Journal of Forest Economics 26 (2017) 46–55

Contents lists available at ScienceDirect

Journal of Forest Economics

journal homepage: www.elsevier.com/locate/jfe

Forest pests and home values: The importance of accuracy in damage assessment and geocoding of properties

Klaus Moeltner[a,*], Christine E. Blinn[b], Thomas P. Holmes[c]

[a] Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA, USA
[b] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA
[c] USDA Forest Service, Southern Research Station, Research Triangle Park, NC, USA

ARTICLE INFO

Article history:
Received 6 December 2016
Accepted 2 February 2017

Keywords:
Mountain Pine Beetle
Matching estimator
Measurement error
Aerial Detection Survey

ABSTRACT

We examine the impact of measurement errors in geocoding of property locations and in the assessment of Mountain Pine Beetle-induced tree damage within the proximity of a given residence on estimated losses in home values. For our sample of homes in the wildland-urban interface of the Colorado front range and using a novel matching estimator with Bayesian regression adjustment we find that both types of errors can lead to substantial biases in estimated losses. Our results confirm that the Forest Service's Aerial Detection Survey is generally too coarse to be informative for property valuation that depends on highly localized spatial data.

© 2017 Department of Forest Economics, Swedish University of Agricultural Sciences, Umeå. Published by Elsevier GmbH. All rights reserved.

Introduction

The accuracy of address-based geocoding (ABG), as used in many geographic information system (GIS) software packages, has been subject to scrutiny in recent years by spatially focused disciplines such as geography and epidemiology. The emerging consensus in these areas is that ABG can lead to measurable positional errors, when compared to ground-truthing or interpretation via aerial imagery. Furthermore, these errors tend to be more pronounced in rural areas with lower population density.

For example, Cayo and Talbot (2003) examine 3000 addresses of residential properties in upstate New York and find a median error, in terms of the Euclidean distance between the true and the presumed location of the actual housing structure, of 38 meters (m), 78 m, and 201 m in urban, suburban, and rural settings, respectively. Schootman et al. (2007), in turn, report a mean error of 42 m and a median deviation of 31 m for 299 addresses in urban St. Louis, Missouri. Zandbergen (2007), considering over 100,000 addresses in urban/suburban Orange County, Florida, discovers a mean error of 66 m, with a standard deviation of over 430 m.

As mentioned in Burra et al. (2002), and discussed in detail in Zandbergen (2007) these positional errors can introduce serious bias into an underlying statistical analysis if they are systematic in nature, which is often the case in a given empirical application. For example, in Zandbergen (2007)'s context ABG generally "moves" residential locations closer to major traffic arteries than they really are. This leads to an over-estimation of the number of children exposed to traffic-related air pollution. Similar biases are reported in Harada and Shimada (2006) and in Hay et al. (2009) in the context of spatial densities and hot spots for crime locations.

Somewhat surprisingly given the recent proliferation in spatially explicit property valuation studies, the potential inaccuracy of ABG has largely been ignored in the applied economics literature. As we illustrate in this study, pinpointing the accurate location of a private residence can be very important if the effect of environmental changes in the immediate vicinity of a home constitutes the central focus of a given analysis.

In our application we consider tree mortality induced by the Mountain Pine Beetle (MPB) within a close perimeter of a given residence in the wildland-urban interface (WUI) of the Colorado front range. The MPB (*dentroctonus ponderosae*) is a native forest pest that has shown explosive population growth in recent years, with annual forest destruction rates in North America and Canada comparable to that of all wildfires combined (Carroll et al., 2004; Bentz et al., 2010; Sims et al., 2010; Man, 2012).

This poses a second spatially explicit risk of bias in estimated real estate capitalization rates if tree mortality itself is assessed with systematic error. Specifically, we examine the home-level accuracy of MPB damage as determined by the U.S. Forest Service's Aerial Detection Survey (ADS) relative to case-by-case photo-interpretation by a remote sensing expert.

* Corresponding author.

http://dx.doi.org/10.1016/j.jfe.2017.02.002
1104-6899/© 2017 Department of Forest Economics, Swedish University of Agricultural Sciences, Umeå. Published by Elsevier GmbH. All rights reserved.

Using a novel econometric approach based on nonparametric matching with a Bayesian regression adjustment we find that both types of measurement error – imprecise geocoding and erroneous damage assessment – can lead to substantial biases in estimated losses of home values. Specifically, we show that both sources of imprecision can introduce sample attrition by entirely omitting actually affected homes from the analysis, and "swapping errors" by falsely classifying an impacted home as damage-free and vice versa, thus sorting them into the wrong bin for our matching model. This leads to damage estimates in terms of lost property values that can deviate from the expert benchmark by a very large margin.

The next section provides details on remote sensing strategies and econometric modeling. This is followed by an empirical section that discusses data and estimation results. Section four concludes.

Methodology

Remote sensing approaches

We consider four different combinations of remote sensing tools to determine the geocoded location of a property, as well as MPB damage in the proximity of a given home. Our flagship approach is based on visual inspection of both home location and surrounding tree health by a remote sensing expert. This provides benchmark estimates for MPB effects on home values that we assume are largely free of spatial measurement errors. The specific steps for this combination of expert photo-interpretation and visual geo-coding (*PI, vis*) are as follows:

1. Use 2011 maps of the National Agriculture Imagery Program (NAIP) with three color bands (red, green, blue) and a spatial resolution of one meter covering our housing market areas of interest.
2. Add preliminary geocodes for each property in our sales data using ESRI's "StreetMap premium for ArcGIS" package (2012, release 2), which is based on home addresses.
3. Use Google Earth (GE), Bing Maps (BM), and GIS imagery provided by local county assessor's offices to visually determine the correct location for a given address. Update geocodes as needed.[1]
4. Using ArcView, create a 100 m buffer around each corrected geocode.[2]
5. Visually inspect tree mortality within this buffer using a combination of all available imagery (NAIP, GE, BM, assessor maps) and declare a home as impacted (that is "treated") if *any* tree damage is detected, irrespective of intensity. We then assign a binary code of 0 (1) to damage-free (impacted) properties. This is essentially the same approach as taken by Backsen and Howell (2013), who use NAIP maps of the same type and resolution as ours to assess MPB damage in the Black Hills of South Dakota in 2010. Using ground-truthing to verify the reliability of this method, they find the photo-interpretation results to be 90–95% accurate using a binary "impacted / not impacted" rule such as ours.[3]

Our second remote sensing approach, labeled "*PI, ABG*" follows steps (1.) and (2.) from above, but skips the geocode verification process in step (3.). Instead, the 100 m buffers are drawn around the generic address-based geocodes provided by StreetMap. Tree mortality is then assessed via visual inspection within that buffer. Thus, this strategy exposes our home sales analysis to the risk of measurement error due to "misplaced properties," as discussed below in more detail.

