

Measuring the Value of U.S. National Parks using Hedonic Property Value Models

Jeffrey Zabel, Professor
Department of Economics
Tufts University, Medford MA
jeff.zabel@tufts.edu

Christoph Nolte, Assistant Professor
Department of Earth & Environment
Boston University
chnolte@bu.edu

Robert Paterson, Principal
Industrial Economics, Inc
Cambridge MA
rwp@indecon.com

Abstract

The National Park Service (NPS) is expected to calculate the economic benefits of national parks for use when evaluating policy and management options of the parks (Loomis 2002). We estimate the marginal willingness to pay (MWTP) to live close to a national park that is capitalized into house prices for 18 national parks. We find that estimated effects can be significant and are heterogeneous across different types of parks. However, these effects can also be attributed to other factors in the housing market which brings the significance of these estimated impacts into question.

1. Introduction

America's national parks are a significant national treasure. There are 424 national parks that cover approximately 85 million acres. The U.S. National Park Service (NPS) recognizes many different protected area designations of which National Park is only one (there are 63 National Parks in the U.S.).¹ Recreational visitation rates peaked at 331 million in 2016 as the number of national parks continues to rise on a yearly basis (see Figure A1 in the Data Appendix). Yet the NPS's total (real) budget has been stagnant for many years despite this increased visitation rate and this has resulted in a significant maintenance backlog.^{2,3} Banzhaf (2022) and papers referenced therein provide detailed analyses of different means for raising revenues to finance the use of public lands.

NPS has a specific mandate to preserve the ecological and historical integrity of national parks and to make them available and accessible for public use.⁴ The NPS and other federal resource management agencies are expected to estimate the economic benefits of these parks for use when evaluating policy and management options of the parks (Loomis 2002). There are numerous ways to do this. Contingent valuation (CV), travel cost, and hedonics are methods that economists have used to value national parks specifically, and natural resources more generally.

These techniques are relevant to valuing different aspects of parks. CV is useful for estimating the non-use or existence value of parks. For example, Haefele et al. (2020) surveyed a sample of U.S. households and estimated that Americans place a value of \$92 billion on all U.S. national parks; \$62 billion is for national park lands, waters, and historic sites and \$30 billion is for NPS Programs such as recreational activities, educational programs, and protection of natural landmarks and historic sites. Travel cost methods are used to capture the use value of the parks. Data on individuals' park visitation and associated travel costs are used to estimate the value of a

park trip. For example, Parsons et al. (2020) construct models of visitation to a suite of seven national parks in the four corners state. Depending on model specification they estimate trip values ranging from approximately \$30 to over \$300 for Grand Canyon National Park.

The hedonic method is used to estimate the marginal willingness to pay (MWTP) to live close to a national park that is capitalized into house prices. This captures the utility from having a national park in one's neighborhood (versus living farther away from the park). Mueller et al. (2021) is the first study to use the hedonic method to value the MWTP to live near a U.S. National Park. They focus on two parks near Tucson, Arizona: Saguaro National Park West and Tucson Mountain Park. They find that homeowners are willing to pay, on average, \$124,000 to live within 1 km of Saguaro National Park West, and \$160,000 to live within 1 km of Tucson Mountain Park (\$2020).

Given the variation in national park settings and characteristics, it is hard to generalize from individual case studies the MWTP to live near all other national parks. In this study, we provide estimates of the MWTP to live near a national park for a variety of settings. One common roadblock to doing this is getting the transaction data in the proximity of multiple parks. We are able to overcome this roadblock by using the ZTRAX dataset, Zillow's U.S. transaction and tax assessment data for 1987-2021 that is made available to researchers. We use the ZTRAX data to estimate the MWTP using the hedonic model for numerous U.S. national parks. This is useful since it applies the same framework and data to get MWTP estimates for many national parks that are directly comparable. It is also particularly interesting since national parks exist in very different settings. Some are grand settings such as those for Yosemite and Yellowstone National Parks, whereas others are small units in urban settings such as the John F. Kennedy and Longfellow National Historical Sites in Boston. While we use a common approach to estimate the

MWTP to live near national parks we do not expect that the results will be similar across all these parks. We see it as a real advantage of our study that we estimate impacts across such a varied set of NPS units. This includes three National Parks, six National Historic Sites, five National Historical Parks, a National Battlefield Park, a National Forest, and two National Monuments. They are in nine states: Alabama, Arizona, Massachusetts, Michigan, North Carolina, Nevada, New Jersey, Ohio, and Tennessee. Furthermore, five of these national parks were either established or redesignated as national parks during the data covered by ZTRAX.

Why would households be willing to pay more to live close to a national park? Reasons include easier access to the park's amenities, a better view and/or a nicer neighborhood due to the natural amenities the park provides, and a guarantee that the open space will not be developed. But living close to a national park could also have negative consequences such as traffic and associated air pollution and the general inconvenience of congestion in the neighborhood. Hence, determining the net impact requires an empirical analysis.

Walls (2022) points out two important issues that must be addressed to provide credible estimates of the MWTP to live near national parks. The first is the bias that arises when unobserved local characteristics are correlated with the proximity to national parks. One solution to this is the difference-in-difference (DID) estimator. But this is only applicable when there is a change in the national park designation and then this provides an estimate of the change in MWTP from before and after the designation. The problem here is that most national parks have been in existence for many years that predates the start of most house price transactions samples. One of the advantages of having access to the ZTRAX data for such a long time period is that we can include areas that were more recently established or redesignated as national parks to be able to apply the DID framework. We estimate such a model for five national parks.

Applying DID means that we can estimate the value-added when the national park is established or redesignated. Why might households' MTWP change once the site is designated as a national park, particularly given that the site was open space to begin with? Reasons include more public knowledge about the park and its amenities, more park amenities or special programs arising from a larger operating budget, and more restrictions on activities in the park. National Parks face the most restrictions on activities, but the exact details of these restrictions vary from park to park. For example, recreational activities such as motorboats and snowmobiles are generally allowed in national recreation areas and national seashores but not in National Parks. Extraction of natural resources is generally prohibited in National Parks though there are exceptions (Szabó and Ujhelyi 2022).

Walls (2022) also raises the point made by Kuminoff & Pope (2014) and others that when using the DID framework, the change in the treatment variable can result in a new market equilibrium. Not accounting for this change in the hedonic house price function can bias the results. We address this problem by allowing for the coefficients in the hedonic model to vary before and after the national park is established or redesignated.

We add a third concern about the validity of the coefficient estimates to be correctly interpreted as the MWTP to live near the national park: changes in the coefficient estimates related to living near the park (particularly before and after a designation) can also reflect changes due to the housing cycle versus actual changes in the MWTP to live near the park. Zabel (2015) finds that the MWTP for open space in the Greater Boston Area is countercyclical. This can reflect changes in the marginal buyer and their MTWP rather than a change in the population MWTP for open space. We will take this into consideration when evaluating the change in the

MWTP to live close to a national park over time by allowing the MWTP to vary by year and then comparing this time series to the local housing cycle.

In one example, we consider ten National Historical Parks and National Historic Sites in an urban setting. Because none of these national parks were designated within the time period covered by our transaction data, we use “pseudo sites” as control groups that account for changes in the coefficients for living near the national park due to changes in the housing market. Then the difference in the coefficient estimates for the actual national parks and the “pseudo sites” provide a more accurate estimate of the change in the MWTP to live near the parks over time.

In addressing these three concerns, we feel that we are making an important contribution to the literature on using hedonic models to not only estimate the MWTP to live near national parks but open space in general. In fact, we show that what can originally be seen as significant positive MWTP to live near national parks can be attributed to not fully addressing these three issues.

The main research questions that we answer in this study are 1) What are the MWTPs to live near national parks as estimated using the hedonic method? 2) Do these estimates vary across the different types of national parks? and 3) What is the value added for areas that are established or redesignated as national parks during our sample period?

The estimates of the impact of living near 18 national parks show considerable variation across the different types of NPS units we analyzed. They even differ across national parks in the same area. Saguaro National Park consists of two nearby parks; East and West. The impacts of living near Saguaro National Park East are large and significant in some years whereas they tend to be small and insignificant for Saguaro National Park West. And even for the same National Park where, for Great Smokey Mountains National Park, the estimated impacts are different on

the Tennessee and the North Carolina sides of the park. The results also show significant differences across the five areas that were established or redesignated as national parks. They include positive treatment effects, no effects, delayed negative effects, and possible negative effects.

The key takeaway from our analysis is to be careful about interpreting results without addressing important concerns that can contaminate the estimates. This is a cautionary tale; what appears to be significant MWTP to live near a national park can also be explained by other factors related to the housing market including unobserved neighborhood characteristics (that are correlated with the measures that capture distance to the park) and the housing cycle.

Section 2 provides a review of relevant literature. Section 3 describes the ZTRAX data and the special link that provides the geocoding for the open space (including national parks) data. Section 4 develops the hedonic property value model and how living near national parks is specified. In this section, we develop alternative approaches to identifying the MWTP to live near a national park including the DID framework and allowing the MTWP to live near the park to vary annually as in an event study model. Section 5 gives the results of applying the hedonic property value method to these 18 national parks. Section 6 provides concluding remarks.

