Shedding light on large-scale solar impacts: an analysis of property values and proximity to photovoltaics across six U.S. states

Salma Elmallah, Ben Hoen, K. Sydny Fujita,
Dana Robson, Eric Brunner

Berkeley Lab Webinar

March 22, 2023

This work was funded by the U.S. Department of Energy Solar Energy Technologies Office under Contract No. DE-AC02-05CH11231.

Meet the authors

Salma Elmallah
Graduate Student Research Assistant, Berkeley Lab

Ben Hoen
Research Scientist, Berkeley Lab

Sydny Fujita
Principal Scientific Engineering Associate, Berkeley Lab

Dana Robson
Program Manager, Berkeley Lab

Eric Brunner
Professor of Economics and Policy, University of Connecticut

Disclaimer

Copyright Notice

Shedding light on large-scale solar impacts: contents

- **Introduction**
  - Motivation
  - Study overview
  - Policy context
- **Data**
  - Overview of data sources
  - Characterizing our sample by land use, urbanicity, and size
- **Methods**
  - Introducing the difference-in-difference approach
  - Introducing the model and robustness checks
- **Results**
  - Average effect of LSPVP construction on home prices
  - Average effect of LSPVP construction on home prices by state, land use, and urbanicity
- **Takeaways, limitations, and future work**

ELECTRICITY MARKETS & POLICY

Introduction

Motivation

- **The local economic impact of large-scale solar projects are broad and often underexamined**
  - They include job creation, tax revenue, generation revenue in cases of cooperative ownership, and home price impacts
- **Many states rely on solar to meet climate goals, but meet increasing permitting constraints**
  - These constraints may exacerbate concerns about property value impacts among other local economic impacts
- **Recent research provides mixed evidence on the property value impacts of solar**
  - Studies based in the U.S., specifically, MA and RI (Gaur and Lang, 2020) and NC (Abashidze, 2019), and the Netherlands (Dröes and Koster, 2021), find a statistically significant negative effect for homes near solar projects compared to homes further away
  - One study, based in the UK, finds no statistically significant effect of LSPVP proximity on home property values (Jarvis, 2021)