The third approach, "*ADS, vis*" corrects generic geocodes via expert interpretation as in the benchmark strategy, but uses results from ADS fly-overs to mark a property as impacted (binary code of "1") or not (binary code of "0"). Specifically, a home is declared as affected by tree mortality if its (corrected) 100 meter buffer intersects with a damage polygon from any of the 2007–2010 ADS outings if the sale occurred before June 1, 2011, and any of the 2007–2011 ADS fly-overs if the home sold after May 31, 2011.[4] Since 2007 marks the year in which MPB infestations turned large-scale in the Colorado front range, and each annual fly-over only captures newly diseased trees, this strategy can be expected to cover all relevant cumulative damage surrounding a given property. However, due to the lack of precision of ADS polygons at this refined spatial level as noted in Johnson and Ross (2008), Backsen and Howell (2013) and confirmed by Cohen et al. (2016) this strategy poses the risk of measurement error for our home sales analysis due to erroneous damage assignments.[5]

The fourth and final remote sensing interpretation, labeled "*ADS, ABG*," deviates from the benchmark in both dimensions – geocoding and damage assessment. It combines generic StreetMap geocoding with ADS-informed damage assessment, and is such susceptible to both types of potential measurement errors.

Econometric approach

Our general estimation approach relies on matching techniques, that is a direct comparison of sales price between "treated" homes (with impacted trees within 100 m) and matched "control" homes (without impacted trees within the 100 m buffer). In theory, if a matched control home is identical in relevant observed and unobserved dimensions to a treated home except for tree damage, the difference in price must reflect the damage effect. In practice, controls will rarely be a perfect match for a given treated home. However, residual differences in observables can be controlled for with an auxiliary regression, as shown in Abadie and Imbens (2011) and explained below in more detail. In our case this auxiliary regression also controls for unobservable spatial and seasonal effects via inclusion of corresponding binary indicators.

Overall, this matching approach with regression correction is a robust alternative to a full-fledged hedonic regression model relating sales price to housing attributes and tree damage. Such hedonic

[1] Specifically, the location of a given home was viewed on the NAIP map simultaneously with BM or GE imagery. This minimizes the risk of residual georegistration errors imported via BM or GE, since NAIP has reliable image georegistration.

[2] The 100 m buffer was chosen based on Cohen et al. (2016) who do not detect any significant impacts on home values from MPB-induced tree mortality beyond the 100 m perimeter.

[3] We only consider imagery with known acquisition date. For each home sale, we use the image(s) closest to the sales date to determine tree damage.

[4] This strategy is based on in-person discussions with expert entomologists, according to whom new MPB damage should be clearly visible by June of a given year. Specifically, an attacked tree's needles turn red within the first year of infestation, and gray within three to four years. Either type of discoloration was considered evidence of impacted trees in our imagery-based damage assessment. In contrast, the ADS focuses on new damage (red needles) in a given fly-over year. In our analysis we use cumulative damages from all recent years of fly-overs to implicitly capture trees in all stages of discoloration.

[5] Johnson and Ross (2008), using ground-truthing for 233 sample plots in 2005 find that the ADS accurately characterized the presence of MPB damage in only 61% of cases. They consider these classification errors "to be excessive for use at fine spatial scales" (p. 216). Similarly, Backsen and Howell (2013) assess the accuracy of ADS to detect recent MPB damage at only 25% at the plot level, and at 65% at a larger, 300 foot buffer level. That notwithstanding, Price et al. (2010) use ADS results to determine the number of MPB-affected trees within different perimeters of homes sold between 1995 and 2006 in Grant County, CO, and generally find negative marginal per-tree effects on home values. In contrast, Cohen et al. (2016), for their sample of home sales between 1999 and 2011 in Larimer and Boulder County, CO, were not able to obtain meaningful results using ADS data. Instead, they settle on a repeat-sales analysis using host tree GIS layers to determine MPB-induced real estate effects.

regressions have been found to be highly sensitive to specification errors, as discussed in Kuminoff et al. (2010) and Abbott and Klaiber (2013). Matching avoids these specification risks by using observed explanatory variables primarily to select matched controls, and by balancing the sample of treated and matched controls such that the treatment effect becomes de facto independent of all other observables. As discussed in Ho et al. (2007), this pre-balancing of the data as part of the matching process is vitally important to assure unbiased estimation.

Each remote sensing approach identifies a set of treated observations $i = 1 \ldots n_T$ (with impacted trees within 100 m) and a set of eligible control observations $j = 1 \ldots n_C$.[6] The first step of our econometric approach is to identify a sub-set of these controls that can be matched to the treated while optimizing on covariate balance. To be specific, we seek to select control observations that closely match one or more treated homes in observable dimensions (e.g. age, square footage, lot size, etc.), and as a matched sample *at large* exhibit similar distributions for observables compared to the sample of treated.

We follow Diamond and Sekhon (2013) and use a weighted nearest neighbor criterion to select a balanced sample of controls, that is:

$$d_{ij} = \sqrt{\left(\mathbf{x}_i - \mathbf{x}_j\right)' \mathbf{V}^{-1/2} \mathbf{W} \mathbf{V}^{-1/2} \left(\mathbf{x}_i - \mathbf{x}_j\right)}, \tag{1}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are, respectively, vectors of observable home characteristics for treated observation i and some control observation j, $\mathbf{V}$ is the sample variance-covariance matrix for the entire set of treated and eligible controls and all explanatory variables, and $\mathbf{W}$ is a diagonal weight matrix that assigns relative weights to each covariate in the distance formula. Diamond and Sekhon (2013) developed a genetic search algorithm "GenMatch" that converges towards the optimal $\mathbf{W}^*$, that is the weight matrix that leads to a sub-set of controls that is optimally balanced with the treated.

For example, a researcher may choose the identity matrix as starting matrix for $\mathbf{W}$. She then applies the distance formula in (1) to match each treated home with a pre-specified number (one in our case) of control homes that produce the smallest values for d_{ij}. After all treated have been matched in that fashion, allowing for controls to be matched to more than one treated, a second criterion is used to assess *balance*, that is the similarity in distribution of observables between the *entire sample* of treated and the matched sample of controls. As discussed in Diamond and Sekhon (2013) and Sekhon (2011), this distributional proximity criterion for a given variable could be a normalized difference, basic summary statistics for an underlying quantile-quantile (QQ) plot, or a basic p-value for difference in means, which we employ in our analysis. The GenMatch routine then updates $\mathbf{W}$ in an attempt to improve balance, and repeats the matching process. The entire sequence of steps is iterated until no further improvement is reached for the chosen convergence criterion.