2. Literature Review

There is a relatively well-developed literature that has addressed economic aspects of National Parks and other NPS units (see Walls 2022 for an overview). Several studies have examined regional economic impacts of designating new units, a particularly controversial issue in the western U.S. For example, Jakus and Akhundjanov (2019) considered nine monument designations in five western states that occurred between 1982 and 2001 and found no evidence

of a change in county-level income associated with designation. Walls et al. (2020) looked at 14 monument designations in eight states between 1991 and 2014 (two sites overlap with Jakus and Akhundjanov) and report increases in employment and businesses established, but no effect on wage income following designation within 25 kilometers of the monuments.

Using a comprehensive dataset covering all NPS units over time, Szabó and Ujhelyi (2022) estimate the impact of a National Park designation. They find that this increases visitation rates by 17% after two years. It also leads to a 4% increase in local employment and a 5% increase in local income within four years after the designation. They also find that the increase in visitation rates can result in negative local externalities. Traffic fatality rates increase by 10% within four years after designation but there is no significant impact on local air pollution. Some prior studies have also estimated the effect of redesignation from national monument to National Park on visitation. Weiler and Seidl (2004) and Weiler (2006) found a statistically significant increase of 12,000 to 13,000 visitors in the year following redesignation while controlling for trends in population, overall NPS system visitation, per capita GDP, and size of the unit. The authors examined eight units redesignated between 1979 and 2000. Their results suggest the rate of increase in visitation ranges from 0.4% (Saguaro) to 16% (Great Basin).

Revealed and stated-preference techniques have also been used to estimate the value of trips to NPS units, as well as changes in park quality. Heberling and Templeton (2009), Neher et al. (2013), Benson et al. (2013) and Parsons et al. (2020) all use recreational demand models to estimate trip values, which generally fall in the range of \$100 to \$200 per trip. Boyle et al. (2016) utilize a choice modeling approach to estimate the value of visibility improvements in National Parks and certain wilderness areas expected to result from the U.S. Environmental Protection Agency's Regional Haze Rule. They find that households in the southwestern and southeastern

U.S. are willing to pay between \$80 and \$120 annually, on average, for the most dramatic visibility improvements at parks in their region. This corresponds to a return to “natural conditions,” as defined in the Rule. Haefele et al. (2020) also use choice modeling techniques in a general population survey to estimate values for NPS units and programs. Their aggregate estimate of the value of all NPS lands and programs is \$92 billion, which includes both use and nonuse values. This research appears in a compendium edited by Bilmes and Loomis (2020) that also includes studies on the value of NPS educational resources and carbon sequestration in NPS units, among other topics.

Turning to hedonics, the method has been used with some frequency to investigate the property value impacts associated with proximity to protected federal lands, including national wildlife refuges (Neumann et al. 2009; Liu et al. 2013) and national forests and wilderness areas (Hand et al. 2008; Ham et al. 2015). However, Mueller et al. (2021) is the first study to use the hedonic method to value the MWTP to live near a National Park. They focus on the two parks: Saguaro National Park West and Tucson Mountain Park. They use a nearest neighbor matching technique to estimate the impact of living within 1 km compared to living more than 1 km from each site using transaction data for single family homes from 2015 to 2020. They do the same for living within 2 km and 3 km of the parks. Their matching procedure is conditional on house and local census characteristics. They find that homeowners are willing to pay, on average, \$124,000 to live within 1 km of Saguaro National Park West, and \$160,000 to live within 1 km of Tucson Mountain Park (\$2020). These amounts are \$48,000 and \$63,000 for residing within 1 and 2 km of these two parks and \$22,000 for living within 2 and 3 km of Saguaro National Park West. On aggregate this is \$227.4 mil. and \$133.3 m. for households living within 1 km of Saguaro National Park West and Tucson Mountain Park, respectively.

Here, taking advantage of the national scale of the ZTRAX data, we consider numerous National Parks in various settings across the U.S. A study that is similar to ours that is carried out on a national scale in England is Gibbons et al. (2014). They estimate the house price impact associated with proximity to a variety of natural amenities such as habitats, designated areas, domestic gardens, and 10 national parks. The data period is relatively short (1998-2008) so they cannot rely on changes in these amenities to estimate the proximity effects. They can only rely on fixed effects approximating labor market areas and other controls to isolate the impacts of the natural amenities.⁵ They do find that the proximity to a number of these amenities has a significant and positive relationship with house prices. Houses in census wards that are totally inside a national park are valued 4.8% higher than houses just outside the national park but this impact is not significant. The distance to the nearest national park is also included. The price of a house that is 1 km closer to the nearest national park is 0.24% higher, all else equal. This is a small difference.

3. Data

Sales data for single family homes are from the ZTRAX-based PLACES dataset.⁶ ZTRAX is the largest database of property records made available free of charge to U.S. academic researchers. It consists of three databases including a property transaction database, a tax assessment database, and historical versions of tax assessment databases to allow tracking of changes to the database at the property record level. ZTRAX contains tax assessor data on property characteristics, geographic information, and valuation for approximately 150 million parcels in over 3,100 U.S. counties, in addition to more than 400 million public transaction records. Researchers at Boston

University have linked most ZTRAX records for the contiguous U.S. to county-level digital parcel maps using text-based tax assessor parcel identifiers and conversion algorithms to develop the Private-Land-Conservation Evidence System (PLACES) (Nolte et. al 2023, this issue).⁷ All geoprocessing (e.g., distance to national parks, open space, roads, lakes, and rivers), was implemented in PLACES.

We will estimate the MWTP to live near a national park under various settings and types of national parks to see how this MWTP can differ. The first consideration for choosing national parks for this study is that there must be a large enough housing market near the park and hence a large enough number of sales during the time covered by the data to obtain precise estimates of MWTP to live near the park. Ideally, the national park is established during the period covered by the data so that we can apply the difference-in-difference (DID) framework to identify changes in MWTP. Another case where we can apply DID is when existing units (e.g., national monuments) are redesignated as National Parks. These parks were part of the park system with “lesser” designations such as National Preserve, National Recreation Area, or National Seashore.

National parks can only be designated by Congress and the designation process can be quite drawn out (Szabó and Ujhelyi 2022). This means that the designation is unlikely to be tied to the state of the housing market which will likely have changed since the beginning of this process. Hence, we can treat the designation as plausibly exogenous in the context of the hedonic house price model. Furthermore, given this drawn-out process, one might expect that it will take years for the full impact to materialize, which is a reason for letting the MWTP to live near the park to vary over time.

We include five national parks that were established or re-designated during our data period.

- Birmingham Civil Rights National Monument, Alabama (BICR) is only 0.88 acres in size and mostly is located within the larger 36-acre Birmingham Civil Rights District, which was designated in 1992 by the City of Birmingham. It was established as a National Monument on January 13, 2017, so 2017 is the designated treatment year.
- Cuyahoga Valley is a 32,572-acre park that is located between Akron and Cleveland. It was originally designated as a National Recreation Area in 1974 and then was redesignated as Cuyahoga Valley National Park, Ohio (CVNP) in 2000.
- Paterson Great Falls (PGF) is a 77-foot waterfall on the Passaic River in New Jersey. It was established as a National Historical Park on Nov 11, 2011, so we choose 2012 as the treatment year.
- River Raisin National Battlefield Park (RRNB) was the site of the Battle of Frenchtown in Michigan and is the only national battlefield marking a site of the War of 1812. It was established on Oct 22, 2010, so we choose 2011 as the treatment year.
- Tule Springs Fossil Beds National Monument, Nevada (TSFB) is a 22,650-acre park that is located 20 miles from Las Vegas. It was established on Dec 19, 2014, so the treatment year is set at 2015.

Most other national parks were established before the period our data covers. In these cases, we rely on other means for identifying the MWTP to live near a national park. First, we allow the impacts to vary over time to see how preferences for living near a national park might change over time such as during the COVID-19 pandemic. As stated earlier, any such changes can be, in part, driven by the housing cycle. We can use the change in overall market prices as the “control” and compare this change to that of the MWTP to live near the park. Then we can

consider the differences between these two measures over time when identifying the change in the MWTP to live near the park.

One case we analyze is a set of 10 National Historic Sites (NHS) and National Historical Parks (NHP) in the Greater Boston Area (GBA: Essex, Middlesex, Norfolk, Plymouth, and Suffolk counties). Usually, an area is designated as a NHS due to its single historical feature. Such Sites in the GBA include the Salem Maritime and Saugus Iron Works National Historic Sites. NHPs tend to be larger in area than NHSs and are not designated because of a single historical feature. We choose this case as these units are in a densely populated urban area, which is quite different than the relatively rural areas where many national parks tend to be found. These parks are also relatively small compared to the typical national park; three rank among the smallest 30 national parks: John F Kennedy NHS (0.09 acres), Boston African American NHS (0.59 acres), and Longfellow House Washington Head Quarters NHS (1.98 acres). It is also possible in this setting to use similar placebo sites to act as a good control group to compare to the national parks. We can do this for the GBA since there are nearby areas that are similar to the national parka that can be used as placebo sites.

Two of the three National Parks we analyze are Saguaro National Park and the Great Smokey Mountains National Park. Saguaro National Park is a large and well-known park near Tucson AZ. It was declared a National Monument on March 1, 1933, and was designated a National Park in 1994. Our data cover the years 1994 – 2019 so we cannot carry out a DID analysis to estimate the impact of these National Park designations. What is interesting about this 92,000-acre park is that it consists of two separate areas—the Rincon Mountain District about 10 miles east of Tucson and the Tucson Mountain District about 10 miles west of the

city. These two areas preserve the giant saguaro cactus and are referred to as Saguaro National Park East and West. Bordering Saguaro West is Tucson Mountain Park. The Pima County Parks Commission was established to oversee the park in April 1929. It is one of the largest natural resource areas owned and managed by a local government in the U.S (approximately 20,000 acres).⁸ Also nearby is Coronado National Forest. We will include all four areas in our analysis to see if households value living near a National Park versus a National Forrest and a local park differently.