Study overview

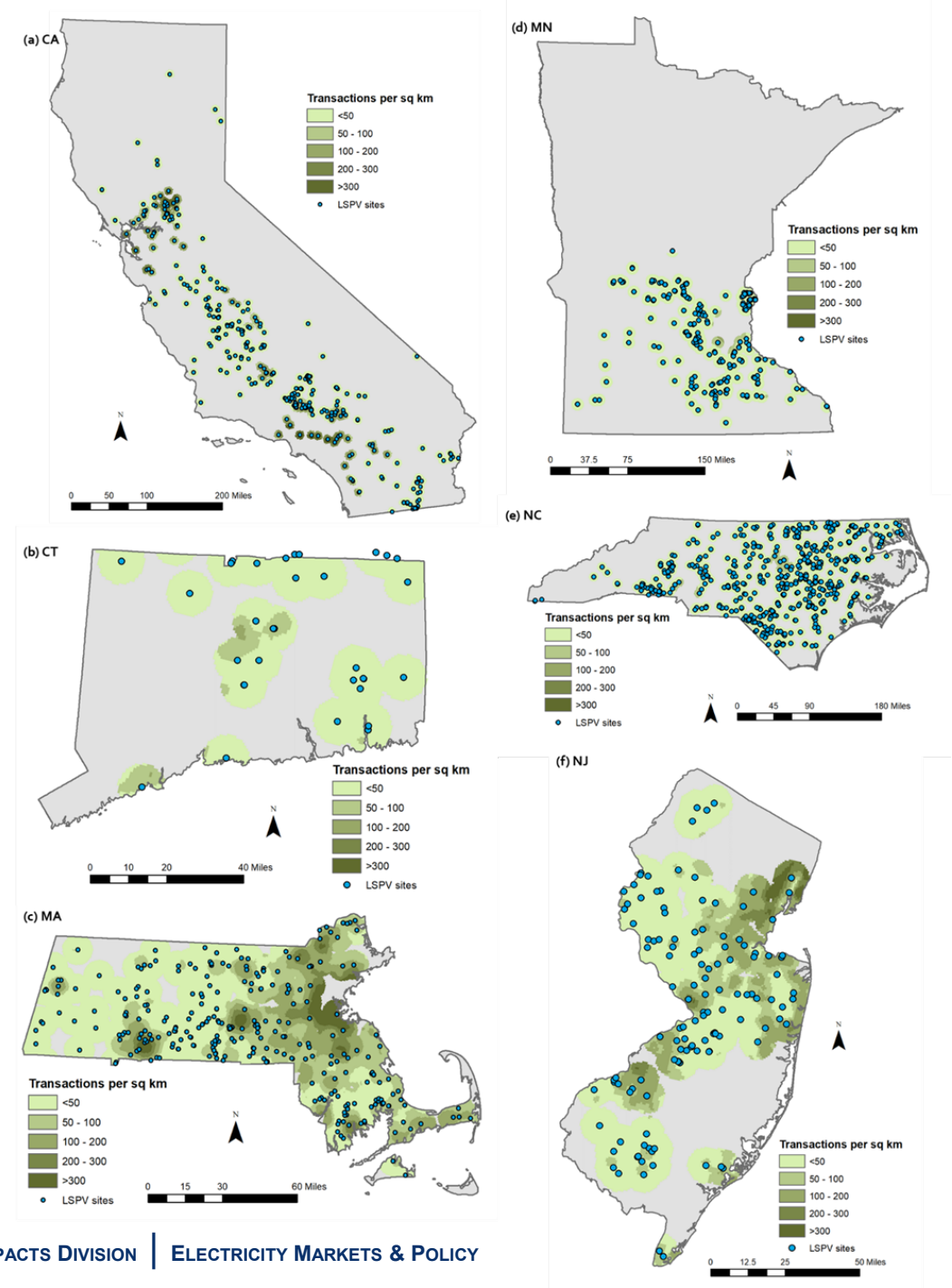
Objective: answer two related research questions: (1) what effect, if any, do LSPVPs have on residential home prices and (2) does the effect of LSPVPs on home prices differ based on the prior land use on which a LSPVP is located, the size of the LSPVP, or the urbanicity of a home's location?
Here, **LSPVP**, or large-scale photovoltaic project, is defined as ground-mounted photovoltaic generation facilities with at least 1 MW of direct current generation capacity.

Approach: we compile a unique dataset that includes home transactions and LSPVP footprints, and use a difference-in-difference approach to assess the impact of LSPVP construction on home prices for our entire sample, by state, by prior land use, by LSPVP size, and by urbanicity.

Scope: we focus on 6 states (CA, CT, MA, MN, NC, NJ) that collectively account for over 50% of the installed MW capacity of LSPVP in the U.S. and are largely understudied with respect to property value impacts of solar.

Policy context in the study area

- **LSPVP development context:** Solar is key to meeting renewable portfolio standard and GHG reduction goals in all 6 states
- **Land use context:**
  - Solar installations convert land from a previous use, and are often the only use possible once the installation occurs (agrivoltaics and some brownfield installations are exceptions).
  - Many states are encouraging solar to be sited in urban or semi-urban areas and therefore near homes
- **Community concerns**: Because a home is often the most valuable asset neighbors own, concerns about property value impacts of abbutting and neigboring properties are often a top concern for communities considering hosting a LSPVP