Upon convergence of GenMatch, we obtain the optimal weight matrix $\mathbf{W}^*$, as well as the optimally balanced sample of controls. A generic average treatment effect on the treated (ATT) can then be obtained as

$$\hat{\tau}_{att} = \frac{1}{n_T} \sum_{i=1}^{n_T} \left(y_i - \hat{y}_{i0}\right) \tag{2}$$

where y_i is the actual price of the treated home, and $\hat{y}_{i0}$ is the counterfactual price of i, that is what the home would have sold for in absence of nearby tree damage. In this basic matching version, we have $\hat{y}_{i0} = y_j$, that is the counterfactual price is simply the price of the matched control home.[7]

However, as discussed in Abadie and Imbens (2011), if a given pair i and j is not matched perfectly on observables (say i has slightly higher square footage than j), the resulting ATT will be biased as it captures more than the pure treatment effect. To attenuate this risk of bias, Abadie and Imbens (2011) propose a bias-corrected matching estimator (BCME) that builds on an underlying auxiliary regression model. Specifically, the expression for the counterfactual price of treated home i is now given as

$$\hat{y}_{i0} = y_j + \hat{\mu}\left(\mathbf{x}_i\right) - \hat{\mu}\left(\mathbf{x}_j\right), \tag{3}$$

where $\hat{\mu}(\mathbf{x}_i)$ is the predicted sale price of a home with characteristics $\mathbf{x}_i$ (i.e. the features of the treated home in question) from a linear regression of sales price on housing characteristics using only matched controls.[8] Conversely, $\hat{\mu}(\mathbf{x}_j)$ is the predicted sale price of a home with characteristics $\mathbf{x}_j$ (i.e. the features of the matched control home in question). As shown in Abadie and Imbens (2011), this produces a doubly-robust estimate for $\hat{y}_{i0}$ in the sense that estimation bias goes to zero if either i and j are a perfect match (in which case we have $\mathbf{x}_i = \mathbf{x}_j$ and the regression correction cancels out), or if i and j are an imperfect match ($\mathbf{x}_i \neq \mathbf{x}_j$), but the regression model produces unbiased estimates of predicted prices.[9] The bias-corrected ATT can then again be written as an average of all home-specific estimated price differentials as shown in (2), with the expression in (3) taking the place of $\hat{y}_{i0}$.

Obtaining standard errors for the BCME is somewhat cumbersome in a classical estimation framework. As discussed in Abadie and Imbens (2011), the bootstrapping approach is generally inappropriate. Abadie and Imbens (2006) develop an analytical solution for the variance of the BCME that is asymptotically unbiased. However, with the relatively small samples available for some of the housing markets considered in our empirical application, invoking asymptotic properties is somewhat tenuous.

We therefore opt instead for a Bayesian estimation framework, which does not rely on large-sample results. Specifically, we first balance the sample and find the set of optimally matched controls, as outlined above. We then use Bayesian posterior simulation via Gibbs Sampling to estimate the auxiliary regression model. This produces a large number (say R) draws of predicted price estimates $\hat{\mu}(.)$ in (2), and – in consequence – R draws of the ATT via (3) and (2). The entire set of draws for the ATT characterizes its posterior distribution, which can then be used for inference on central tendency and variability. The detailed estimation steps for this Bayesian econometric approach are given in Appendix A.[10]

Potential effects of measurement errors

Each type of spatial measurement error – wrong geocodes for the location of a home, and/or erroneous tree damage assessment in its proximity – can introduce one or more of the following

[6] Henceforth, we will distinguish between eligible and matched controls. The former are homes within a given market that were not identified as treated by a given remote sensing approach, while the latter are a subset of these eligible controls that was actually matched to the treated and is thus used for final inference.

[7] In the few cases where two or more control homes exhibit the exact same distance metric d_{ij}, the counterfactual price for the treated home is computed as the average over the equidistant matched controls.

[8] As described in Abadie and Imbens (2011), each control observation in this auxiliary regression is weighted by the number of times it was chosen as a matched control. In our case with 1:1 matching this essentially amounts to the number of treated observations for which it was selected as a suitable control.

[9] As noted in Abadie et al. (2002) the BCME is relatively robust to misspecifications of the auxiliary regression function even in a small-sample context as the regression model only needs to adjust for relatively small differences in covariates.

[10] This Bayesian strategy follows, in spirit, that taken by Chib and Greenberg (2010) in the context of propensity score matching.

undesired selection effects for our matching strategy: (i) Actually impacted ("treated") homes can be classified as not impacted and entirely omitted from the analysis if they are not chosen as matched "controls" in the balancing step, (ii) actually treated homes can be classified as not impacted and wrongfully moved to the bin of matched controls, and (iii) actual control homes can be classified as impacted and moved to the bin of treated.

Effects (ii) and (iii), in isolation of (i), would be expected to lead to a downward bias for the ATT (in absolute terms), as less affected and thus more valuable homes are moved to the pool of treated, and/or more affected and thus depreciated homes enter the pool of controls. We find this to be the driving force behind biased results for our "*PI, ABG*" remote sensing approach in our market-specific applications. Effect (i) would, in theory, not bias results if the remaining set of actually treated is still representative of the broader population of affected homes. However, if the omission of treated homes introduces an additional selection effect, bias can result. We find this to be generally the case for our ADS-based approaches, which are likely to select the most heavily impacted sub-set of treated homes. Naturally, this leads to an upward bias (in absolute terms) of the ATT. In some of our market area/remote sensing approach combinations all three effects are at work, which, as expected, leads to mixed results compared to the benchmark ATT depending on the relative strength of the selective omission effect (i) or swapping effects (ii) and (iii).

Empirical application

Data

Given limited budgets for the time-consuming process of expert geocoding and damage interpretation, and the quality of available imagery to accurately assess tree mortality, we settle for 2011 as the time frame for our analysis. To start, we consider all arms-length sales of single-family residential homes in the five front-range counties of Larimer, Boulder, Gilpin, Jefferson, and Clear Creek that are located within 100 m of host trees typically targeted by the MPB. Thus, each home has ex ante a nonzero probability of being impacted by MPB-infested trees in its proximity.[11]

This produces 908 properly coded properties scattered along a 150 mile corridor of the Colorado front range between Fort Collins to the north and Colorado Springs to the south. As is the case with any property valuation analysis, for our matching approach to produce representative results treated and matched control homes must be located in the same housing market, in the sense of being governed by the same consumer preferences and corresponding hedonic price surface (e.g. Kuminoff et al., 2010). This proved to be challenging as many of the considered properties are located in sparsely populated remote communities or along major highway corridors transecting the Rockies. We initially partition the research area into seven spatial sub-regions, loosely based on geographic proximity and comparable home values. Two of these, surrounding the municipalities of Golden and Boulder, respectively, exhibit relatively densely clustered home sales compared to the rest of the sample, and have at least minimally sufficient sample sizes. We will therefore give special attention to these two markets in our analysis. However, for comparison purpose, we also estimate a model using the full data, with spatial fixed effects for the seven sub-regions added to the auxiliary regression. While this full model likely violates the common-market assumption in the sense of spatially overlapping consideration sets of prospective home buyers, it can still provide meaningful guidance on the type of biases introduced by inaccurate geocoding or damage assessment.

Golden is a western suburb in the greater Denver area. Our Golden housing market also comprises the adjacent municipality of Morrison, and the nearby community of Evergreen. Overall, this sub-sample includes 510 correctly coded properties. The City of Boulder lies approximately 25 miles further to the north. Our Boulder sample also comprises homes from the nearby municipalities of Longmont and Lyons, leading to 106 retained observations for this market. While these sample sizes are small to modest, our econometric approach is geared towards the efficient use of smaller samples.

Fig. 1 provides a GIS map of our housing sample, along with an overview of cumulative MPB damage from 2007 through 2011 as assessed by the ADS. Properties located in our "Golden' zone are marked as white triangles, those in the Boulder zone as white circles. The remaining properties, which are included in our full sample analysis, are shown as dark dots on the map.