Great Smokey Mountains National Park (GSMNP) was established as a National Park on June 15, 1934, and it is America’s most visited National Park. With a size of 522,427 acres, it is the largest national park analyzed in this study. And it is the only National Park that is free to visit all year. What is interesting about GSMNP is that it borders two states, North Carolina (NC) and Tennessee (TN). So, we estimate separate hedonic equations for each state.

The list of national parks that we include in our analysis and the years covered for each park are given in Table 1. The maximum period is 1987-2021 that is used for the Greater Boston Area. We include sales data from the counties in which the national park is located, as well as all adjacent counties. We apply the same process for cleaning the data to each national park. These steps are listed in the Data appendix in Table A1. We use the same variables in all regressions to provide uniformity to the analysis.

[Insert Table 1 here]

We include control variables in the hedonic regressions related to the distance to the nearest lake, river, and highway, age of the unit, the number of bathrooms, bedrooms, and total rooms, and the unit and lot sizes. Many of these control variables have missing and/or extreme values and the number of these missing and/or extreme values can differ substantially across

the data sets since the data are reported by each state.⁹ This is an important problem with tax assessor data compiled across many county offices, including ZTRAX (Nolte et al. 2023, this issue). We do not want to drop these observations given that they include useful information on the key relationship of interest: the house price and proximity to the national park. For lake, river, and highway access and lot size, we generate a separate flag for each variable for missing observations and set the missing values to zero. For building size, we generate a flag for values less than 500 square feet as well as for missing observations and set these values to zero. For age (of the house), number of bedrooms, bathrooms, and total rooms, we include indicators of ranges of values for these variables. This means that extreme (and implausible) values are included in the indicator of the largest (or smallest) values (e.g., greater than or equal to five bathrooms) which minimizes their influence on the results. For unit age, we include an indicator of observations where age is less than zero or missing. For beds and baths, we include an indicator of observations where the value equals zero or is missing. For total rooms, we include an indicator of observations where the value is less than or equal to three or is missing (we view three or fewer rooms as implausible for single family houses). Included in the Data Appendix are summary statistics for each data set (including sample size) and maps of each area.

For the parks that were established or redesignated during the period covered by the data, we can check to see if there are any supply-side effects by looking at changes in new sales before and after a national park is established or redesignated for near (treatment) versus far (control) areas from the site. If this action results in significantly more new housing in the area near the park, this could blunt any price effects and vice versa. We also look at changes in the number of transactions near versus far from the site before and after a national park is

established or redesignated. We provide the results for each of these five parks in Table A2 in the Data Appendix. Generally, there is little evidence of significant differences in the before and after percent changes in either new supply or transactions for the near versus far areas.

We also look at changes in the means of observable housing characteristics to see if there was a change in the types of houses that were sold in the treatment and control areas. We run DID regressions such that the estimated coefficient on the interaction term of the before and after periods and treatment indicator is the key statistic. Table A2 in the Data Appendix provides the p-value for this coefficient estimate and the semi-elasticity; the coefficient estimate divided by the mean of each variable in the period before the treatment for the control group (a measure of economic significance). Note that some elements in the table are missing because these variables are missing in these cases. This is due to the varying availability of information on these variables by site in the ZTRAX data. We also include the change in price in this table as a simple indicator of the main treatment effect. We can see that the estimated treatment effect is negative for three of the five parks and it is positive for the other two parks.

For the house characteristics, the p-values are a mix of significant and not significant (at the 1% level) values. The fact that many are statistically significant should not necessarily be taken as evidence of a meaningful difference given the large sample sizes. Therefore, we also include the semi-elasticity, which is not a function of sample size. None of these semi-elasticities are greater than 20%, an often-used cutoff for economic significance. And there are only three values that are greater than 10%, all for lot size; two are positive and one is negative (but not statistically significant). The two positive values indicate that the difference in lot sizes before and after the treatment was larger for the treatment versus the control group.

Overall, this does not provide much evidence of significant differences in the number and/or types of houses that were sold before and after the treatment year for the treatment and control groups. This means that the estimated MWTPs are little (if at all) affected by these differences.

4. Hedonic Model and the Specification of Open Space

Because there is no market for local public goods and local amenities or disamenities, a standard economic approach to measuring their value is based on indirect information in the housing market. This method uses the relationship between the price of a house and its associated characteristics to identify the value that is placed on these characteristics through market transactions. Characteristics that have been valued extensively in this manner include school quality, crime, environmental disamenities (air pollution, hazardous waste sites, and contamination that is the byproduct of firm activity) and local amenities such as open space. This relationship is specified using what is known as the “hedonic house price model.”

Typically, the dependent variable is specified in logarithmic form:

$$\log(P_{ijt}) = \beta_0 + X_i\beta_1 + N_{ij}\beta_2 + f(O_{ij};\delta) + u_j + v_t + e_{ijt}, \quad i = 1, \dots, n \text{ and } t = 1, \dots, T \quad [1]$$

where P_{ijt} is the real price of property i , in block group j , at time t (we use monthly data), X_i is a vector of characteristics for property i ¹⁰, N_{ij} is a vector of neighborhood (local) amenities, $f(O_{ij};\delta)$ is some function of proximate open space, O_{ij} , u_j and v_t are block group and time fixed effects, and e_{ijt} is a standard error term. For this analysis, all the neighborhood characteristics are distances to nearby landmarks (lake, river, highway) so they do not vary over time. There are n housing transactions in the sample and T time periods (month by year). The parameters β_1 , β_2 , and δ are the values that the market places on the house characteristics X_i ,

neighborhood amenities N_{ij} , and open space O_{ij} , respectively. House prices are put in real terms using the local CPI. Across most of our analyses, January 2020 is the base month.

Given that it can be difficult to observe all measures of neighborhood quality, we use block group fixed effects in the model as this captures factors that affect all houses in that block group. Note that because u_j does not depend on i or t , it captures measures of neighborhood quality in j that do not vary over time or within j , whereas N_{ij} will control for changes in neighborhood quality within the block group. Note that identification of the impact of proximity to open space on house prices comes from differences in O_{ij} within the block group.

It is important to recognize that open space is complex and multi-dimensional. So, it is important to control for other measures of open space when estimating the MWTP to live near national parks. In addition to measures of the distance to the nearest national park, NPD, we include the percent of open space that is protected within 5000 meters, PROT_5000. We start by specifying δ_1 and δ_2 to be constant over time as this is the simplest case to characterize the variation in PROT_5000 and NPD that is used to identify and estimate these parameters. We refer to the version of equation (1) with the following specification for open space as Model 1

$$f(O_{ijt}; \delta) = \text{PROT_5000}_i \delta_1 + \text{NPD}_{ij} \delta_2 \quad [2]$$

Changes in PROT_5000 are likely to be endogenous, so we fix the value at the beginning of the period. Note that since PROT_5000 varies by housing unit, it varies within the block group, so δ_1 can be estimated.

We expect that the marginal valuation of living near a national park will only be significant for a reasonably small radius around the site. For Saguaro National Park, Mueller et al (2021) use up to 3 km. That is, it is assumed that the MWTP to live close to the park is zero for units that are more than 3 km from the park. Furthermore, they use three distance bands: ≤ 1

km, 1 km – 2km, 2 km – 3 km. We use a similar procedure for specifying the proximity to a national park. The actual distance bands that we use will vary by national park unit. This will depend on the number of units sold within specific distance bands (we want to have enough sales to be able to precisely estimate the MWTP in that distance band) and how far the impact of the national park on house prices extends. Note that the control group consists of houses outside these distance bands and in the counties included in each national park data set. In the above example, these would be units more than 3 km from the national park. As a robustness check, we also estimate a model with the actual distance to the site rather than using the distance bands. We show that this has little impact on the results (see Section 5E).

Our goal is to produce an estimate of δ_2 that will measure the MWTP to live near a national park. But because the locations of national parks are not randomly assigned, we need to consider other unobservable factors that might be correlated with both house prices and NPD that would then bias our estimates. We include block group fixed effects which can mitigate a significant part of this bias. This means that for national parks that are not established nor redesignated during the sample period, identification of δ_2 comes from block groups where there is variation in the distance bands, that is, from sales of houses on either side of the border between two distance bands in the same block group. Note that bias in the estimate of δ_2 can arise if houses in these block groups that are closer to the national park and in the nearer distance band contain higher neighborhood quality (in unobservable ways) than those that are in the farther distance band.

We do include five national parks that were either established or re-designated during the period of our data so we can specify and estimate a DID model for these cases. Then δ_2

varies from before and after the designation. We refer to the version of equation (1) with the following specification for open space as Model 2

$$f(O_{ijt}; \delta) = \text{PROT_5000}_i \delta_1 + \text{NPD}_{ij} \delta_{21} + \text{NPD}_{ij} \cdot \text{AFTER}_t \delta_{22} \quad [3]$$

where $\text{AFTER} = 1$ after the park was established or re-designated, and 0 otherwise. In response to Walls' (2022) critique that when using the DID framework, the change in the treatment variable can result in a new market equilibrium, we allow all the coefficients in the hedonic model to vary before and after the designation (including the block group fixed effects).

