning)		
Let burn	0.0735	(0.0545)	Campfire	−0.0329	(0.0798)
ln(Prescribed fire)	−0.0015	(0.0021)	Cigarette	−0.2339	(0.0978)**
Climate &weather			Debris burning	−0.1472	(0.0394)***
KBDI	0.0000	(0.0001)	Arson	−0.0688	(0.0317)**
La Nina	0.0054	(0.0609)	Equipment	−0.1210	(0.0567)**
El Nino	0.0287	(0.0183)	Railroad	−0.1175	(0.0946)
Humidity	0.0006	(0.0008)	Children	−0.1776	(0.0463)***
Spread index	−0.0013	(0.0012)	Unknown	−0.0567	(0.0384)
ln(Wind speed)	−0.0186	(0.0200)	Misc	−0.1223	(0.0412)***
Wind direction (base = calm)			ln(Population)	0.0100	(0.0223)
East	0.3666	(0.1856)**	ln(Income)	0.0567	(0.0484)
North	0.3962	(0.1873)**	ln(Report time)	0.0608	(0.0063)***
Northeast	0.3481	(0.1861)*	Private ownership	−0.0041	(0.0374)
Northwest	0.3309	(0.1881)*	Federal ownership	−0.2052	(0.0951)**
South	0.3082	(0.1871)*	*County (base = Alachua)*		
Southeast	0.3927	(0.1863)**	Baker	−0.4505	(0.2083)**
Southwest	0.3447	(0.1868)*	Bradford	−0.3936	(0.3426)
Variable	0.3315	(0.1845)*	Brevard	−0.4191	(0.2056)**
West	0.4149	(0.1869)**	Clay	−0.2733	(0.1603)*
Fuel			Columbia	0.2635	(0.8101)
Fuel type (base = palmetto-gallberry)			Duval	−0.2204	(0.1649)
Dense pine	0.0028	(0.0308)	Flagler	−0.6736	(0.1775)***
Swamp	0.1134	(0.0394)***	Indian river	−0.7456	(0.3242)**
Blowy leaf	−0.0702	(0.0467)	Lake	−0.6212	(0.1861)***
Grassy fuels	−0.1717	(0.0291)***	Marion	−0.3932	(0.0696)***
Muck	0.1427	(0.0813)*	Nassau	−0.2213	(0.1717)
Other	−0.0370	(0.0445)	Okeechobee	−0.5677	(0.3267)*
Vegetation build-up	0.0001	(0.0004)	Orange	−0.1469	(0.1916)
Longitude	−0.0021	(0.0009)**	Osceola	−0.4555	(0.2103)**

Table 2 Continued

Variables	Coefficient	SE	Variables	Coefficient	SE
Latitude	−0.0047	(0.0006)***	Polk	−1.1278	(0.3298)***
Elevation	−0.0048	(0.0015)***	Putnam	−0.0266	(0.0670)
Slope	0.0133	(0.0579)	St. Johns	−0.3309	(0.1730)*
Upland forest	0.0471	(0.0768)	St. Lucie	−0.8511	(0.4593)*
Agricultural lands	−0.0669	(0.0828)	Seminole	−0.2513	(0.1957)
Rangelands	0.0728	(0.1083)	Sumter	−0.4931	(0.2602)*
Residential area	0.0016	(0.0788)	Union	0.0044	(0.8096)
Water	0.3143	(0.2594)	Volusia	−0.7402	(0.1805)***
Wetland forest	0.1001	(0.0868)	*Fire district (base = District 6)*		
Forest density	−0.0017	(0.0006)***	Fire district 07	0.0152	(0.1316)
Previous fire ignitions	−0.0033	(0.0020)*	Fire district 08	−0.4602	(0.1983)**
Previous neighboring fire ignitions	−0.0046	(0.0028)	Fire district 10	0.1972	(0.1882)
ln(Previous fire intensity-acres)	−0.0092	(0.0019)***	Fire district 11	−0.2641	(0.2236)
ln(Previous neighboring fire intensity-acres)	−0.0026	(0.0025)	Fire district 12	−0.6047	(0.2088)***
ln(Road density)	−0.0160	(0.0053)***	Fire district 14	0.1000	(0.3159)
ln(Previous non-hazard prescribed fire)	0.0028	(0.0030)	Fire district 15	−1.7740	(0.8495)**
ln(Previous neighboring non-hazard prescribed fire)	−0.0025	(0.0026)	Fire district 16	−0.4900	(0.3208)
			Fire year (base = 1996)		0.0000
Balance			Fire year 1997	−0.0109	(0.0449)
ln(Previous fire intensity-acres)^2	0.0009	(0.0002)***	Fire year 1998	0.0176	(0.0461)
ln(Previous neighboring fire intensity-acres)^2	0.0010	(0.0002)***	Fire year 1999	−0.0448	(0.0438)
ln(Report time)^2	−0.0062	(0.0024)**	Fire year 2000	−0.0567	(0.0379)
			Fire Year 2001	−0.1197	(0.0386)***
			Month (base = October)		
			November	−0.0267	(0.0882)
			December	0.0086	(0.0870)
			January	−0.0764	(0.0783)
			February	−0.0829	(0.0795)