Descriptive analysis

Table 1 summarizes classification errors *prior to any balancing and matching* for the full sample of 908 observations. As is evident from the table, the benchmark approach identifies 304 impacted properties and 604 homes free of tree mortality that are thus eligible to serve as control observations for our matching analysis. The remote sensing approaches based on generic geocoding ("*PI, ABG*" and "*ADS, ABG*") produce numerous geocoding errors that, on average, "move" a property 73 m (240 feet) from its actual location, with a standard deviation of 113 meeters (371 feet). The smallest errors are quite negligible, but on the upper end of the range the misplacement of properties can be quite pronounced, amounting to over 1800 m (1.12 miles). As has been noted elsewhere (e.g. Zandbergen, 2007), the bulk of these ABG errors stems from positioning homes along the center line of street segments, when they are in fact laterally tucked away from the actual roadway. A second source of positional error is the longitudinal misplacement of homes *along* a given roadway due to the presumed uniform spacing of addresses within a given road segment.

Overall, these geocoding errors alone lead to a misplacement of 21 actually treated homes into the control group, and ten actual controls into the treatment group for a total of 31 allocation errors, as is evident from the "*PI, ABG*" column of Table 1. As can be seen from the last two columns of the table, imprecise tree damage interpretation using ADS data leads to a much higher number of mis-classifications, close to one third of the entire sample. The bulk of these are due to interpreting treated homes as infestation-free (288 and 291 cases, respectively). However, the two ADS-based methods also produce a sizable share of properties that are misinterpreted as treated (65 and 63, respectively). These error patterns largely persist for the sub-markets of Golden and Boulder, as discussed below.

Table 2 shows basic sample statistics for the full sample and the two sub-areas. As can be seen from the table, for the sample at large control homes are of comparable age and size relative to treated properties, but fetch a higher average price. Furthermore, a relatively larger share of eligible control homes have experienced a recent wildfire in their proximity.[12] Treated and control homes sold at comparable rates across annual quarters, as is evident from the second set of rows under the FULL SAMPLE heading.

These sample patterns generally also hold for the two sub-areas, though control homes in Golden exhibit a smaller average lot size

[11] The few instances of flip-sales, that is homes that sold multiple times in the same year, were not included in our analysis as they were most likely purchased as investment properties and not to serve as a primary family residence.

[12] We include this five-year, five-kilometer wildfire indicator in our auxiliary regression as it has found to affect home values in Cohen et al. (2016).

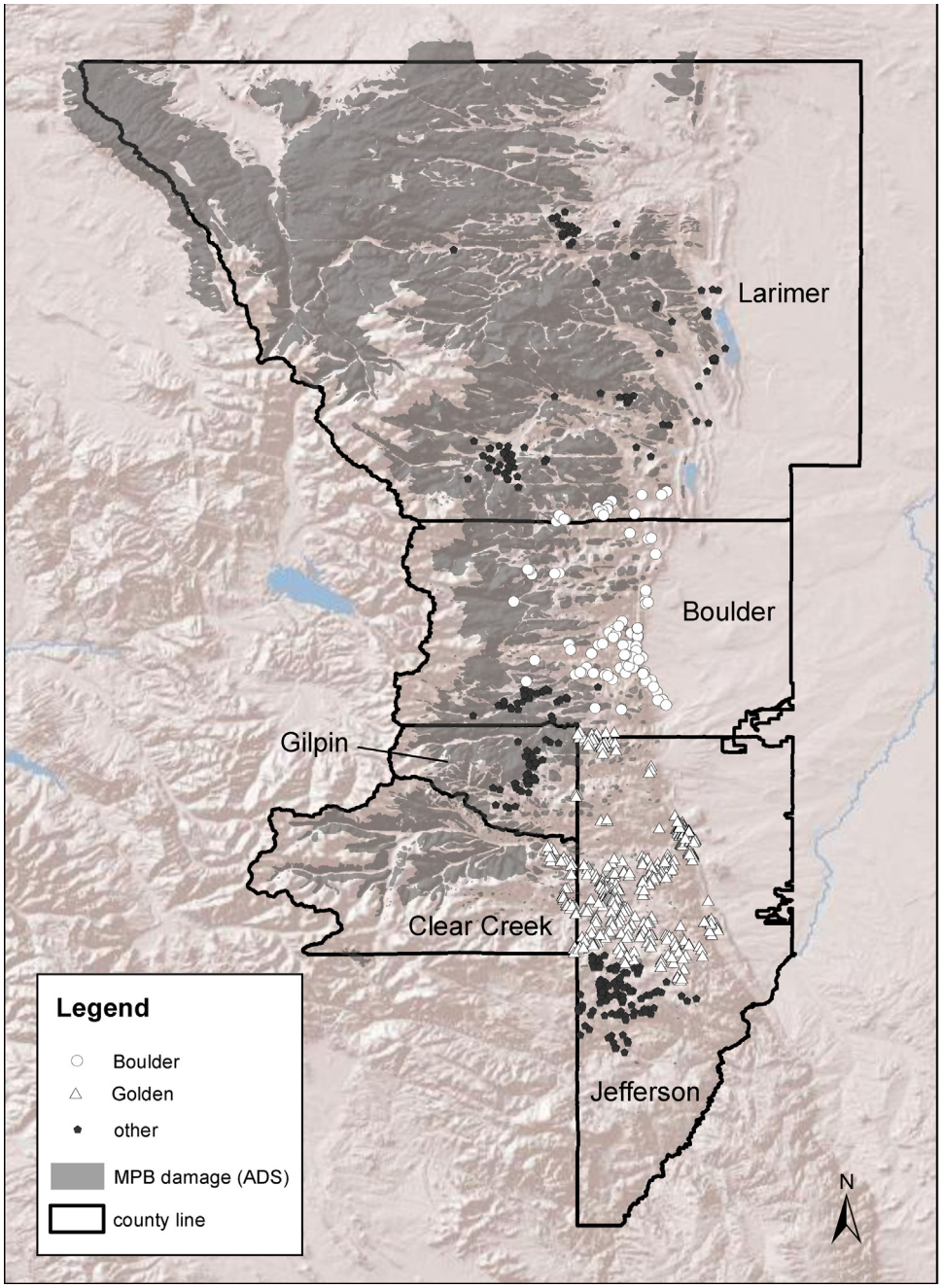

Fig. 1. Overview of research area.

Table 1
Remote sensing errors.

	PI, vis.	PI, GIS	ADS, vis.	ADS, GIS
Treated	304	293	146	139
Eligible controls	604	615	762	769
Actual control flagged as treated	–	10	65	63
Actual treated flagged as control	–	21	223	228
Total classification errors	–	31	288	291
Error distances (m)				
Mean	–	73	–	73
std	–	113	–	113
min	–	0	–	0
max	–	1815	–	1815

PI: damage assessment via expert photo-interpretation; ADS: damage assessment via Aerial Detection Survey; ABG: address-based geocoding using generic GIS software; vis.: expert geocoding using photo-imagery.

Table 2
Sample statistics.