Since we now compare differences over time within block groups, the above source of bias will only be an issue if the relative MWTP for the unobserved quality between houses closer and farther away from the national park increases from before to after designation.¹¹ And then even if there is no heterogeneity in the distance band within the block group, the change in δ_{22} from before and after the designation can still be identified.¹² This would involve using a version of the standard two-way fixed effects (TWFE) estimator where the control group are houses outside the farthest distance band and included in the counties selected for each national park data set.

The above discussion of Model 2 motivates a generalized version of the standard DID model such that δ_2 (and δ_1) is allowed to vary in every period (yearly in this case). We refer to the version of equation (1) with the following specification for open space as Model 3

$$f(O_{ijt}; \delta_t) = \text{PROT_5000}_i \delta_{1t} + \text{NPD}_{ij} \delta_{2t} \quad [4]$$

Model 3 incorporates both cases where the national parks were either established or re-designated during the period of our data and where the park's designation occurred prior to the start of the data period. In the former case, allowing δ_{2t} to vary for each year produces what is

referred to as the “event study” model in the DID literature. This allows for a test for parallel pre-trends and for a time-varying treatment effect. The latter seems very reasonable in this case as it will likely take many years for the impact of the national park designation to be fully capitalized into house prices. In the case of the national parks that were established before the beginning of the sample data, there is no change in the characteristics of the park, but the variation in δ_{2t} can result from changes in preferences that would cause these impacts to vary. For example, there is some evidence that the COVID-19 pandemic had led to increased preferences for living near open space (Malik et al. 2023, this issue).

How are the parameters identified when estimating Models 1 -3? For simplicity, assume there are two distance bands (NPD = 0 or 1). First, consider Model 1 (equation 2) where δ_2 is fixed. We rely on block groups with a change in the distance band

$$\begin{aligned} & E\left[\log(P_{ijt}) \mid \text{NPD} = 1, v_t = 1, X_0, N_0, u_j\right] - E\left[\log(P_{ijt}) \mid \text{NPD} = 0, v_t = 1, X_0, N_0, u_j\right] \\ &= (\beta_0 + \delta_2 + X_0\beta_1 + N_0\beta_2 + u_j + v_t) - (\beta_0 + X_0\beta_1 + N_0\beta_2 + u_j + v_t) \\ &= \delta_2 \end{aligned} \quad [5]$$

For Model 2 (equation 3), assume the designation happens in period t_0 . δ_{22} is identified by block groups with a change in the distance band for $t \geq t_0$ as in equation (5). But we also rely on transactions before and after the designation in the treatment and control groups. This includes block groups that do not contain a change in the distance band (and for simplicity, assume that the other coefficients are constant from before and after the designation)

$$\begin{aligned}
& E\left[\log\left(P_{ij,t+k}\right) \mid \text{NPD} = 1, \text{AFTER} = 1, v_{t_0+p} = 1, X_0, N_0, u_j\right] \\
& - E\left[\log\left(P_{ijt}\right) \mid \text{NPD} = 1, \text{AFTER} = 0, v_{t_0-q} = 1, X_0, N_0, u_j\right] \\
& = \left(\beta_0 + \delta_{21} + \delta_{22} + X_0\beta_1 + N_0\beta_2 + u_j + v_{t_0+p}\right) - \left(\beta_0 + \delta_{21} + X_0\beta_1 + N_0\beta_2 + u_j + v_{t_0-q}\right) \\
& = \delta_{22} + \left(v_{t_0+p} - v_{t_0-q}\right)
\end{aligned}$$

[6]

and

$$\begin{aligned}
& E\left[\log\left(P_{ij,t+k}\right) \mid \text{NPD} = 0, v_{t_0+p} = 1, X_0, N_0, u_m\right] - E\left[\log\left(P_{ijt}\right) \mid \text{NPD} = 0, v_{t_0-q} = 1, X_0, N_0, u_m\right] \\
& = \left(\beta_0 + X_0\beta_1 + N_0\beta_2 + u_m + v_{t_0+p}\right) - \left(\beta_0 + X_0\beta_1 + N_0\beta_2 + u_m + v_{t_0-q}\right) \\
& = v_{t_0+p} - v_{t_0-q}
\end{aligned}$$

where p and q are integers greater than zero. The difference between these two differences identifies δ_{22} . Here we expect a change in MWTP as a response to the “treatment” which are the changes in the site characteristics that are a result of the national park designation. Note that the first of the above two expressions fixes the block group at u_j and the second at u_m to highlight that these can be different block groups (even though it is possible they could be the same block group).

In Model 3 (equation 4), we can identify δ_{2t} using block groups with a change in the distance band. But we can also identify changes in δ_{2t} using block groups without a change in the distance band.

$$\begin{aligned}
& E\left[\log\left(P_{ij,t+k}\right)\middle|NPD = 1, v_{t+k} = 1, X_0, N_0, u_j\right] - E\left[\log\left(P_{ijt}\right)\middle|NPD = 1, v_t = 1, X_0, N_{0j}, u_j\right] \\
& = \left(\beta_0 + \delta_{2,t+k} + X_0\beta_1 + N_0\beta_2 + u_j + v_{t+k}\right) - \left(\beta_0 + \delta_{2,t} + X_0\beta_1 + N_0\beta_2 + u_j + v_t\right) \\
& = \delta_{2,t+k} - \delta_{2,t} + \left(v_{t+k} - v_t\right)
\end{aligned}$$

and

[7]

$$\begin{aligned}
& E\left[\log\left(P_{ij,t+k}\right)\middle|NPD = 0, v_{t+k} = 1, X_0, N_0, u_m\right] - E\left[\log\left(P_{ijt}\right)\middle|NPD = 0, v_t = 1, X_i, N_{ij}, u_m\right] \\
& = \left(\beta_0 + X_0\beta_1 + N_0\beta_2 + u_m + v_{t+k}\right) - \left(\beta_0 + X_0\beta_1 + N_0\beta_2 + u_m + v_t\right) \\
& = v_{t+k} - v_t
\end{aligned}$$

where k is an integer greater than zero. The difference in these two differences identifies $\delta_{2,t+k} - \delta_{2,t}$. In this case u_j and u_m must be different block groups. $\delta_{2,t+k} - \delta_{2,t}$ is identified by changes over time in household preferences for living near a national park such as what might have happened during the COVID-19 pandemic. Note that one can estimate Model 3 using only the block groups in which there is no change in the distance bands.

A concern is that changes in β_{3t} can arise due to changes in housing market conditions rather than changes in the MWTP to live near the park. Zabel (2015) provides evidence that the impact of open space on house prices is counter-cyclical. This is driven by the WTP for the marginal buyer who can change across the housing cycle. So, the population WTP may not change but the marginal buyer selected from the population does change with the housing cycle. To get a counter-cyclical movement in the MWTP, it could be the case that when the market is booming, all potential buyers enter the market, and the marginal buyer is likely to be one with a relatively lower income with a lower WTP for open space. When the housing market is in decline, there are fewer buyers in the market and the marginal buyer is then likely to be one with a relatively higher income who can still afford to buy a house. This marginal buyer

will have a higher WTP for open space. We will show examples of where the MWTP to live close to a national park appears to be procyclical and hence the marginal buyer follows a reverse pattern. So, we can observe changes in the estimated values for δ_{2t} even though the population WTP might not change due to the different, non-random, samples of buyers over the housing cycle.

As noted earlier, factors that would lead to a positive MWTP to live close to a national park include better access to the park's amenities, a better view and/or a nicer neighborhood due to the natural amenities near the park, and a guarantee that the open space will not be developed. Factors that would lead to a negative MWTP to live close to a national park include traffic and the resulting air pollution and the general inconvenience of many visitors in the neighborhood. So, whether the MWTP is positive, negative, or zero is an empirical question.

What changes might be valued when the national park is established or redesignated? These include more public knowledge about the park and its amenities, more park amenities or special programs from a larger operating budget, and more restrictions on activities in the park. National Parks face the most restrictions on activities, but the exact details of these restrictions vary from part to park. For example, recreational activities such as motorboats and snowmobiles are generally allowed in national recreation areas and national seashores but not in National Parks. Extraction of natural resources are generally prohibited in National Parks though there are exceptions (Szabó and Ujhelyi 2022).

5. Results

Depending on the circumstances, we estimate the different versions of the underlying hedonic model (Models 1-3) that includes housing and neighborhood characteristics, year by month dummies, and block group fixed effects. The open space measure PROT_5000 and the distance bands to the nearest national park are included and their corresponding coefficients are first fixed and then allowed to vary over time. As previously mentioned, the distance bands vary across the national parks. We want to make sure that there are enough sales in each band to accurately identify the treatment effects. Because the density of housing is different across the parks, this is what leads to the use of different sized distance bands to make sure we have an adequate number of sales in each band.

Generally, these open space measures are not very correlated with the housing and neighborhood characteristics, so multicollinearity issues are not a problem of concern.¹³ What is interesting about the estimates of the time-varying impacts of PROT_5000 is that they take on significant negative and positive values and tend to be counter-cyclical. The yearly impacts are jointly significant at the 1% level in each case (these results and all regression results are available upon request).

One might be concerned that since PROT_5000 (the percent of open space that is protected within 5000 meters) accounts for NPS units, then including it in the regression may mute the impacts of living near the sites of interest. In fact, these separate measures of open space are not very correlated, and as we show in the section addressing robustness checks (5E), dropping PROT_5000 has little impact on the estimates for the distance bands.