Table 2 Continued

Variables	Coefficient	SE	Variables	Coefficient	SE
			March	−0.1190	(0.0807)
			April	−0.1151	(0.0789)
			May	−0.1346	(0.0781)*
			June	−0.1504	(0.0794)*
N	7480		July	−0.0803	(0.0798)
F	9.4300		August	0.0214	(0.0806)
R-Squared	0.1184		September	0.0204	(0.0914)

***, ***, * Denotes significance at the 0.01, 0.05, 0.10 level, respectively

Table 3 1st propensity score model, natural log of prescribed fire propensity score estimator

Variables	Coefficient	SE	Variables	Coefficient	SE
Intercept	−1.6920	(5.3223)	*Ignition*		
Climate &weather			ln(Population)	−0.0565	(0.1239)
La Nina	0.1630	(0.3304)	ln(Income)	1.1880	(0.2679)***
El Nino	0.4719	(0.0936)***	Private ownership	−1.7712	(0.2057)***
Fuel			Federal ownership	0.5713	(0.5280)
Fuel type (base = palmetto-gallberry)			*County (base = Alachua)*		
Dense pine	−0.0378	(0.1703)	Baker	2.7421	(1.1555) **
Swamp	−0.0158	(0.2176)	Bradford	5.5685	(1.9012)***
Blowy leaf	−0.0025	(0.2579)	Brevard	6.4124	(1.1381)***
Grassy fuels	−0.4219	(0.1584)***	Clay	1.5324	(0.8888)*
Muck	−0.0434	(0.4494)	Columbia	−4.9883	(4.5007)
Other	0.1660	(0.2409)	Duval	−0.5123	(0.9132)
Longitude	−0.0274	(0.0051)***	Flagler	4.8811	(0.9833)***
Latitude	0.0050	(0.0035)	Indian river	5.4507	(1.7989)***
Elevation	−0.0130	(0.0083)	Lake	1.6054	(1.0324)
Slope	−0.3241	(0.3211)	Marion	0.8263	(0.3854) **
Upland forest	0.8549	(0.4238) **	Nassau	−0.2038	(0.9519)
Agricultural lands	1.1064	(0.4562) **	Okeechobee	5.3113	(1.8134)***
Rangelands	2.5178	(0.5985)***	Orange	5.2736	(1.0620)***
Residential area	1.1339	(0.4355)***	Osceola	6.7939	(1.1637)***
Water	−0.2857	(1.4396)	Polk	4.0212	(1.8287) **
Wetland forest	0.6703	(0.4803)	Putnam	1.2965	(0.3714)***
Forest density	0.0115	(0.0035)***	St. Johns	1.8569	(0.9596)*
Previous fire ignitions	−0.0174	(0.0106)	St. Lucie	4.8394	(2.5469)*
Previous neighboring fire ignitions	0.0810	(0.0152)***	Seminole	6.0739	(1.0831)***
ln(Previous fire intensity-acres)	−0.0016	(0.0088)	Sumter	1.1778	(1.4444)
ln(Previous neighboring fire intensity-acres)	0.0081	(0.0110)	Union	−2.8737	(4.4971)
ln(Road density)	−0.0374	(0.0295)	Volusia	5.0361	(0.9999)***
ln(Previous non-hazard prescribed fire)	0.1458	(0.0168)***	*Fire district (base = District 6)*		
ln(Previous neighboring non-hazard	0.2123	(0.0144)***	Fire district 07	−0.6148	(0.7284)