Variable	Treated observations				Eligible control observations			
	Mean / pct.	std	min	max	mean	std	min	max
	Full sample (treated = 304, eligible controls = 604)							
Price (000s)	$355	$191	$77	$1,300	$460	$364	$55	$4,313
Age	30.67	16.61	0	92	30.87	23.53	0	151
sqft (00s)	18.80	8.46	3.68	66.24	21.29	9.57	4.71	71.47
Lot sqft. (000s)	114.96	163.19	0.87	1590.30	101.68	259.06	0.00	2132.70
Bedrooms	3.02	0.94	0	7	3.24	0.96	1	9
Bathrooms	2.37	0.97	0	6	2.75	1.18	0	9
Near fire	11.84%				23.84%			
Sold 1st quarter	21.71%				25.00%			
Sold 2nd quarter	31.25%				25.17%			
Sold 3rd quarter	29.93%				34.11%			
Sold 4th quarter	8.22%				8.61%			
	Golden (treated = 161, eligible controls = 349)							
Price (000s)	$413	$208	$78	$1,300	$469	$255	$81	$2,200
Age	30.92	16.61	3	89	28.93	21.56	0	136
sqft (00s)	20.77	9.11	5.76	66.24	23.10	9.56	4.71	71.47
Lot sqft. (000s)	100.53	154.02	0.87	1590.30	66.24	188.45	0.00	2132.70
Bedrooms	3.27	0.93	1	7	3.42	0.96	1	9
Bathrooms	2.68	0.92	1	6	3.04	1.16	1	9
Near fire	6.83%				22.64%			
Sold 1st quarter	19.88%				23.50%			
Sold 2nd quarter	32.92%				27.22%			
Sold 3rd quarter	27.95%				34.38%			
Sold 4th quarter	8.07%				8.60%			
	Boulder (treated = 20, eligible controls = 86)							
Price (000s)	$315	$108	$159	$589	$737	$705	$108	$4,313
Age	40.30	15.19	9	70	38.13	25.85	3	131
sqft (00s)	15.55	5.22	7.38	24.57	21.11	11.61	5.86	61.27
Lot sqft. (000s)	154.82	232.83	17.09	916.50	158.97	393.20	0.71	1742.40
Bedrooms	3.05	1.05	1	5	3.21	0.93	1	5
Bathrooms	2.00	0.86	1	3	2.77	1.22	1	7
Near fire	45.00%				66.28%			
Sold 1st quarter	30.00%				23.26%			
Sold 2nd quarter	15.00%				27.91%			
Sold 3rd quarter	40.00%				34.88%			
Sold 4th quarter	15.00%				9.30%			

std: standard deviation; min (max): minimum (maximum); pct.: percent.

than treated properties, and control homes in Boulder have higher average square footage than treated residences. We use the basic housing attributes captured in the table in our balanced matching procedure, and add to these the fire and sales quarter indicators in the auxiliary regression model.

Table 3 shows sample counts before and after matching, as well as *post-matching* counts for the three types of errors that can occur using one of the flawed interpretation methods. The first block of rows refers to the full sample. The first two rows, which give the number of identified treated and eligible control properties, are repeated from Table 1. The third row shows the remaining control observations after matching, that is the set of controls that actually feed into the econometric analysis and the estimated treatment effect. In all cases the number of these matched controls is smaller than the number of treated. This is as expected, since the balancing and matching algorithms allow for replacement in the selection of controls, following best practices in the matching literature (e.g. Ho et al., 2007; Abadie and Imbens, 2011).

Rows four through six provide a detailed look at omission and swapping errors. As is evident from the table, all three non-benchmark approaches introduce all three types of errors, although error counts are substantially smaller for the "*PI, ABG*" version compared to the ADS-based approaches. Specifically, the latter drop over 50% of actually treated from the analysis. As discussed above this not only hurts efficiency via reduced sample sizes, but can introduce selection bias if this attrition is non-random. Post-matching swapping errors are also considerable for the ADS-based strategies, with approximately 17% of actually treated moved to the bin of matched controls, and approximately 10% of controls interpreted as treated.

Analogous results for Golden are captured in the second block of rows. As can be seen from the table, the benchmark remote sensing approach identifies 161 treated observations, which leaves 349 eligible control homes. After balanced matching 113 of these are actually selected as matched controls, with some serving as matches for multiple treated. In comparison, the "*PI, ABG*" interpretation completely omits five treated homes from the analysis, erroneously sorts three treated homes into the bin of matched controls, and misinterprets two control homes as treated. As for the sample at large, errors introduced by the ADS-based methods are more staggering, leading to the omission of close to 80% of actually treated in both cases. In addition, several treated homes (12 and 15, respectively) are misplaced into the matched control bin, and a comparable number of control cases (11 and 10, respectively) are falsely used as treated properties.

Results for Boulder are given in the bottom block of rows of Table 3. The benchmark interpretation produces 20 treated homes, 86 eligible controls, and 15 matched controls. Working with flawed geocodes in the "*PI, ABG*" interpretation drops four treated altogether, and misplaces another treated and three controls. While these figures may seem diminutive, they still constitute a sizable share of all available observations for this small sample. As is the case for the full sample and Golden, the ADS-based approaches drop even more treated (eight in both cases) and induce some erroneous swapping of actually treated (one in each case) and actual controls (10 in each case).

Table 3
Sorting effects based on remote sensing approaches.

	PI, vis.	PI, ABG	ADS, vis.	ADS, ABG
	Full sample			
Treated	304	293	146	139
Eligible controls	604	615	762	769
Matched controls	206	200	120	114
Unused actually treated	–	10	173	183
Actually treated used as matched controls	–	11	50	45
Actual controls used as treated	–	10	65	63
	Golden			
Treated	161	155	32	28
Eligible controls	349	355	478	482
Matched controls	113	114	30	28
Unused actually treated	–	5	128	128
Actually treated used as matched controls	–	3	12	15
Actual controls used as treated	–	2	11	10
	Boulder			
Treated	20	18	21	21
Eligible controls	86	88	85	85
Matched controls	15	15	18	18
Unused actually treated	–	4	8	8
Actually treated used as matched controls	–	1	1	1
Actual controls used as treated	–	3	10	10

PI: damage assessment via expert photo-interpretation; ADS: damage assessment via Aerial Detection Survey; ABG: address-based geocoding using generic GIS software; vis.: expert geocoding using photo-imagery.

As discussed previously, swapping effects, by themselves, should unambiguously bias the ATT towards zero. In theory, the omission effect, by itself, can be bias-neutral if omissions are completely random. However, the bias can go towards either direction if omitted treated homes differ from the sample at large in relevant dimensions. In our case, such a dimension could be the *intensity* of infestation, that is the actual count or proportion of affected trees surrounding a property. For example, ADS fly-overs may focus primarily on the most heavily affected areas and miss low-intensity impacts. This would bias the ATT away from zero and lead to an over-estimation of losses. Similarly, if homes are "moved" from their woody surroundings to the center of a highway by ABG, it is feasible that the erroneous new location will "pick up" some tree damage only in the most heavily affected areas. Our expert assessment did not capture actual counts of diseased trees due to budget and time limitations, which preempts a rigorous quantitative verification of intensity-based selection.