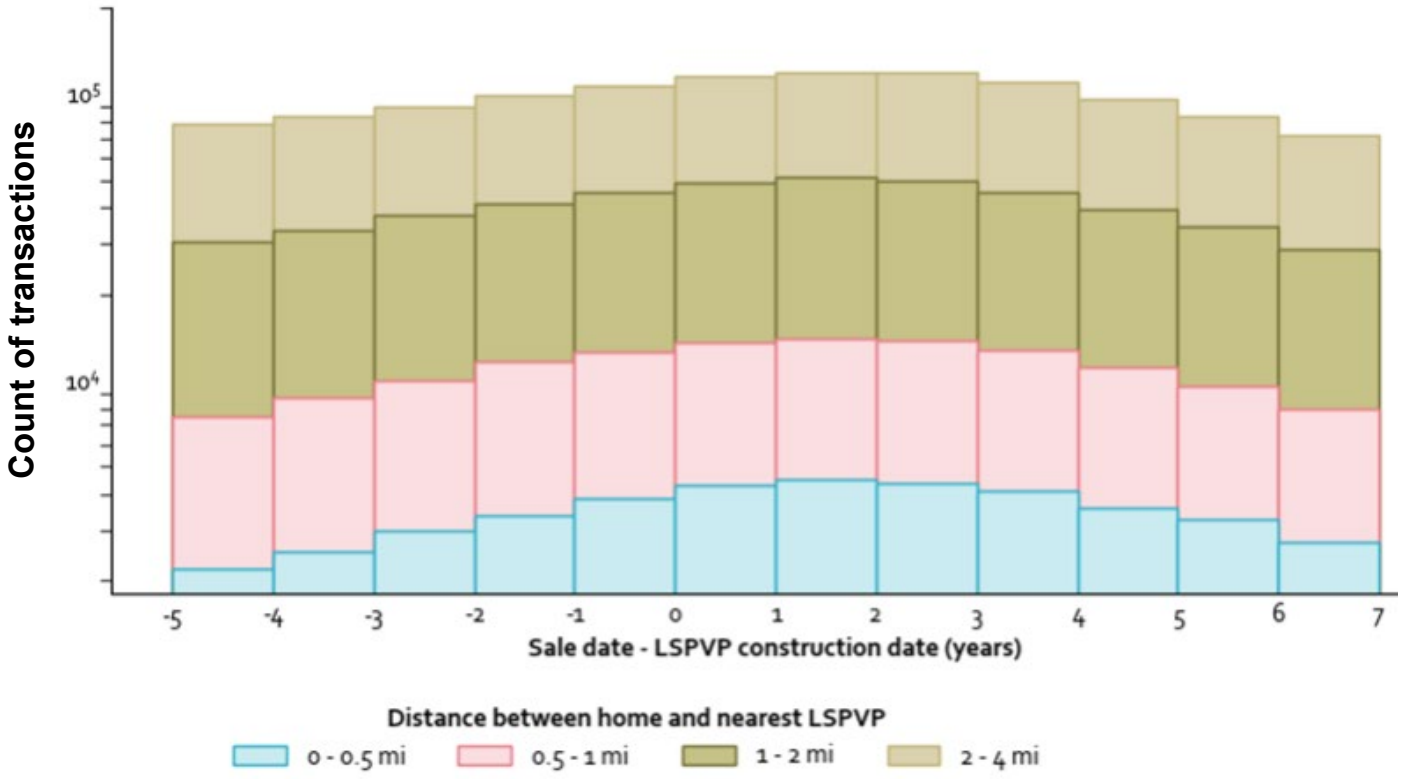


ELECTRICITY MARKETS & POLICY

Data, geographical coverage, and sample period

- **Data:** we combined the following five data sources for this analysis
  - **Polygon footprints** of LSPVPs along with their installed capacity, operation start date, and area; polygons were drawn based on validated LSPVP point location data from EIA Form 860 and satellite imagery
  - **Home transaction data** from CoreLogic that includes sales prices and property characteristics
  - **Land cover data** from the USGS National Land Cover Database, which was used to determine land use prior to LSPVP development
  - **Urban, rural, or urban cluster designations** from the US Census Bureau
  - **Amenity and disamenity (A/D) data:** several landscape characteristics that could positively or negatively impact the price of a home (e.g. flood zone designation, proximity to a landfill)
- **Geographical coverage**: we only included
  - LSPVPs and properties in the 6 study states: CA, CT, MA, MN, NC, NJ
  - LSPVPs with an installed capacity over 1 MW, and eliminated rooftop installations
  - Homes within 4 miles of an LSPVP
- **Sample period**
  - CoreLogic data cover Jan. 2004 through Sept. 2020; solar installations through Dec. 2020
  - We consider only homes that transact within 6 years of LSPVP construction, assuming that construction begins 1 year before operation