Table 3 Continued

Variables	Coefficient	SE	Variables	Coefficient	SE
Prescribed fire)			Fire district 08	−0.9484	(1.0992)
			Fire district 10	−1.7482	(1.0416)*
			Fire district 11	−1.1342	(1.2389)
			Fire district 12	−3.7378	(1.1556)***
			Fire district 14	−3.0532	(1.7506)*
			Fire district 15	−4.3040	(4.7209)
			Fire district 16	−2.4835	(1.7790)
			Fire year (base = 1996)		
			Fire year 1997	−0.8950	(0.2245)***
			Fire year 1998	−1.8606	(0.2375)
			Fire year 1999	0.3916	(0.2110)*
			Fire year 2000	0.8359	(0.1866)***
			Fire year 2001	0.2632	(0.2013)
			Month (base = October)		
			November	0.3991	(0.4871)
			December	0.4495	(0.4786)
			January	0.5458	(0.4267)
			February	0.9517	(0.4318) **
			March	0.6258	(0.4392)
			April	0.5679	(0.4309)
			May	0.1785	(0.4255)
			June	0.1877	(0.4229)
N	7480		July	0.1910	(0.4297)
F	20.9600		August	0.2253	(0.4355)
R-Squared	0.1771		September	−0.1388	(0.5018)

***, **, * Denotes significance at the 0.01, 0.05, 0.10 level, respectively

Table 4 Propensity score wildfire model results showing the treatment effect and associated standard errors (in parentheses)

Response time propensity score	Prescribed fire propensity score		
	Lower third	Middle third	Upper third
Lower third	ln(Response time) = 0.4176*** (0.1091) ln(Prescribed fire) = −0.0005 (0.0279) R2 = 0.4651 n = 850	ln(Response Time) = 0.1235 (0.1109) ln(Prescribed Fire) = −0.0436** (0.0206) R2 = 0.5671 n = 731	ln(Response time) = 0.1026 (0.1111) ln(Prescribed fire) = −0.0268 (0.0169) R2 = 0.4560 n = 913
Middle third	ln(Response time) = 0.2858** (0.1265) ln(Prescribed fire) = −0.0246 (0.0333) R2 = 0.3966 n = 906	ln(Response time) = 0.3649*** (0.1317) ln(Prescribed Fire) = −0.0266 (0.0225) R2 = 0.4827 n = 868	ln(Response time) = 0.2900** (0.1470) ln(Prescribed fire) = −0.0136 (0.0220) R2 = 0.4566 n = 719
Upper third	ln(Response time) = 0.5097*** (0.1366) ln(Prescribed fire) = −0.0216 (0.0367) R2 = 0.4470 n = 738	ln(Response time) = 0.4874*** (0.1216) ln(Prescribed fire) = 0.0357 (0.0224) R2 = 0.4304 n = 894	ln(Response time) = 0.2787** (0.1363) ln(Prescribed fire) = −0.0084 (0.0203) R2 = 0.4150 n = 861

***, **, * Denotes significance at the 0.01, 0.05, 0.10 level, respectively. The weighted average across blocks is 0.3175 (0.0420 standard error) and −0.0138 (0.0085 standard error) for response time and prescribed fire, respectively are generally significant across all subgroups, except those groups with significant prescribed fire effects. Focusing only on the subgroups with significant results, the weighted average elasticity is 0.3754, with subgroup variation ranging from 0.2787 to 0.5097.

Across the nine subgroups, the strength of the treatment effect of one of the management variables appears to be inversely related to size of the propensity score of the other. This implies that the effect of response time on fire intensity-acres is greater for fires where less expected prescribed fire occurs. Further, prescribed fire appears to be effective only for those fires with quicker expected fire crew response times. The prescribed fire treatment effect is significant (or weakly significant) only in the groups where the treatment effect for response time is not, and vice-versa. Because wildfires appear to be affected by either suppression or prescribed fire, but not both, this suggests that prescribed fire and suppression act as substitutes rather than complements.