Econometric analysis

As mentioned above, we match homes on age (in years since construction), square footage (in units of 100 square feet), lot size (in 1000s of square feet), number of bedrooms, and number of bathrooms. To check on balance within the GenMatch algorithm, we also include all possible squared and interaction terms for these variables, as suggested in Diamond and Sekhon (2013). Example balancing results for the benchmark case are given in Appendix B. The original set of matching variables are also included in the auxiliary regression that underlies the regression-adjusted ATT, with the addition of binary indicators for the occurrence of a wildfire within five kilometers and the past five years, as well as the annual quarter in which a give sale occurred. The full-sample regression also contains sub-regional fixed effects.

As described above, the auxiliary regression is estimated in a Bayesian framework, with diffuse normal priors for all coefficients (mean = 0, variance = 10^9), and a flat inverse-gamma prior for the variance (shape = scale = 1/2). We let our Gibbs Sampler burn in for 5000 draws to reduce the effect of starting values, and retain 10,000 additional draws of parameters for inference. Specifically, we use each draw of regression coefficients to generate an estimate of predicted prices for a given treated home and its matched control, that is the terms $\hat{\mu}(\mathbf{x}_i)$ and $\hat{\mu}(\mathbf{x}_j)$ in Eq. (3). These predicted prices then feed into an estimate of the overall ATT via Eq. (2). Ultimately, therefore, our Bayesian model de facto produces 10,000 draws to characterize the distribution of the sought ATT.

Estimation results are captured in Table 4. As before, the table exhibits three blocks of rows for the full sample, Golden, and Boulder, respectively. The columns show the posterior mean, standard deviation, proportion of posterior distribution to the right of zero ($p > 0$), and the lower and upper bound of the 95% highest posterior density interval (HPDI). The $p > 0$ statistic provides an at-a-glance assessment if the ATT is unambiguously negative ($p > 0$ is close to zero), positive ($p > 0$ is close to one), or indeterminate in sign ($p > 0$ is close to 0.5). Thus, it can be used as an informal indication of significance, in slight abuse of the classical definition. The HPDI, in turn, gives the narrowest possible set of bounds that contain 95% of the posterior distribution. While conceptually and computationally different, it can be interpreted akin to a classical confidence interval in the sense that there is a 95% chance that the (presumable fixed but unknown) ATT lies between these bounds. The last column in the table gives percentage deviation of estimates flowing from the flawed remote sensing approaches relative to those generated by the benchmark.

As can be seen from the table, all three *benchmark* ATTs are clearly negative, though the HPDI does include zero in the case of Boulder, which has the smallest sample size. Generally, though, this supports existing research that finds that MPB-induced tree damage has a detrimental effect on home values (Price et al., 2010; Cohen et al., 2016). For the sample at large, the average loss is $10,471 per home, which amounts to approximately 2.5% of the sample mean, counting all treated and all eligible controls. In contrast, average losses in Golden amount to over $23,000, or 5.1% of the sample mean. Boulder posts the highest losses with a posterior mean of over $124,000, or close to 19% of the sample mean.

Turning to the "*PI, ABG*" remote sensing version, the swapping effect dominates any potential selection effects through omission for Golden and Boulder, and thus introduces a pronounced positive deviation from the benchmark for the posterior mean for both areas (108% and 35%, respectively). The opposite holds for the full sample, where the "*PI, ABG*" approach produces a posterior mean that shows a negative deviation of −79% from the benchmark. Thus, in this case

Table 4
Estimated losses in home values ($000's).

Method	Mean	std	$p>0$	Low	Up	Diff. bench. (%)
	Full sample					
PI, visual	−10.471	(2.926)	0.000	−16.517	−5.110	–
PI, ABG	−18.746	(3.192)	0.000	−24.828	−12.387	−79.03%
ADS, visual	−23.835	(2.299)	0.000	−28.312	−19.376	−127.63%
ADS, ABG	15.506	(55.595)	0.607	−91.265	123.971	248.09%
	Golden					
PI, visual	−23.097	(5.871)	0.000	−34.791	−11.562	–
PI, ABG	1.912	(6.252)	0.621	−10.563	13.898	108.28%
ADS, visual	−14.430	(5.130)	0.004	−24.869	−4.573	37.52%
ADS, ABG	−91.314	(112.915)	0.204	−311.240	136.497	−295.35%
	Boulder					
PI, visual	−124.338	(81.445)	0.058	−285.075	40.801	–
PI, ABG	−81.503	(40.797)	0.024	−164.931	−0.792	34.45%
ADS, visual	−173.582	(26.180)	0.000	−227.700	−124.026	−39.60%
ADS, ABG	−159.567	(22.131)	0.000	−201.598	−114.280	−28.33%

PI: damage assessment via expert photo-interpretation; ADS: damage assessment via Aerial Detection Survey; ABG: address-based geocoding using generic GIS software; vis.: expert geocoding photo-imagery; $p>0$: proportion of posterior distribution greater than zero; low (up): lower (upper) bound of 95% highest posterior density interval; diff. bench: % difference relative to benchmark (PI, visual).

intensity effects via involuntary selection of retained treated appear to outweigh swapping effects.

The "*ADS,visual*" approach leads to negative deviations from the benchmark for the full sample (−128%) and Boulder (−40%), and a positive deviation for Golden (38%). The "*ADS,ABG*" version, in turn, yields a negative deviation for Boulder (−28%) and Golden (−295%, although with 20% of posterior draws in the positive domain), and a positive deviation for the full sample (248%, though 40% of the posterior is located in the negative domain). Thus, while some of these estimates exhibit less-than-desirable posterior precision, we obtain a mixed picture for these ADS-based effects, with selection effects potentially dominating over swapping effects.

Our benchmark results for MPB-induced losses in home values are within the ballpark range of those found by Cohen et al. (2016) for Boulder county and its northern neighbor, Larimer county. Specifically, Cohen et al. (2016) estimate MPB-induced losses of 12.5% for the typical Boulder county home located near host trees (damaged or not) compared to a home *outside* of the host tree zone for the 2011 calendar year. In contrast, our analysis de facto compares homes with *damaged* host trees to homes with *healthy* host trees. To the extent that homes in a healthy forest surrounding are typically more highly priced than homes outside the host tree zone in that area, our larger loss estimate of 19% for Boulder appears reasonable. It also comes close to Cohen et al. (2016)'s estimate of a 22% erosion of home values in 2011 for the typical host-tree-zone property in adjacent Larimer county.

Conclusion

We take a closer look at the possible estimation biases than can occur in a property valuation study with focus on localized forest health due to imprecise, address-based geocoding and overly coarse tree damage assessment via the Forest Service's Aerial Detection Survey.

For our sample of over 900 single family residences in the Colorado front range, we find that both geocoding and damage assessment errors can introduce substantial biases for residential capitalization rates. In theory, either type and combination of spatial mis-information can introduce selection bias by retaining only a subset of treated that is no longer representative of the broader population, as well as swapping bias by sorting homes into the wrong bin in an econometric matching framework. In our application, we find that ADS-based approaches to damage assessment are susceptible to both types of errors, leading to substantial over- or under-estimation of lost values.