Most of the areas surrounding the units that we include in our study followed a housing cycle that peaked before the Great Recession, then bottomed out around 2010 and have recovered since then. We generate house price indices for each area that we can compare to the annual estimates for the MWTP to live near the NPS unit.

Established/Redesignated Sites

We first consider five national parks that were either established or redesignated within the sample periods covered by the ZTRAX data. This allows us to carry out DID analyses (Model 2) that are better able to control for unobservable neighborhood effects than the standard hedonic model (Model 1). We allow all the coefficients in the hedonic model to vary before and after the designation given that this can result in a new market equilibrium (Kuminoff & Pope 2014). In all five cases, the coefficients from before and after designation are significantly different supporting the claim that the designation results in a new housing market equilibrium. The main difference when compared to estimating the DID model assuming one equilibrium (constant coefficients before and after the designation) is that the treatment effect is much less precisely estimated since the effective sample size is limited to the post-treatment years versus the full sample.

The five national parks that we analyze, and their dates of establishment or redesignation are given in Table 1. We choose a treatment year based on these dates as the basis for estimating the treatment effects for each site. We first use the standard DID model (Model 2) where we estimate the change in the MWTP to live near the site from before and after the designation, δ_{22} .

We then estimate the “event study” model where δ_{2t} is allowed to vary in every year (Model 3). This approach allows one to see the evolution of the treatment effect that is obscured when using the DID approach that only estimates a single treatment effect. We normalize the impact in the year prior to the designated treatment year to be 0. We then subtract the estimated δ_{2t} coefficient from the year prior to the treatment year from the estimated MWTP coefficient for all other years. We test for pre-trends for each distance band for each site and in almost all cases we do not reject parallel trends at the 1% significance level (the case where we do reject parallel trends is discussed below). We present the estimated annual treatment effects for each of the five national parks in Figure 1. Note that the house price index is also set to zero in the year of designation to make it directly comparable to the treatment effects.

[Insert Figure 1 here]

Birmingham Civil Rights National Monument, Alabama (BICR) was established January 13, 2017, so 2017 is the designated treatment year. The data cover the years 1998-2021 and the number of observations is 275,643. We use distance bands of ≤ 3 km, 3-5 km, 5-7 km and 7-10km. The treatment effects estimated using the DID model (Model 2), while positive and as large as 20%, are not significant (Panel A of Table 2). The imprecision in these estimates is possibly due to the relatively few post-treatment years (5) versus pre-treatment years (19). The results from estimating Model 3 are shown in the first row of Figure 1. While there is some evidence of a positive and sustained treatment effect for three of the four distance bands that reach as high as 20% for the two closest distance bands, none of the individual estimates are significant nor are the five annual treatment effects jointly significant for any of the distance

bands. Note that we restrict Figure 1 to the years 2012-2021 as this provides a better picture of the treatment effects after designation starting in 2017.

[Insert Table 2 here]

Cuyahoga Valley National Park, Ohio (CVNP) was redesignated in 2000 so 2001 is considered the treatment year. The data cover the years 1993-2021 and the number of observations is 492,635. We use distance bands of ≤ 1 km, 1-2 km, 2-3 km, and 3-5km. The treatment effects estimated using the DID estimator are small and not significant (Panel A of Table 2). When we allow the impacts to vary by year (Model 3), the treatment effect is initially negative and declines for the first five years, then gradually increases to around zero in 2015, and then declines again to close to 20% by 2021 for the ≤ 1 km distance band (second row of Figure 1). The individual point estimates are only significant in the final years though they are not jointly significant (at the 1% level). For the other three distance bands the treatment effect is initially not significant but turns negative around 2006, declines until around 2011 and then levels off at a negative value of -10%, -15%, and -15%, respectively (second row of Figure 1). Is this evidence of a negative treatment effect? There may be weak evidence of this for the closest distance band. Note that this is the only example of a redesignation, and this negative effect could be due to increased visitation rates and traffic (Szabó and Ujhelyi 2022) relative to little perceived benefit of proximity from the redesignation. The negative impacts for the other three distance bands do not happen until 2006 and they coincide with the drop in the housing market, so this is less compelling evidence of an actual decline the MWTP to live near CVNP.

Paterson Great Falls National Historical Park, New Jersey (PGF) was established on Nov 11, 2011, so we choose 2012 as the treatment year. The data cover the years 1994-2021

and the number of observations is 779,069. We use distance bands of ≤ 1 km, 1-2 km 2-3 km, and 3-5km. Using the DID estimator, the estimated treatment effects for the closest three distance bands are negative, significant, and larger than 10% in magnitude.

When we allow the impacts to vary by year (Model 3), the picture is quite different. Initially, the impacts are large and significant but starting around 2006, they decline until the year of establishment in 2011, then they increase monotonically and reach impacts of 81%, 54%, and 37% by 2021 (third row of Figure 1). Starting in 2006, one can see that these treatment effects closely follow the housing cycle; between 2006 and 2021 the correlation of the price index with the three distance bands is 0.85, 0.87, and 0.90, respectively. So, it appears that the changes in the MWTP follow the housing cycle, and this just reflects a selection effect (of the marginal buyer). These are very large effects, yet the house price index only increases by 28% as of 2021. Hence, one can argue that the treatment effect significantly outpaces the price index and indicates a positive net effect of 53%, 26%, and 9% for the ≤ 1 km, 1-2 km, and 2-3 km distance bands, respectively. The result for the ≤ 1 km band seems implausibly large. Yet note that, like the house price index, these impacts are also not significant until around 2018. For this to be caused by an increase in the MWTP to live near PGF would mean that there is a six-year delay in when this takes effect. It seems more likely that this increase was driven by the concomitant rise in the housing cycle.

River Raisin National Battlefield Park, Michigan (RRNB) was established on Oct 22, 2010, so we choose 2011 as the treatment year. The data cover the years 1994-2021 and the number of observations is 575,846. We use distance bands of ≤ 3 km, 3-5 km, and 5-7 km. The DID estimate of the treatment effect is -10.2% for the nearest distance band and 8.3% and 6.5% for the other two though none are significant.

When we allow the impacts to vary by year (Model 3), the treatment effects eventually turn increasingly negative after the designation and by 2021 are -30.7% , -33.0% , and -40.6% and marginally significant (fourth row of Figure 1). These treatment effects exhibit countercyclical behavior compared to the local housing cycle (the correlations of the three treatment effects and the price index are -0.96 , -0.95 , and -0.86 , respectively, since 2012 and are also significantly negative before 2003). House prices rose by 127% over this period. So, what appears to be an increasingly negative and large treatment effect could just be (countercyclical) housing cycle effects. The estimated distance band impacts are missing for 2003-2008 because of a lack of data for those years the results are only presented starting in 2009.

Tule Springs Fossil Beds National Monument, Nevada (TSFB) was established on Dec 19, 2014 so we choose 2015 as the treatment year. The data cover the years 1994-2018 and the number of observations is 1,480,403. We use distance bands of ≤ 3 km, 3-5 km, and 5-10 km. Based on Model 2, the treatment effects for these 3 distance bands are each around -10% but only the one for the farthest distance band is significant. When we allow the impacts to vary over time, we get a similar story. The treatment effects are each negative and not significant though they do increase and are close to zero for each of the three distance bands in 2018. (fifth row of Figure 1). Note though that the impact was increasing starting in 2012 in all three cases, and this is evidence of significant positive pre-trends for all three distance bands that also lines up with the increase in house prices after the market started to rebound after the Great Recession.

These results provide examples of the value added for areas that were established or redesignated as national parks. They include delayed significant, positive treatment effects (PGF), weak evidence of positive treatment effects (BICR), weak evidence of delayed negative

treatment effects (CVNP) and weak evidence of negative treatment effects (RRNB, TSFB). But upon further inspection, one is led to the conclusion that at most one of these results (PGF) shows any strong evidence of significant (positive or negative) MWTP to live near a national park.

Greater Boston Area

The Greater Boston Area (GBA) dataset is comprised of 10 NPS units with either National Historic Site (NHS) or National Historical Park (NHP) designations. The data cover the years 1987-2021 and the number of observations is 966,983. We include distance bands to the nearest NHS and to the nearest NHP. The bands are indicators of being within 1 km, 1-2 km, and 2-3 km from the nearest NHS/NHP.

First, we estimate Model 1 where the coefficients for the distance bands do not time-vary. All six estimates for the three distance bands for the sites and parks are small and not significant (see Table 2). Next, we estimate Model 3 where the coefficients for the distance bands time-vary. Figure 2 provides the annual estimated percent change in price for the distance to the nearest NHS and NHP for each of the three distance bands along with confidence intervals for each estimate. Many of the yearly impacts are not significant until around 2010; they turned positive and significant since then. The largest impact is around 15%. The yearly impacts are jointly significant at the 1% level for each of the three distance bands for the nearest NHS and NHP. Note that the increase in these estimates starting in 2011 coincides with the upturn in the GBA housing market after the Great Recession (see Figure 2).

[Insert Figure 2 here]

These results appear to indicate that households have increasingly valued living near these sites and parks in the last 8-10 years. And some impacts have reached as high as 10-15%. Though there is no evidence of an increase since the beginning of the COVID-19 pandemic (for 2020 and 2021).

One problem with the interpretation of these results is that the sites and parks were always in existence over the time period our data covers. As stated earlier, any changes in the MWTP would therefore come from changes in preferences versus changes in the characteristics of the parks. But these increases in the estimated impacts could also come from relatively higher increases in either the neighborhood quality itself or in the MWTP to live in the neighborhoods that are nearer the park and in the closer distance band. Again, this could reflect the cyclic nature of the housing market versus actual changes in the MWTP to live near the parks.