6 Discussion

The question examined here is whether or not wildfire management pays off. Statistical difficulties in modeling wildfire behavior at policy relevant scales make quick answers challenging. After accounting for potential endogeneity and nonlinearities of prescribed fire and fire crew response time with wildfire behavior, I find evidence that quicker response times limit wildfire size and intensity, and that prescribed fire may provide beneficial effects against wildfire extent and intensity up to three years of its application, especially in combination with quick response time.

The propensity score blocking model yields evidence that wildfire management (fire crew response time and prescribed fire) is effective at limiting wildfire size and intensity. The model findings also suggest nonlinearity between wildfire behavior and wildfire management. This result implies that the returns to management may not be homogeneous across the landscape and that these models could be used by policymakers to target (locate) areas with the highest returns to wildfire management investment.

Prestemon et al. (2002) and Mercer et al. (2007) both evaluate the effectiveness of prescribed fire on wildfire behavior. I am not aware of any other research that has quantified the effectiveness of fire crew response time, initial attack, or suppression in general on wildfire behavior. Overall, Prestemon et al. (2002) finds very little statistical evidence that hazard mitigating prescribed fire reduces wildfire size. In fact, they find some significant positive correlations between prescribed fire and fire size. They state a possible explanation is an omitted variable problem (unobserved confounder). Mercer et al. (2007) find significant correlations between hazard mitigating prescribed fire and wildfire intensity-weight acres. The average prescribed fire treatment effect elasticity (the average prescribed fire effect over three years) is 0.27. Because the prescribed fire variable used in Mercer et al. (2007) is defined as the natural log of the ratio of prescribed fire to forested area (in acres), a 1% increase in the ratio of prescribed fire to forested area is associated with a 0.27% reduction in wildfire intensity-weighted acres. Based on the prescribed fire and forest acre data present in their table 1 (Volusia County, Florida, 1994–2001), a 1% increase in prescribed fire acres is associated with a 0.0110% reduction in wildfire intensity-weighted acres. The significant block group treatment effects in the propensity score model are significantly different than the Mercer et al. (2007) finding. This research coupled with Prestemon et al. (2002) and Mercer et al. (2007) suggest that prescribed fire does appear to limit wildfire intensity, but may be less effective at limiting fire size. It also suggests that OLS models may under-estimate the true impact of prescribed fire, as compared with propensity score models.

The estimated financial returns to wildfire management appear great, but have large margins of error. From 1993 to 2001, approximately 669,290 acres were prescribed burned in the SJRWMD for hazard mitigation purposes. Cleaves et al. (2000) estimate the cost of prescribed fire to be \$26.30 an acre in the southeast U.S., thus roughly \$17,602,327 was spent on prescribed fire in the SJRWMD. Based on the propensity score blocking prescribed fire elasticity of -0.0138, the SJRWMD experienced 30,651,129.12 kW-acre/meters less wildfire (1.38% less). Mercer et al. (2007) estimate the market value of damage from wildfire to be \$0.88 per kW-acre/meter, based on the findings of Butry (2001). Using this value as an approximation of the unit cost of wildfire, prescribed fire in the SJRWMD saved \$26,972,993.63 in wildfire damages. Thus, for every \$1 spent on prescribed fire treatments, \$1.53 in wildfire damage was avoided. But given the elasticity variations across propensity score groups, the savings from prescribed fire applications are likely to vary geographically.

It is not possible to calculate a similar estimate for suppression response (it is not known how large a fire would have become without any response). However, it is possible to estimate what the financial return would have been, from 1996 to 2001, if suppression response would have been 10% quicker. Using the propensity score blocking estimated response time elasticity of 0.3175 and the fire damage estimate of \$0.88 kW-acre/meter, a 10% decrease in response time (a total reduction of 543.7 hours) would have resulted in \$62,057,420 reduction in fire damages. Unfortunately, a dollar equivalent to the 543.7 hour reduction is not available. Future research might focus on evaluating the costs associated with suppression response.

In addition, future research should focus on exploring other methods (perhaps less parametric) to condition wildfire behavior on propensity score, and explore the role spatial scale plays in the modeling of the effectiveness wildfire management. The spatial scale used here is much finer than Prestemon et al.(2001, 2002), which finds fairly weak evidence of the effectiveness of prescribed fire. However, the challenge is to find the appropriate modeling scale that provides insight into the wildfire problem, while being policy relevant.