While our empirical application is germane to forest health, the importance of accurate geocoding likely carries over to other property valuation contexts that rely on localized distances to environmental (dis-)amenities or hot spots, such as polluted river segments (Cho et al., 2011), hazardous waste sites (Guignet, 2013; Mastromonaco, 2014), noise sources (Pope, 2008), urban recreation sites (Phaneuf et al., 2008), or air and water quality monitoring stations (Leggett and Bockstael, 2000; Poor et al., 2007; Neill et al., 2007; Anselin and Lozano-Gracia, 2008). It may be prudent in such cases to obtain a small sample of expertly geocoded properties and decide if ABG-related position errors are small and/or random enough to be ignorable for a primary valuation analysis.

A natural extension of our work would be a more careful, imagery-based measurement of the *intensity* of infestation surrounding a given property, in terms of the number or proportion of diseased trees within a close buffer. This would allow for a more rigorous examination and quantification of our presumed selection effect towards more heavily impacted areas associated with the ADS efforts. It would also pave ways towards investigating if the relationship between infested trees and home values may be nonlinear or exhibit threshold effects. Another interesting extension might be the outsourcing of image interpretation for dead trees to online labor markets, as recently implemented by Yu et al. (2015) in the context of cloud detection in Landsat scenes.

Acknowledgements

This work was supported by the United States Forest Service, Southern Research Station [cooperative agreement number 12-JV-11330143-106].

Appendix A. Econometric details

The GenMatch balancing algorithm returns the original sample of treated, $i=1\ldots n_T$, and a matched sample of unique controls, $j=1\ldots n_M$, with possible replications, such that $n_M \le n_T$. Each matched control receives a weight K_j that indicates for how many treated it serves as a match. For most matched observation, K_j will be unity.

Next, we build the auxiliary regression model, using only the matched controls. The models follows a basic weighted OLS specification, that is:

$$\tilde{y}_j = \tilde{\mathbf{x}}_j'\boldsymbol{\beta} + \epsilon_j, \quad \text{with}$$
$$\tilde{y}_j = \sqrt{K_j}y_j, \quad \tilde{\mathbf{x}}_j = \sqrt{K_j}\mathbf{x}_j \quad \text{and} \tag{A.1}$$
$$\epsilon_j \sim n\left(0, \sigma^2\right)$$

where y_j is the observed sales price of home j, $\mathbf{x}_j$ is a vector of home characteristics, wildfire indicator, and quarter-of-sales indicators as described in the main text. The error term ϵ_j follows the usual normal distribution with zero mean and common variance σ^2.

For the entire sample of matched controls, the likelihood function is thus given as:

$$p\left(\tilde{\mathbf{y}}|\boldsymbol{\beta}, \sigma^2, \tilde{\mathbf{X}}\right) = (2\pi)^{-n_M/2}\left(\sigma^2\right)^{-n_M/2} \exp\left(-\frac{1}{2\sigma^2}\left(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)'\left(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)\right), \tag{A.2}$$

where $\tilde{\mathbf{y}}$ denotes the entire vector of weighted home prices for the matched controls, and $\tilde{\mathbf{X}}$ is the weighted sample matrix of explanatory variables.

Next, we employ the standard normal prior for $\boldsymbol{\beta}$ and inverse-gamma prior for σ^2, i.e.:

$$\boldsymbol{\beta} \sim n\left(\boldsymbol{\mu}_0, \mathbf{V}_0\right), \quad \text{or}$$
$$p\left(\boldsymbol{\beta}\right) = (2\pi)^{-k/2}|\mathbf{V}_0|^{-1/2} \exp\left(-\frac{1}{2}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_0\right)'(\mathbf{V}_0)^{-1}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_0\right)\right)$$
$$\sigma^2 \sim ig\left(\nu_0, \tau_0\right), \quad \text{or}$$
$$p\left(\sigma^2\right) = \frac{\tau_0^{\nu_0}}{\Gamma\left(\nu_0\right)}\left(\sigma^2\right)^{-(\nu_0+1)} \exp\left(-\frac{\tau_0}{\sigma^2}\right), \quad \text{with}$$
$$E\left(\sigma^2\right) = \frac{\tau_0}{\nu_0 - 1}, \quad V\left(\sigma^2\right) = \frac{\tau_0^2}{(\nu_0 - 1)^2(\nu_0 - 2)}, \tag{A.3}$$

where k is the dimension of $\boldsymbol{\beta}$. As is evident from (A.3) we parameterize the inverse-gamma with shape ν_0 and scale τ_0, following Gelman et al. (2004).

Combining the likelihood with these priors and dropping all nuisance terms yields the posterior kernel

$$p\left(\boldsymbol{\beta}, \sigma^2|\tilde{\mathbf{y}}, \tilde{\mathbf{X}}\right) \propto \left(\sigma^2\right)^{\frac{-n_M - 2\nu_0 - 2}{2}} \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2}\left(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)'\left(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)\right.\right.$$
$$\left.\left. + \left(\boldsymbol{\beta} - \boldsymbol{\mu}_0\right)'(\mathbf{V}_0)^{-1}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_0\right)\right)\right) \tag{A.4}$$

Following standard procedures, we break this posterior kernel into conditional densities for $\boldsymbol{\beta}$ and σ^2, given as:

$$p\left(\boldsymbol{\beta}, |\sigma^2, \mathbf{y}, \tilde{\mathbf{X}}\right) \sim n\left(\boldsymbol{\mu}_1, \mathbf{V}_1\right), \quad \text{with}$$
$$\mathbf{V}_1 = \left(\mathbf{V}_0^{-1} + \frac{1}{\sigma^2}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}, \quad \text{and} \tag{A.5}$$
$$\boldsymbol{\mu}_1 = \mathbf{V}_1\left(\mathbf{V}_0^{-1}\boldsymbol{\mu}_0 + \frac{1}{\sigma^2}\tilde{\mathbf{X}}'\tilde{\mathbf{y}}\right)$$

and

$$\sigma^2|\boldsymbol{\beta}, \tilde{\mathbf{y}}, \tilde{\mathbf{X}} \sim ig\left(\nu_1, \tau_1\right), \quad \text{with}$$
$$\nu_1 = \frac{2\nu_0 + n}{2}, \quad \text{and} \tag{A.6}$$
$$\tau_1 = \tau_0 + \frac{1}{2}\left(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)'\left(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)$$

Our Gibbs Sampler then draws sequentially and repeatedly from these two conditionals. as discussed inter alia in Koop (2003) and Koop et al. (2007), after a sufficient number of iterations the sampler will converge to the desired joint posterior given in (A.4).

As described in the main text, for each retained draw of $\boldsymbol{\beta}_r$ from the rth round of the sampler, we compute the predictive constructs $\hat{\mu_r}(\tilde{\mathbf{x}}_i) = \tilde{\mathbf{x}}_i'\boldsymbol{\beta}_r$ and $\hat{\mu_r}(\tilde{\mathbf{x}}_j) = \tilde{\mathbf{x}}_j'\boldsymbol{\beta}_r$. These then feed into the rth draw of the ATT via Eqs. (2) and (3) of the main text.

Table B.1
Balancing results for benchmark case.