This relative increase in the MWTP for nearer neighborhoods will not be accounted for in the control group (houses further than 3 km from the parks) if it does not also experience such a relative increase in neighborhood quality. And this is possible since the control group neighborhoods are likely to be different than those closer to the parks. To address this problem, we use nearby areas with similar neighborhood quality to where the parks are located to act as “placebo” sites. Then the value of living nearer this location would not include the proximity to a NHS or NHP. By choosing locations with similar neighborhood quality, we assume that these placebo sites have a similar variation in neighborhood quality around the site so that the variation in the MWTP to live near these placebo sites that is due to changes in neighborhood quality will mimic that for the park sites. Then the difference in the impacts of the parks and

these placebo sites will be the estimate of the MWTP to live near the park as this nets out the change in MWTP due to relative changes in neighborhood quality.

We match the census block groups overlapping the NHS or the NHP with another nearby block group based on their characteristics which we obtain from the 1990 Decennial Census since it is at the beginning of the sample data period and block groups are not available in the 1980 Decennial Census.

To be out of each other's sphere of influence, the placebo sites must be more than 6 km away from the NHS/NHP. To be this far away, we find that these placebo sites are not located in the same town. We only consider block groups that are at least 6 km and at most 12 km away from a given NHS/NHP. This is the "outer ring", as shown in Figure A8.2 in the Data Appendix in green. The process of selecting a placebo site involves first standardizing the five block group characteristics of interest: percent nonwhite, percent of those 25 or older with a bachelor's degree, percent of those working in a professional occupation, the unemployment rate, and the median household income.

Next, we run a regression at the block group level of the median house price on these characteristics. Then we take the square root of the weighted average of the squared difference in the observed value of each characteristic for the block group of each NHS/NHP and each block group that is in the outer ring (within 6 and 12 kms) where the weights are the absolute values of the estimated coefficients from the above regression, $\hat{\beta}_i$. This similarity metric, SM, is then calculated as

$$SM(Y, k) = \sqrt{\sum_{i=1}^5 |\hat{\beta}_i| (x_{ki}^S - x_{Yi}^S)^2} \quad [8]$$

where Y is the block group from the NHS/ NHP and k indexes the possible placebo block groups. We multiply $SM(Y,k)$ by the distance from each possible placebo block group to the NHS/NHP, $dist(Y,k)$, and then choose as the matched placebo block group the one with the smallest value of this term which we refer to as the “placebo choice metric” (PCM)

$$PCM(Y, k^*) = \min(SM(Y, k) \cdot dist(Y, k)) \quad [9]$$

Our reason for multiplying the similarity metric by the distance to the NHS/NHP is to give preference to closer block groups. Some of the parks/sites are made up of multiple block groups and we choose the matched placebo block group as the one that has the minimum value of the PCM across all these park/site block groups. We also generate a much simpler similarity metric which is just the absolute difference in median household incomes between NHS/NHP block groups and possible placebo block groups.

We generate the same three distance rings for the placebo sites and add them to the hedonic regression for Model 3. The estimated impacts for the placebo sites match up very closely with those for the NHS/NHP estimates. The differences in impacts between the placebo sites and the NHS/NHPs are presented in Figure 3. They are generally small, not individually significant, and do not follow any discernable pattern (though the ones for the ≤ 1 km and 1-2 km bands are jointly significant at the 1% level for the NHSs as well as the ≤ 1 km band for the NHPs). This indicates that what initially looks like increasingly positive MWTP to live near the NHS/NHPs in recent years largely reflects the change in the value of living near a similar location. This change likely results from the evolving housing cycle. The results using the simpler similarity metric are quite comparable and they are presented in Figures A8.3 and A8.4 in the Data Appendix.

[Insert Figure 3 here]

Saguaro National Park

Saguaro National Park consists of two noncontiguous parts; Saguaro National Park East and West. We also include two other nearby parks, Tucson Mountain Park (that borders Saguaro West) and Coronado National Forest, to see how the MWTP differs for areas that are not National Parks. The data cover 1994-2019 and the number of observations is 283,046.

The results for the distance bands when estimating Model 1 (constant coefficient) are given in Panel B of Table 2. We see that the impact is positive, significant, and large for the ≤ 2 km band and smaller, positive, and marginally significant of the 2-3 km band for Saguaro East. Whereas the impacts of the distance bands for Saguaro West are negative but not significant. There are large positive and significant impacts for Coronado National Forest whereas there are small, negative but not significant results for Tucson Mountain Park.

We next estimate Model 3 that allows for time-varying coefficients. Figure 4 provides the estimated impacts for these four parks. We see that the estimates to live near Saguaro National Park East are initially small and not significant but then turn large, positive, and significant in 2000, reach a height of 55% in 2010, then decline but still are mostly significant through 2018 for the ≤ 2 km distance band. One can see, though, that these positive and significant impacts are countercyclical to the housing cycle which peaks in 2006, falls by around 50% by 2011 and then increases by approximately 50% by 2019. So, it is not clear that these are actual changes in the population MWTP to live near these parks or just selection effects of the marginal buyer across the housing cycle. For Saguaro National Park West, the annual impacts are generally not individually significant, and if anything, seem to follow the housing cycle in a countercyclical manner.

[Insert Figure 4 here]

The annual impacts for Coronado National Forest are initially marginally significant at best until around 2006, then turn large and significant and peak in 2011 with impacts as large as 45% in the ≤ 1 km distance band. These impacts then monotonically decline to around 10% in 2019. Like Saguaro East, these results also display a countercyclical movement. The estimates for the distance band for living near Tucson Mountain Park are mostly not significant, but displayed a procyclical movement, falling to negative territory around 2012 before recovering to near zero in 2019.

Overall, we find what initially look like positive and significant estimates of the impacts of living near Saguaro National Park West can also be attributed to the housing cycle. Interestingly, the same pattern and magnitude of impacts was found for Coronado National Forest, showing that similar impacts can be found for another type of NPS unit (National Forest). And somewhat surprisingly, there is no evidence of a similar impact of living near Saguaro National Park East (nor for Tucson Mountain Park). It is left for future research to discern why such different results are obtained for the two separate parts of Saguaro National Park.

Recall that Mueller et al. (2021) estimate an impact of proximity around \$124,000 for houses within 1 km of Saguaro National Park West, and \$160,000 for houses within 1 km of Tucson Mountain Park (\$2020). The comparable values for living within 2 km of the parks are \$48,000 and \$63,000. How does this compare to our estimates? First, the average sales price within 1 km of Tucson Mountain Park is \$566,000. An impact of \$160,000 translates to 28%. Over the 2015-2019 period, we find an average impact of 8.6% for living within 1 km of Tucson Mountain Park (p -value=0.28) but a very small negative but not significant effect for the 1-2 km distance band.

Great Smokey Mountains National Park

Great Smokey Mountains National Park (GSMNP) borders both Tennessee (TN) and North Carolina (NC). The number of transactions is much larger (358,668) and the period covered is longer (1994-2021) for TN than for NC (49,043; 2005-2021). Since there is a smaller percentage of transactions close to the park in NC versus TN, we need to include wider distance bands for the NC regression. The distance bands for Tennessee are of ≤ 1 km, 1-3 km, and 3-5km while those for North Carolina are ≤ 5 km and 5-10km. Still, the housing cycles for the areas covered in the two states are very similar (peaking in 2006 and bottoming out in 2013).

The impacts of living near GSMNP in TN and NC are given in Figure 5. On the North Carolina side, the estimated impacts of living near GSMNP using Model 1 are negative and large for both distance bands. When allowing the impact to vary by year (Model 3), the impacts for both distance bands are negative and large (approximately -20%) are significant starting in 2014 despite the small number of sales in each distance band. On average, houses that are more than 10 km from GSMNP have more bedrooms and are on larger lots. And while we control for these characteristics in the regression, it can signify other unobservables that would make units farther away more valuable than units within 10 km of GSMNP.

[Insert Figure 5 here]

On the Tennessee side, the estimated impacts of living near GSMNP using Model 1 are around -5% for all distance bands though only the impact for the 3-5km distance band is significant (see Panel B of Table 2). When allowing the impact to vary by year (Model 3), they are initially large, negative, and significant for all four distance bands but increase to small and not significant by 2006 (the peak of the housing cycle) and remain small and not significant until

2019. The impacts for the three closest distance bands turn positive and are significant in 2021. This is some evidence that MWTP has increased since COVID-19 but notice that prices also increased dramatically in the past few years near the TN side of GSMNP. In fact, the correlation between the estimates in the ≤ 1 km, 1-3 km, and 3-5 km distance bands and the price index is 0.85, 0.82, and 0.80, respectively which is evidence that the rise of this impact is due to the comparable rise in the housing market.

Robustness Checks

In this sub-section, we address some of the assumptions we have imposed on our empirical analysis. First, we run the regressions using census tract versus block group fixed effects. Second, we run the regressions without the PROT_5000 variable to see if this is picking up some of the impacts of the NPD variables. Third, we alter the specific distance bands used in each case to see if the results are robust to different specifications. Fourth, we only use block groups with no change in distance bands so that the identifying variation only comes from the establishment or redesignation of the site. Fifth, we use distance to the site rather than the distance bands and limit the distance to units within a maximum linked to the distance covered by the distance bands. We specify proximity to a site to generate a continuous measure of "distance". Assume that the maximum distance band covers through 5km Then $PROX = 5 - \text{distance}$ for distance to site ≤ 5 and 0 otherwise. Proximity is a "good" versus a "bad" (like distance) and hence is similar to the distance bands as we expect the (positive) impact to decline with distance to the site.