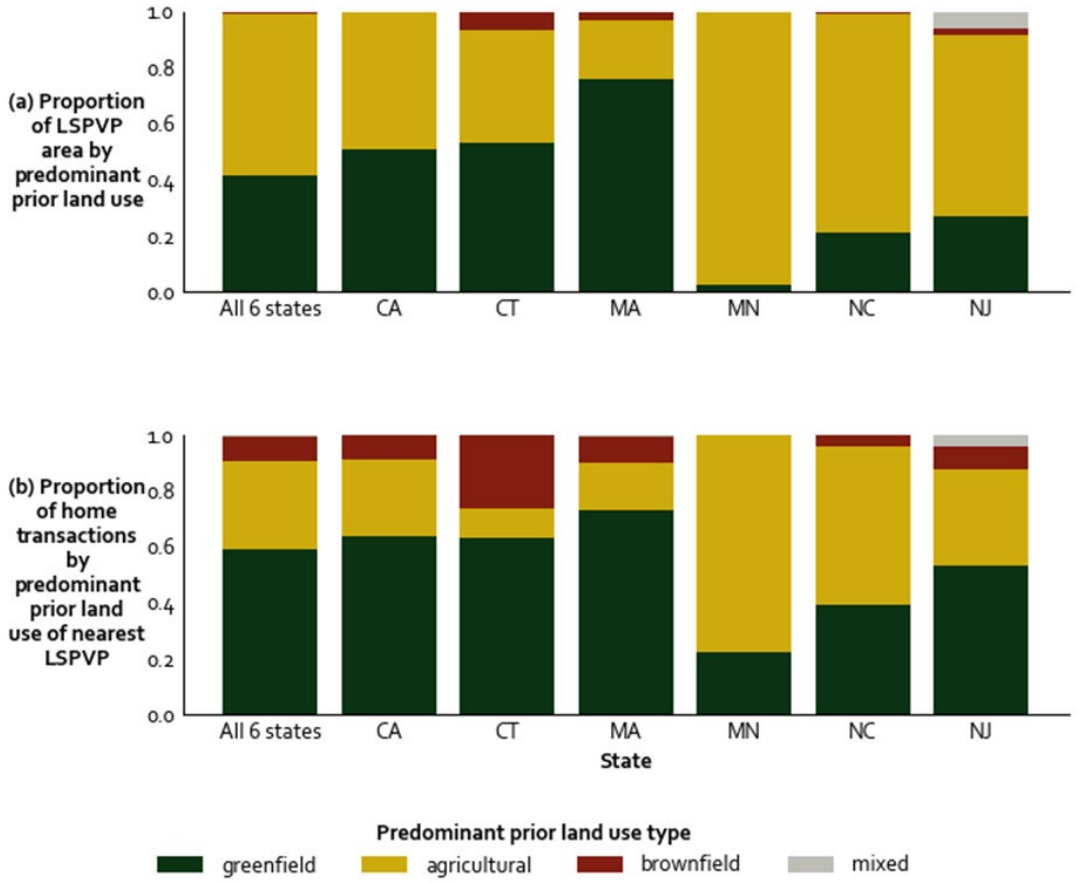
Analytic dataset: 1,836,053 transactions near 1,522 different LSPVPs are distributed among distance bins and years since construction