Acknowledgements This research was funded by the National Fire Plan (USDA Forest Service) as part of a project titled, “Quantifying Trade-offs of Alternative Vegetation Management Strategies, Wildfire, and Suppression in Fire Prone Regions of the U.S.,” while the author was employed with the Southern Research Station, USDA Forest Service. The author thanks Marcia L. Gumpertz, Professor, Department of Statistics, North Carolina State University; Walter N. Thurman, Professor, Department of Economics, North Carolina State University; and Jeffrey P. Prestemon, Research Forester, Southern Research Station, USDA Forest Service for their helpful comments in the development of this manuscript. Any remaining errors in this article are the responsibility of the author.

References

Amacher GS, Malik AS, Haight RG (2005) Not getting burned: the importance of fire prevention in forest management. Land Econ 81:284–302

Behrman JR, Cheng Y, Todd PE (2004) Evaluating preschool programs when length of exposure to the program varies: a nonparametric approach. Rev Econ Stat 86:108–122

Bellman RE (1961) Adaptive control processes. Princeton University Press, Princeton

Brose P, Wade D (2002) Potential fire behavior in pine flatwood forests following three different fuel reduction techniques. For Ecol Manage 163:71–84

Butry DT (2006) Estimating the efficacy of wildfire management using propensity scores. Ph.D. Dissertation, North Carolina State University, p 110

Butry DT, Mercer DE, Prestemon JP, Pye JM, Holmes TP (2001) What is the price of catastrophic wildfire? J Forestry 99:9–17

Butry DT, Pye JM, Prestemon JP (2002) Prescribed fire in the interface: separating the people from the trees. In: Outcalt KW (ed) Proceedings of the eleventh biennial southern silvicultural research conference. General technical report SRS-S48, Department of Agriculture, Forest Service, Southern Research Station, Asheville, pp 132–136

Cleaves DA, Martinez J, Haines TK (2000) Influences of prescribed burning activity and costs in the national forest system. General technical report SRS-S37, U.S. Department of Agriculture Forest Service

Davis LS (1965) The economics of wildfire protection with emphasis on fuel break systems. Resources Agency, Department of Conservation, Division of Forestry, Sacramento, State of California

Dehejia RH, Wahba S (1999) Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. J Am Stat Assoc 94:1053–1062

Dehejia RH, Wahba S (2002) Propensity score-matching methods for nonexperimental causal studies. Rev Econ Stat 84:151–161

EarthInfo, Inc (2002) NCDC first order summary of the day. Data on CD

Gorte JK, Gorte RW (1979) Application of economic techniques to fire management: a status review and evaluation. General technical report INT-53, USDA Forest Service, Ogden

Greene WH (2000) Econometric analysis. Upper Saddle River, New Jersey p 1004

Halvorsen R, Palmquist R (1980) The interpretation of dummy variables in semilogarithmic equations. Am Econ Rev 70:474–475

Heckman JJ, Hotz VJ (1989) Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training. J Am Stat Assoc 84:862–874

Heckman JJ, Robb R (1985) Alternative methods for evaluating the impact of interventions. In: Heckman J, Singer B (eds) Longitudinal analysis of labor market data. Cambridge University Press, New York pp 156–246

Hirano K, Imbens GW (2004) The Propensity score with continuous treatment. Draft of chapter for missing data and Bayesian methods in practice: contributions from Donald Rubin's statistical family, Forthcoming from Wiley

Hirsch KG, Podur JJ, Janser RF, McAlpine RS, Martell DL (2004) Productivity of Ontario initial-attack fire crews: results of an expert-judgement elicitation study. Can J For Res 34:705–715

Imai K, Dyk DAvan (2004) Causal inference with general treatment regimes: generalizing the propensity score. J Am Stat Assoc 99:854–866

Imbens GW (2004) Nonparametric estimation of average treatment effects under exogeneity: a review. Rev Econ Stat 86:4–29

Imbens GW (2000) The role of the propensity score in estimating dose-response functions. Biometrika 87:706–710

Joffe MM, Rosenbaum PR (1999) Propensity scores. Am J Epidemiol 150:327–333

Kennard DK (2004) Depth of Burn. Forest Encyclopedia, Available at http://www.forestencyclopedia.net