We limit the models to estimating a single average treatment effect to make the comparisons across the different specifications more straightforward. And because of this restriction, we carry out this exercise for the five national parks that were either established or redesignated within the

sample periods covered by the ZTRAX data. The results are given in Table A11 in the Data Appendix. The first row for each site is using the main model with block group fixed effects. The remaining rows provide results for the four alternative assumptions given above. Generally, these alternative models produce very similar results. This is encouraging as it indicates that our assumptions about the model are quite robust to alternative assumptions.

6. Conclusion

In this study, we have estimated the MWTP to live near a sample of NPS units in the U.S. The availability of the ZTRAX data has made it possible to look at a variety of national parks across the country. To do so, we addressed three concerns that must be met to provide credible estimates of the MWTP to live near national parks, 1) the bias that arises when unobserved local characteristics are correlated with the proximity to national parks, 2) when using the DID framework, the change in the treatment variable can result in a new market equilibrium, and 3) changes in the impacts of living near the park (particularly before and after a designation) can also reflect (cyclical or countercyclical) changes in the housing cycle.

As Walls (2022) points out, few studies have addressed the first two concerns, so we view this as an important contribution to the literature to have addressed these concerns in our empirical analysis. Our approach was to focus on the source of the variation in the distance to the national parks that identified the MWTP to live nearby. And to be clear how changes in the unobserved neighborhood quality could bias these impacts. We also took advantage of the scope of our data to estimate impacts for five national parks that were either established or redesignated during the sample using a DID estimator to estimate the treatment effects of designation. We

addressed the second concern above by allowing the regression coefficients (including the block group fixed effects) to vary in the pre- and post-treatment periods.

To address the third concern, we allowed the impacts to vary over time. Estimating a single (average) impact over longer time periods can appear to support positive impacts from living near a national park that mask cyclical effects in the housing market. In the GBA example, and using a placebo analysis, we showed that the apparent increases in MWTP to live near a NHS or NHP can be explained by increases in the value of neighborhood quality over the housing cycle. In other examples, we showed that changes in these impacts follow comparable changes in the local housing market cycle, so care must be taken in interpreting these as changes in the MWTP to live near a NPS unit.

To estimate the MWTP to live near national parks, we analyzed parks that were established or redesignated during the period covered by our data, parks in an urban setting, a National Park that is split in two parts geographically, and a National Park that spans two states. We view this as another important contribution to the literature to be able to estimate the MWTP to live near a national park in very different settings; from the grandeur of National Parks like the Great Smokies, to a national battlefield marking a site of the War of 1812, to very small urban parks such as the John Fitzgerald Kennedy National Historic Site and the Longfellow House-Washington's Headquarters.

In our study, we answered three questions 1) What are the MWTP to live near national parks as estimated using the hedonic method? 2) Do these estimates vary across the different types of national parks? and 3) What is the value added for areas that are established or redesignated as national parks?

We estimate the impact of living near 18 NPS units. We had no prior expectations that the MWTP estimates would be similar across these sites. We found that these estimates show considerable variation across the different types of national parks that we analyzed. Saguaro National Park consists of two units; East and West. The impacts of living near Saguaro National Park East are large and significant in some years whereas they tended to be small and not significant for Saguaro National Park West. And even for the same National Park, Great Smokey Mountains National Park, the estimated impacts are different on the TN and the NC sides of the park.

We estimate the value added for five areas that were established or redesignated as national parks. The results show significant differences across these five examples. They include positive treatment effects (PGF), no effects (BICR), delayed negative effects (CVNP) and negative effects (RRNB, TSFB).

While many of these cases appear to show significant MWTP to live near a national park, closer inspection reveals this could be due to the comparable cyclical or countercyclical movement in the housing market. This holds for Paterson Great Falls National Historical Park (cyclical), River Raisin National Battlefield Park (countercyclical), Saguaro East National Park and Coronado National (countercyclical), the parks in the Greater Boston Area (GBA, cyclical), and the TN side of the Great Smokey Mountains National Park (cyclical). In the case of GBA, we developed a placebo site procedure that indicated that the increase in the impact of living near the NPS units was likely driven by changes in the MWTP to live near neighborhoods of similar quality. Netting this out for the overall impact of living near the NPS units results in estimates that are much smaller, generally not individually significant, and do not follow any discernable pattern.

The key takeaway is to be careful about interpreting results without addressing important concerns that can contaminate the estimates. This is a cautionary tale; what appears to be significant MWTP to live near a national park can also be explained by other factors related to the housing market including unobserved neighborhood characteristics (that are correlated with the measures that capture distance to the park) and the housing cycle.

Acknowledgements

Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information can be found at <http://www.zillow.com/ztrax>. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group. Parcel data for approximately two thirds of U.S. counties was provided by Regrid through its “Data with Purpose” program (<https://regrid.com/purpose>). Christoph Nolte acknowledges support from the Department of Earth & Environment at Boston University, the Junior Faculty Fellows program of Boston University's Hariri Institute for Computing and Computational Science, the Nature Conservancy, and the National Science Foundation’s (NSF) Human-Environment and Geospatial Sciences (HEGS) program (grant #2149243)”. The National Park Service is acknowledged for providing funding for this analysis. This paper represents our views and not those of the NPS. Paige Smalley provided significant research support for this project.

References

Banzhaf, H. Spencer. 2022. “Financing Outdoor Recreation: An Introduction,” *Land Economics* 98(3): 421-427.

Benson, Charles, Philip Watson, Garth Taylor, Philip Cook, and Steve Hollenhorst. 2013. “Who Visits a National Park and What do They Get Out of It?: A Joint Visitor Cluster Analysis and Travel Cost Model for Yellowstone National Park,” *Environmental Management* 52: 917–928.

Boyle, K., R. Paterson, R. Carson, C. Leggett, B. Kanninen, J. Molenaar and J. Neumann. 2016. “Valuing Shifts in the Distribution of Visibility in National Parks and Wilderness Areas in the United States,” *Journal of Environmental Management* 173: 10-22.

Gibbons, Stephen, Susana Mourato, and Guilherme M. Resende. 2014. “The Amenity Value of English Nature: A Hedonic Price Approach,” *Environmental Resource Economics* 57:175–196

Haefele, M., J. Loomis and L. Bilmes. 2020. Total Economic Valuation of the National Park Service Units and National Park Service Cooperative Programs. Valuing U.S. National Parks and Programs. Routledge. New York.

Ham, C., J. Loomis and P. Champ. 2015. Relative Economic Values of Open Space Provided by National Forest and Military Lands Surrounding Communities. *Growth and Change*. 46(1): 81-96.

Hand, M., J. Thacher, D. McCollum and R. Berrens. 2008. Intra-Regional Amenities, Wages, and Home Prices: The Role of Forests in the Southwest. *Land Economics*. 84(4): 635-651.

Heberling, M. and J. Templeton. 2009. Estimating the Economic Value of National Parks with Count Data Models Using On-Site, Secondary Data: The Case of the Great Sand Dunes National Park and Preserve. *Environmental Management*. 43(6): 619-627.

Jakus, P. and S. Akhundjanov. 2019. The Antiquities Act, National Monuments, and the Regional Economy. *Journal of Environmental Economics and Management*. 95: 102-117.

Kuminoff, N. and J. Pope. 2014. “Do ‘Capitalization Effects’ For Public Goods Reveal the Public’s Willingness to Pay?” *International Economic Review* 55(4): 1227-1250.

Liu, X., L. Taylor, T. Hamilton and P. Grigelis. 2013. Amenity Values of Proximity to National Wildlife Refuges: An Analysis of Urban Residential Property Values. *Ecological Economics*. 94: 37-43.

Loomis, John. 2002. *Integrated public land management*. Columbia University Press: New York.

Malik, Khyati, Sowon Kim, and Brian J. Cultice. 2023. “The Impact of Remote Work on Green Space Values in Regional Housing Markets,” this issue.

Mueller, Julie M., John B. Loomis, Leslie Richardson, and Ryan A. Fitch. 2021 “Valuing impacts of proximity to Saguaro National Park on house prices,” *Applied Econ Perspectives and Policy* 43: 1–14.

Neher, C., J. Duffield and D. Patterson. 2013. Valuation of National Park System Visitation: The Efficient Use of Count Data Models, Meta-Analysis, and Secondary Visitor Survey Data. *Environmental Management*. 52: 683-698.

Neher, C., J. Duffield and D. Patterson. 2013. Valuation of National Park System Visitation: The Efficient Use of Count Data Models, Meta-Analysis, and Secondary Visitor Survey Data. *Environmental Management*. 52: 683-698.

Neumann, B., K. Boyle and K. Bell. 2009. The Effect of National Wildlife Refuge Proximity on Selling Prices of Nearby Residential Properties. *Land Use Policy*. 26: 1011-1019.

Nolte C, KJ Boyle, AM Chaudhry, et al. 2021. "Studying the Impacts of Environmental Amenities and Hazards with Nationwide Property Data: Best Data Practices for Interpretable and Reproducible Analyses," SSRN Electron J.