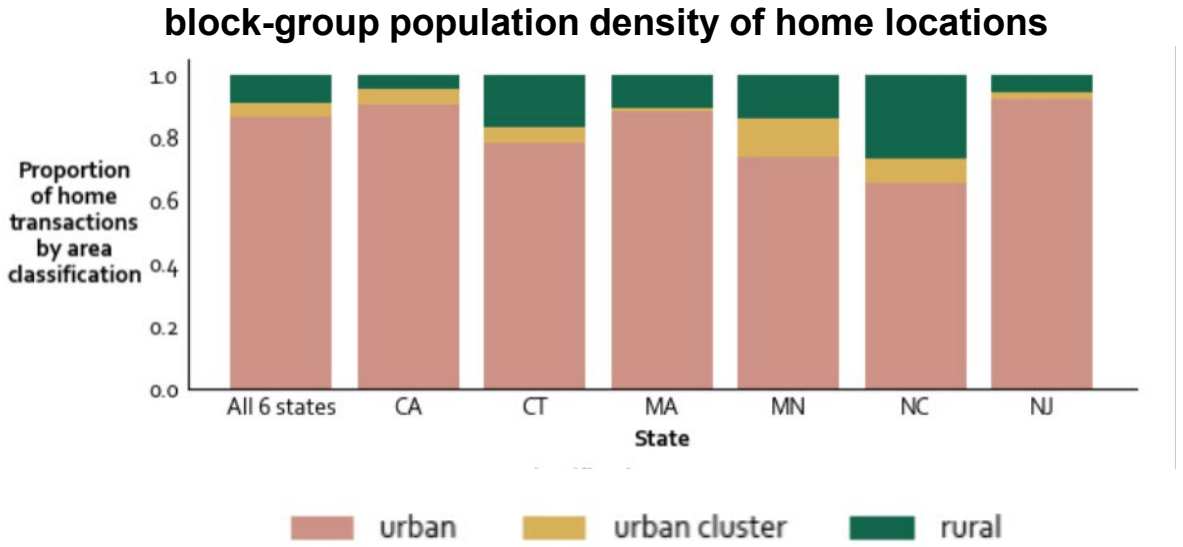


	Number of Transactions		
Distance btw. home and LSPVP	**pre-LSPVP construction**	**post-LSPVP construction**	**Total**
0 - 0.5 mi	15,141	26,989	42,130
0.5 - 1 mi	50,566	83,014	133,580
1 - 2 mi	188,748	297,832	486,580
2 - 4 mi	459,154	714,447	1,173,601
Total	**713,609**	**1,122,282**	**1,835,891**

Note: a table with the counts shown above is in the Appendix

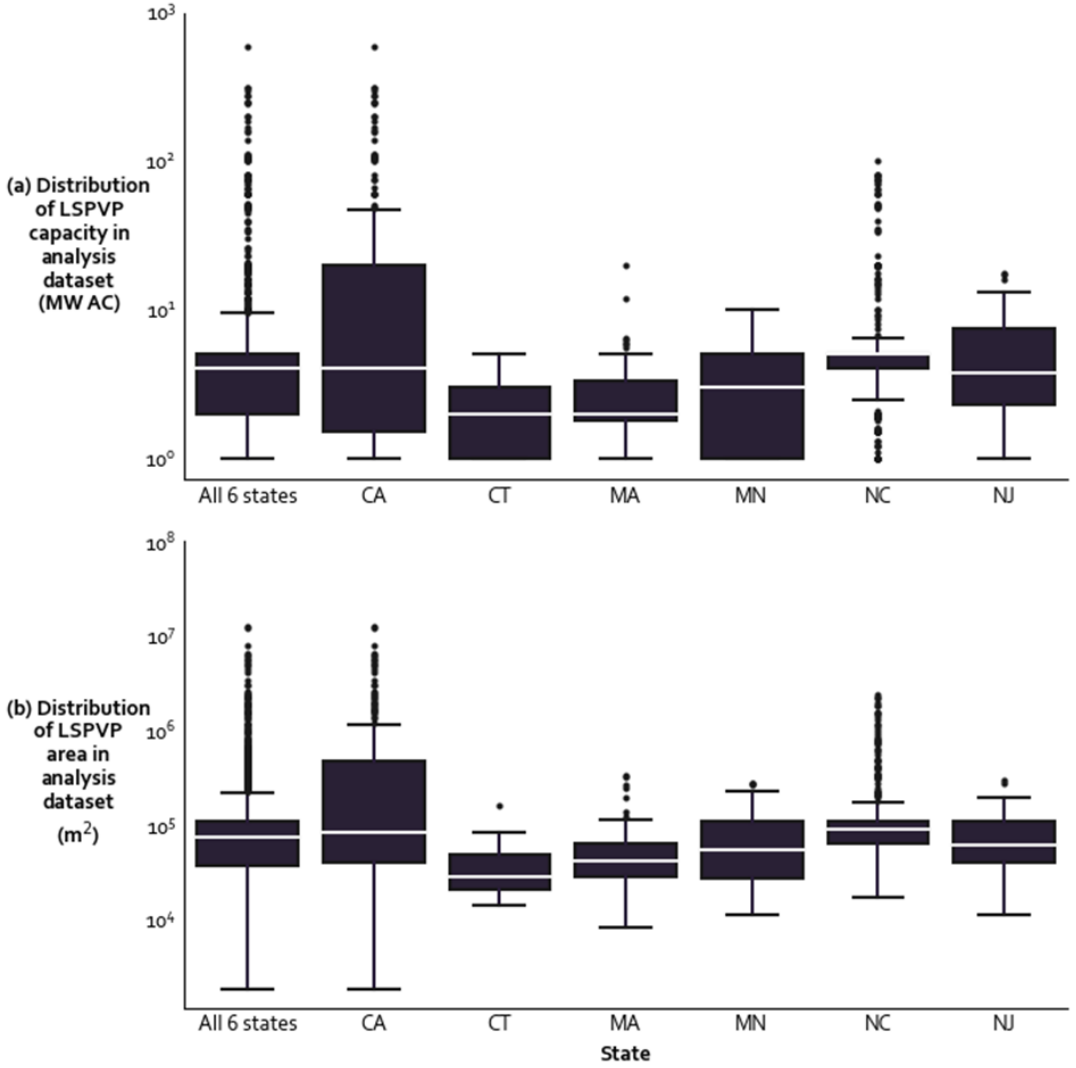
Analytic dataset: individual states vary with respect to prior land use of LSPVP and urbanicity of homes near LSPVP. Most projects are in urban or semi-urban locations