Kennedy PE (1981) Estimation with correctly interpreted dummy variables in semilogarithmic equations. Am Econ Rev 71:801

LaLonde R (1986) Evaluating the econometric evaluations of training programs with experimental data. Am Econ Rev 76:604–620

Lu B, Zanutto E, Hornik R, Rosenbaum PR (2001) Matching with dose in an observational study of a media campaign against drug abuse. J Am Stat Assoc 96:1245–1253

Mercer DE, Prestemon JP, Butry DT, Pye JM (2007) Evaluating alternative prescribed burning policies to reduce net economic damages from wildfire. Am J Agri Econ 89:63–77

National Oceanic and Atmospheric Administration. (2002) El Niño-Southern oscillation sea surface temperature measures. Available at <ftp://ftp.ncep.noaa.gov/pub/cpc/wd52dg/data/indices/sstoi.indices>. Accessed by author on October, 2002
Outcalt KW, Wade DD (2004) Fuels management reduces tree mortality from wildfires in Southeastern United States. South J Appl Forestry 28:28–34
Parks GM (1964) Development and application of a model for suppression of forest fires. Manage Sci 10:760–766
Prestemon JP, Butry DT (2005) Time to burn: modeling wildland arson as an autoregressive crime function. Am J Agri Econ 87:756–770
Prestemon JP, Mercer DE, Pye JM, Butry DT, Holmes TP, Abt KL (2001) Economically optimal wildfire intervention regimes. Paper presented at the annual meeting of the American agricultural economics association, August 5–8, 2001, Chicago, Illinois. 18 pages. Published on the Internet, <http://agecon.lib.umn.edu/cgi-bin/pdf_view.pl?paperid=2872>
Prestemon JP, Pye JM, Butry DT, Holmes TP, Mercer DE (2002) Understanding broad scale wildfire risks in a human-dominated landscape. Forest Sci 48:685–693
Reed WJ (1987) Protecting a forest against fire: optimal protection patterns and harvest policies. Nat Res Model 2:23–53
Rideout DB, Omi PN (1990) Alternative expressions for the economic theory of forest fire management. Forest Sci 36:614–624
Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70:41–55
Rosenbaum PR, Rubin DB (1985) The bias due to incomplete matching. Biometrics 41:103–116
Rosenbaum PR, Rubin DB (1985) Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat 39:33–38
Rubin DB (1990) Formal modes of statistical inference for causal effects. J Stat Plan Inf 25:279–292
Rubin DB, Thomas N (2000) Combining propensity score matching with additional adjustments for prognostic covariates. J Am Stat Assoc 95:573–585
Smith J, Todd P (2005) Does matching overcome lalonde's critique of nonexperimental estimators. J Econ 125:305–353
Sparhawk WN (1925) The use of liability ratings in planning forest fire protection. J Agri Res 30: 693–762
Stewart S, Radeloff V, Hammer R, Fried J, Holcomb S, McKeefry J (2005) Mapping the wildland urban interface and projecting its growth to 2030. USDA Forest Service, North Central Research Station
United States Census Bureau (2004) http://www.census.gov
Wade DD, Brock BL, Brose PH, Grace JB, Hoch GA, Patterson WA, III (2000) Chapter 4: Fire in eastern ecosystems. In: Brown JB, Smith JK (eds) Wildland fire in ecosystems: effects of fire on flora. General technical report RMRS-GTR-42-vol2. USDA Forest Service, Rocky Mountain Research Station, Ogden, p 257
Winship C, Mare RD (1992) Models for sample selection bias. Ann Rev Soc 18:327–350
Yoder J (2004) Playing with fire: endogenous risk in resource management. Am J Agri Eco 86:933–948

Author Biography

David T. Butry is an Economist with the Office of Applied Economics, Building and Fire Research Laboratory, National Institute of Standards and Technology in Gaithersburg, Maryland. He received his Ph.D. in Economics from the North Carolina State University in 2006 and has served as an Economist with the Southern Research Station, USDA Forest Service in Research Triangle Park, North Carolina from 1999–2006. His current research interests include the economics of natural and man-made disasters, spatial statistics, and program evaluation econometrics.