Variable	Mean Before	Mean After	max Before	max After
	Full sample			
Age	6.07	0.94	59.00	5.00
sqft00	2.56	0.91	6.15	8.96
lot000	58.86	20.63	542.40	152.10
	Golden			
Age	5.79	1.68	47.00	5.00
sqft00	2.68	1.01	6.15	5.23
lot000	67.40	19.10	542.40	114.75
	Boulder			
Age	11.43	6.15	61.00	17.00
sqft00	6.60	0.90	36.70	2.38
lot000	120.32	54.12	825.90	387.68

mean (max) = mean (maximum) quantile deviation (5% steps); sqft00 = square footage in units of 100; lot000 = lot size in units of 1000 square feet.

Appendix B. Balancing results

To check on improvements in distributional similarity between treated and controls before and after balancing the sample, we compare the difference in 5%-incremented quantiles between treated and original and matched controls, respectively. The table below shows these diagnostics for the benchmark remote sensing approaches for each sample used in our main analysis, and for the three continuous variables used in matching. Clearly, both mean and maximum quantile deviations decrease markedly after balancing for all variables and applications.

The only exception is a slight increase in the maximum deviation of square footage (from 6.15 to 8.25) for the full sample. However, this comes at the gain of dramatic reductions in maximum deviations for age and lot size.

Table B.1

References

Abadie, A., Imbens, G., 2006. Large-sample properties of matching estimators for average treatment effects. Econometrica 74, 235–267.

Abadie, A., Imbens, G., 2011. Bias-corrected matching estimators for average treatment effects. J. Bus. Econ. Stat. 29, 1–11.

Abadie, A., Drucker, D., Herr, J., Imbens, G., 2002 October. Simple and bias-corrected matching estimators for average treatment effects, nBER Technical Working paper 283.

Abbott, J., Klaiber, H., 2013. The value of water as an urban club good: a matching approach to community-provided lakes. J. Environ. Econ. Manage. 65, 208–224.

Anselin, L., Lozano-Gracia, N., 2008. Errors in variables and spatial effects in hedonic house price models of ambient air quality. Empir. Econ. 34, 5–34.

Backsen, J., Howell, B., 2013. Comparing aerial detection and photo interpretation for conducting forest health surveys. West. J. Appl. For. 28, 3–8.

Bentz, B., Jacques, R., Fettig, C., Hansen, E., Hayes, J., Hicke, J., Kelsey, R., Negrn, J., Seybold, S., 2010. Climate change and bark beetles of the Western United States and Canada: direct and indirect effects. BioScience 60, 602–613.

Burra, T., Jerrett, M., B.R.T., Anderson, M., 2002. Conceptual and practical issues in the detection of local disease clusters: a study of mortality in Hamilton, Ontario. Can. Geogr. 46, 160–171.

Carroll, A., Taylor, S., Régnière, J., Safranyik, L., 2004. Effects of climate change on range expansion by the mountain pine beetle in British Columbia. In: Shore, T., Brooks, J., Stone, J. (Eds.), Mountain Pine Beetle Symposium: Challenges and Solutions. October 30–31, 2003, Kelowna, British Columbia, Natural Resources Canada, Canadian Forest Service, Pacific Forestry Center, Information Report BC-X-399, pp. 223–232.

Cayo, M., Talbot, T., 2003. Positional error in automated geocoding of residential addresses. Int. J. Health Geogr. 2, 1–12.

Chib, S., Greenberg, E., 2010. Bayesian matching for causal inference, paper presented at the Yale Econometrics Research Seminar Series, April 7., pp. 2010.
Cho, S.-H., Roberts, R., Kim, S., 2011. Negative externalities on property values resulting from water impairment: the case of the Pigeon River watershed. Ecol. Econ. 70, 2390–2399.
Cohen, J., Blinn, C., Boyle, K., Holmes, T., Moeltner, K., 2016. Hedonic valuation with translating commodities: Mountain P ine Beetles and host trees in the Colorado Front Range. Environ. Resour. Econ. 63, 613–642.
Diamond, A., Sekhon, J., 2013. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. Rev. Econ. Stat. 95, 932–945.
Gelman, A., Carlin, J., Stern, H., Rubin, D., 2004. Bayesian Data Analysis, 2nd ed. Chapman & Hall/CRC.
Guignet, D., 2013. What do property values really tell us? A hedonic study of underground storage tanks. Land Econ. 89, 211–226.
Harada, Y., Shimada, T., 2006. Estimating the impact of the precision of address geocoding on estimated density of crime locations. Comput. Geosci. 32, 1096–1107.
Hay, G., Kypri, K., Whigham, P., Langley, J., 2009. Potential bias due to geocoding error in spatial analyses of official data. Health Place 15, 562–567.
Ho, D., Kosuke, I., King, G., Stuart, E., 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit. Anal. 15, 199–236.
Johnson, E., Ross, J., 2008. Quantifying error in aerial survey data. Aust. For. 71, 216–222.
Koop, G., Poirier, D., Tobias, J., 2007. Bayesian Econometric Methods. Cambridge University Press.
Koop, G., 2003. Bayesian Econometrics. Wiley.
Kuminoff, N., Parmeter, C., Pope, J., 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? J. Environ. Econ. Manage. 60, 145–160.
Leggett, C., Bockstael, N., 2000. Evidence of the effect of water quality on residential land prices. J. Environ. Econ. Manage. 39, 121–144.
Man, G., 2012. Major forest insect and disease conditions in the United States: 2011, Tech. Rep. FS-1000. United States Department of Agriculture, Forest Service.
Mastromonaco, R., 2014. Hazardous waste hits Hollywood: superfund and housing prices in Los Angeles. Environ. Resour. Econ. 59, 207–230.
Neill, H., Hassenzahl, D., Assane, D., 2007. Estimating the effect of air quality: spatial versus traditional hedonic models. South. Econ. J. 73, 1088–1111.
Phaneuf, D.J., Smith, V., Palmquist, R., Pope, J., 2008. Integrating property value and local recreation models to value ecosystem services in urban watersheds. Land Econ. 84 (3), 361–381.
Poor, P., Pessagno, K., Paul, R., 2007. Exploring the hedonic value of ambient water quality: a local watershed-based study. Ecol. Econ. 60, 797–806.
Pope, J., 2008. Buyer information and the hedonic: the impact of a seller disclosure on the implicit price for airport noise. J. Urban Econ. 63, 498–516.
Price, J., McCollum, D., Berrens, R., 2010. Insect infestation and residential property values: a hedonic analysis of the mountain pine beetle epidemic. For. Policy Econ. 12, 415–422.
Schootman, M., Sterling, D., Struthers, Y., Laboube, T., Emo, B., Higgs, G., 2007. Positional accuracy and geographic bias of four methods of geocoding in epidemiologic research. Ann. Epidemiol. 17, 464–470.
Sekhon, J., 2011. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. J. Stat. Softw. 42, 1–52.
Sims, C., Aadland, D., Finnoff, D., 2010. A dynamic bioeconomic analysis of mountain pine beetle epidemics. J. Econ. Dyn. Control 34, 2407–2419.
Yu, L., Ball, S., Blinn, C., Moeltner, K., Peery, S., Thomas, V., Wynne, R., Cloudsourcing, 2015. Using an online labor force to detect clouds and cloud shadows in landsat images. Remote Sens. 3, 2334–2351.
Zandbergen, P., 2007. Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. BMC Public Health 7, 1–13.