urban: >50,000 pop/mi^2;
urban cluster: >2,500 pop/mi^2 & <50,000 pop/mi^2;
rural: <2,500 pop/mi^2

Analytic dataset: we see less variation between states with respect to installed capacity and area of LSPVPs and all projects are relatively small compared to the full universe of projects in the 6 states and all of the US



All 6 States Distributions		
	LSPVP area (acres)	LSPVP capacity (MW AC)
minimum	0.5	1
5th percentile	3.3	1
10th percentile	4.4	1
25th percentile	6.1	1.5
50th percentile	11.7	2.5
75th percentile	20.3	4.5
90th percentile	35.8	8
95th percentile	77.4	18
maximum	3,140	585.9
mean	20.7	4.4
std. dev.	41.2	7.6

Methods

The difference-in-difference approach

- **The approach:** we use a **difference-in-difference (DiD) estimation** to understand how the construction of a LSPVP impacts sales prices.
  - A DiD estimation allows us to assess how both **proximity to an LSPVP** and **construction of an LSPVP** impact sales price.

- **The model:** we construct a model that relates a **home's sales price** to its distance from an LSPVP and whether it transacted before or after the LSPVP was constructed.

- **Additional controls:** We also control for location and property characteristic variables as well as inflation and deflation in the market.

- **Each project is the unit of analysis:** We compare price differences around the same project of homes near and far from the LSPVP

- **More details:** The detailed model specifications can be found in the appendix.