Parsons, G. C. Leggett, J. Herriges, K. Boyle, N. Bockstael and Z. Chen. 2020. "A Site-Portfolio Model for Multiple-Destination Recreation Trips: Valuing Trips to National Parks in the Southwestern United States," *Journal of the Association of Environmental and Resource Economists*. 8: 1-25.

Szabó Anrea and Gegely Ujhelyi. 2022. "Economic impacts of the US National Park System," Working Paper, University of Houston

Walls, Margaret. 2022. "Economics of the US National Park System: Values, Funding, and Resource Management Challenges," *The Annual Review of Resource Economics* 14:1-18.

Walls, M., P. Lee and M. Ashenfarb. 2020. National Monuments and Economic Growth in the American West. *Science Advances*. 6(12).

Weiler, S. 2006. A Park by Any Other Name: National Park Designation as a Natural Experiment in Signaling. *Journal of Urban Economics*. 60: 96-106.

Weiler, S. and A. Seidl. 2004. What's in a Name? Extracting Econometric Drivers to Assess the Impact of National Park Designation. *Journal of Regional Science*. 44(2): 245-262.

Zabel, Jeffrey. 2015. "The Hedonic Model and the Housing Cycle," *Regional Science and Urban Economics* 54: 74-86.

Table 1: National Park Service Units Included in the Study

Unit Name	Location	Unit Type	Years Covered
Adams NHP	GBA	National Historical Park	1987-2021
Boston African American NHS	GBA	National Historic Site	1987-2021
Boston NHP	GBA	National Historical Park	1987-2021
Frederick Law Olmsted NHS	GBA	National Historic Site	1987-2021
John Fitzgerald Kennedy NHS	GBA	National Historic Site	1987-2021
Longfellow House-Washington's Headquarters	GBA	National Historic Site	1987-2021
Lowell NHP	GBA	National Historical Park	1987-2021
Minute Man NHP	GBA	National Historical Park	1987-2021
Salem Maritime NHS	GBA	National Historic Site	1987-2021
Saugus Iron Works NHS	GBA	National Historic Site	1987-2021
		Established/Redesignated	
Birmingham Civil Rights National Monument	AL	Jan 13, 2017	1998-2021
Tule Springs Fossil Beds National Monument	NV	Dec 19, 2014	1994-2018
Paterson Great Falls National Historical Park	NJ	Nov 7, 2011	1994-2021
River Raisin National Battlefield Park	MI	Oct 22, 2010	1994-2021
Cuyahoga Valley National Park	OH	2000	1993-2021
Great Smokey Mountain National Park	TN, NC	June 15, 1934	TN: 1994-2021 NC: 2005-2021
Saguaro National Park	AZ	March 1, 1933	1994-2018
Coronado National Forest	AZ	April 11, 1902	1994-2018

Table 2: Estimates for Impacts of Living near a National Park

Panel A: DID Results for the Five Established/Redesignated Sites

	Distance Bands			
Birmingham Civil Rights National Monument	1-3km	3-5km	5-7km	7-10km
Estimate	4.128	20.697	8.635	4.157
p-value	(0.812)	(0.210)	(0.626)	(0.672)
Cuyahoga Valley National Park	<1km	1-2km	2-3km	3-5km
Estimate	0.224	2.712	-0.565	-1.550
p-value	(0.973)	(0.636)	(0.899)	(0.546)
Paterson Great Falls National Historical Park	<1km	1-2km	2-3km	3-5km
Estimate	-12.558	-11.245	-10.233	-6.813
p-value	(0.041)	(0.034)	(0.043)	(0.143)
River Raisin National Battlefield Park	1-3km	3-5km	5-10km	
Estimate	-10.155	8.317	6.455	
p-value	(0.686)	(0.763)	(0.816)	
Tule Springs Fossil Beds National Monument	1-3km	3-5km	5-10km	
Estimate	-9.92	-8.61	-11.25	
p-value	0.099	0.144	0.003	

Panel B: Model 1 Results for Remaining Sites

Boston National Historic Sites and Historical Parks								
	Distance Bands							
	<=1 km		1-2 km		2-3 km			
	estimate	p-value	estimate	p-value	estimate	p-value	estimate	p-value
National Historic Sites	0.039	0.252	-0.007	0.724	-0.010		0.516	
National Historical Parks	-0.003	0.028	-0.009	0.553	0.007		0.622	
Saguaro/Coronado/Tucson								
	Distance Bands							
	<=2 km		2-3 km		3-5 km			
	estimate	p-value	estimate	p-value	estimate	p-value	estimate	p-value
Saguaro East	0.113	0.003	0.059	0.102	0.022		0.480	
	Distance Bands							
	<=1 km		1-2 km		2-3 km		3-5 km	
	estimate	p-value	estimate	p-value	estimate	p-value	estimate	p-value
Saguaro West	-0.110	0.063	-0.062	0.197	-0.026	0.567	-0.006	0.870
Coronado	0.141	0.006	0.099	0.008	0.083	0.003	0.051	0.010
Tucson	-0.021	0.615	-0.033	0.426	-0.043	0.167	-0.021	0.372
Great Smokey Mountains National Park								
	Distance Bands							
	<= 1km		1-2 km		2-3 km		3-5 km	
	estimate	p-value	estimate	p-value	estimate	p-value	estimate	p-value
Tennessee	-5.535	0.249	-4.067	0.398	-6.154	0.078	-5.500	0.024

	<= 5km		5-10 km	
	estimate	p-value	estimate	p-value
North Carolina	-24.040	0.000	-12.846	0.020

Impacts for Model 1 are the percent difference in house prices for houses in each distance band compared to houses outside the farthest distance band (the estimate of δ_2 in Model 1).

Figure 1: Treatment Effects for Distance to Established/Redesignated Sites

Notes: Data Source: Zillow ZTRAX, Prices are put in real terms using CPI less shelter from corresponding Census Region. Results from a block group fixed effects regression show the percent difference in house prices for units in each distance band relative to units more than farthest distance band relative to the same difference in the year of designation. The price index is set to zero in the year of designation (dashed vertical line) to make it comparable to the treatment effects.

Figure 2: The Impact of Distance to the Nearest National Historic Site/Historical Park on House Prices for the Greater Boston Area: 1987-2021

Notes: Data Source: Zillow ZTRAX, Prices are put in real terms using cpi less shelter for Boston-Cambridge-Newton MA-NH, Jan 2020=100. Results from block group fixed effects regressions show the percent difference in house prices for units in each distance band relative to units more than 3km from the nearest national historic site/historical park.

Figure 3: Percent Difference in the Impact of Distance to the Nearest National Historic Site/Historical Park versus Placebo Site for the Greater Boston Area: 1987-2021

Note: Data Source: Zillow ZTRAX, Prices are put in real terms using cpi less shelter for Boston-Cambridge-Newton MA-NH, Jan 2020=100. Results from block group fixed effects regressions show the percent difference in the change in house prices for units in each distance" band relative to units more than 3km from the nearest national historic site/historical park versus the same for the placebo site.

Figure 4: The Impact of Distance to Four Parks near Tucson AZ: 1994-2019

Notes: Data Source: Zillow ZTRAX, Prices are put in real terms using CPI less shelter for West Census Region. Results from block group fixed effects regressions show the percent difference in house prices for units in each distance band relative to units more than farthest distance band from the park.

Figure 5: The Impact of Distance to Great Smokey Mountains National Park In Tennessee: 1994-2021 and North Carolina: 2005-2021

Notes: Data Source: Zillow ZTRAX, Prices are put in real terms using cpi less shelter for South Census Region. Results from block group fixed effects regressions show the percent difference in house prices for units in each distance band relative to units more than farthest distance band relative from the parks in Tennessee and North Carolina.

Endnotes

¹¹ National: Park, Monument, Preserve, Historical Park, Historic Site, Military Park, Battlefield, Battlefield Park, Battlefield Site, Memorial, Recreation Area, Seashore, River, Lakeshore, Reserve, Parkway, Scenic Trail, Historic Trail, and Wild and Scenic Rivers and Riverways and International Historic Site.

<https://www.nps.gov/aboutus/national-park-system.htm>

² <https://www.nps.gov/aboutus/budget.htm>

³ <https://www.nps.gov/subjects/infrastructure/maintenance-backlog.htm>

⁴ <https://www.doi.gov/ocl/nps-organic-act>

⁵ These are 'Travel to Work Areas' that are defined in the Census as "zones where at least 67% of the resident population work within the same area, and at least 67% of the employees in the areas live in the area."

⁶ PLACES Lab, Department of Earth & Environment, Boston University. Available: <https://placeslab.org/data/>

⁷ Additional information about ZTRAX/PLACES is available at: <https://placeslab.org/>

⁸ <https://webcms.pima.gov/cms/One.aspx?pageId=1539>

⁹ The extent of missing values for the control variables and the periods covered by the data differ across states depending on the data they provide to Zillow.

¹⁰ Many properties do appear more than once in the dataset. For most of these cases, the housing characteristics are the same. We do not link multiple sales of the same property (estimating a repeat sales model) so we treat them as if they are different properties and hence the property characteristics do not vary with time.

¹¹ Note that this could come from relative increases in the neighborhood quality between houses closer and farther away from the national park or from relative increases in the MWTP for higher neighborhood quality (without any change in the neighborhood quality itself).

¹² As a robustness check, we limit the data to only include block groups with no change in the distance band and this produces similar results (see Section 5E).

¹³ The VIF values are all quite low (close to 1) and the R^2 's from the regression of each open space measure on the housing and neighborhood characteristics are low.