The difference-in-difference approach

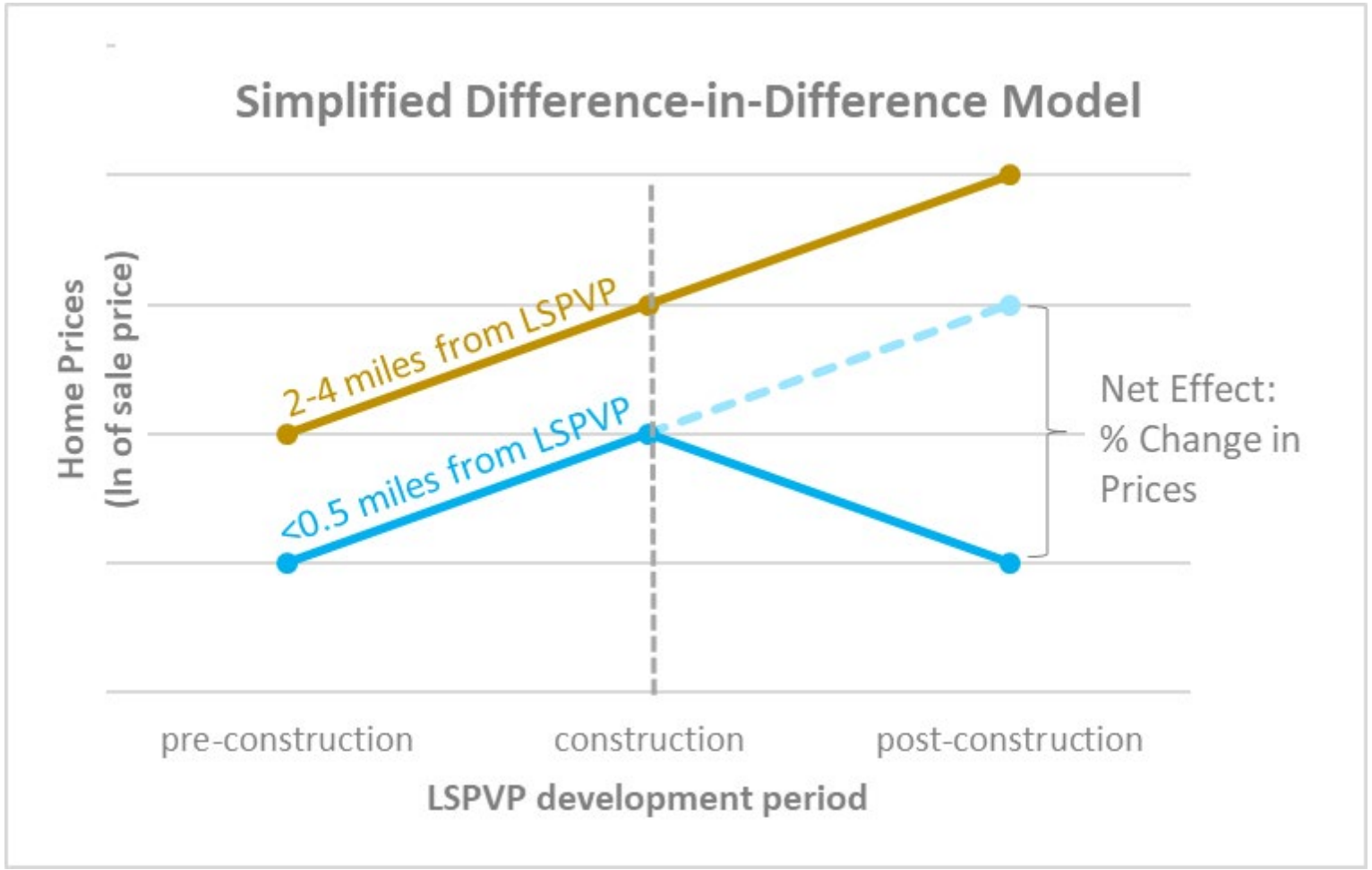
Interpreting the results:

the **% change in price** reported in our results is the

average of the difference in sales price before and after construction for homes close to LSPVP (0-0.5 mi, 0.5-1 mi, or 1-2 mi away)

compared to the…

average of the difference in sales price before and after construction for homes further away (2-4 mi away).

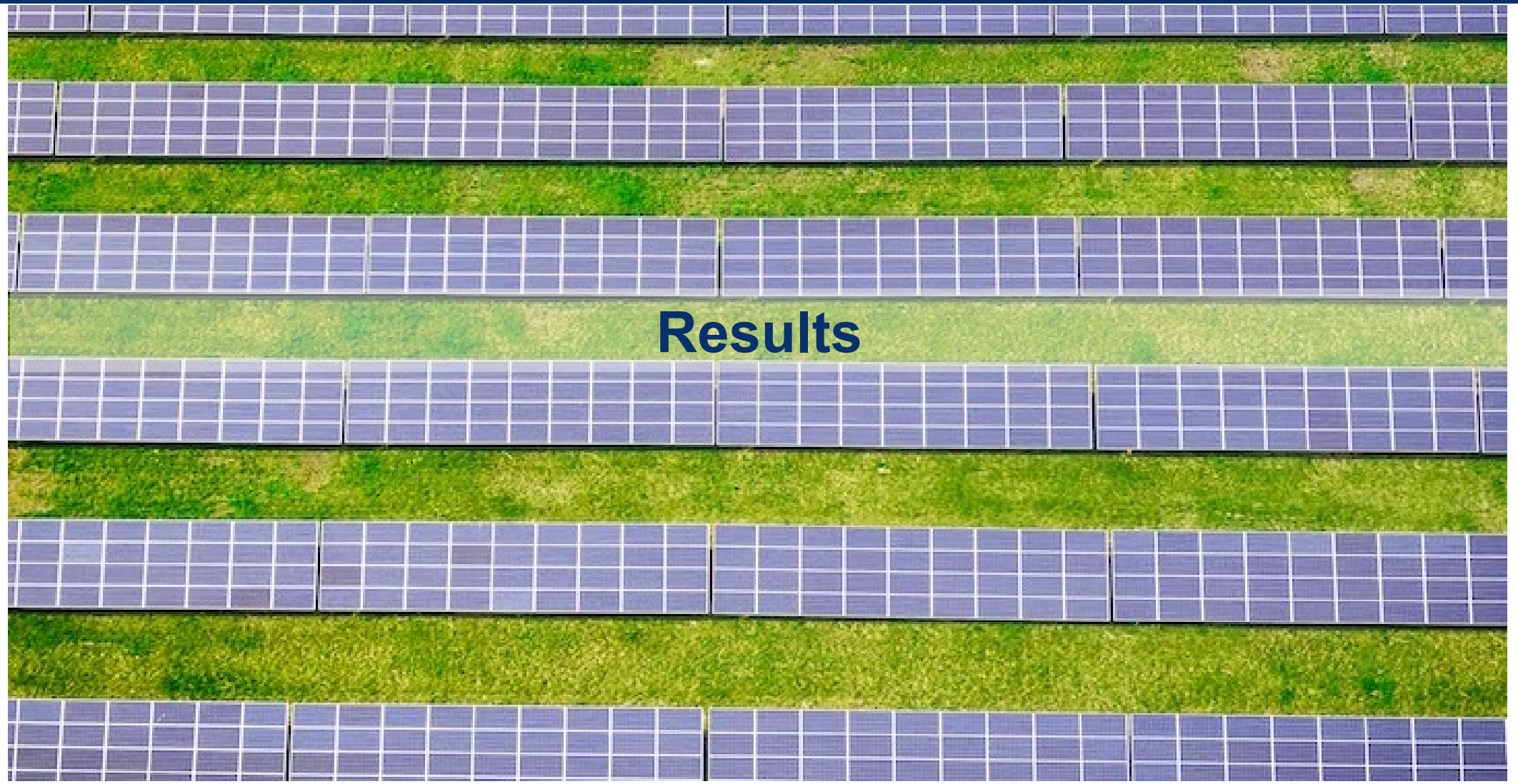


Models estimated

We constructed the following models:

- **Base model:** uses data from all 6 states, comparing price differences for homes close to LSPVP (0-0.5 mi, 0.5-1 mi, 1-2 mi) to homes further away (2-4 mi)
- **Robustness checks:** these are changes to our base model specifications that ensure that our estimates do not change significantly based on modeling choices
  - We add a 0-0.25 mi distance bin to understand if there are effects even closer to LSPVPs
  - We introduce quarter and quarter-by-year control variables to see if a more granular trending of home values impacts our estimates
  - We add proximity to amenities and disamenities (e.g. landfills, coal plants, parks) to account for any potential correlation between home prices and other landscape characteristics
- **Event study:** we compare home prices close to and further from LSPVP for individual years before or after construction to ensure that home prices near LSPVP sites were not already significantly different from home prices further away before the LSPVP was constructed
- **Heterogeneity models:** Estimate separate base models for sub-sample cohorts
  - **by state:** for each state (CA, CT, MA, MN, NC and NJ)
  - **by land use:** prior land use of each LSPVP (agricultural, greenfield, brownfield, mixed)
  - **by urbanicity:** urban, urban cluster, or rural region home locations
  - **by LSPVP area:** for LSPVPs larger or smaller than the sample-wide LSPVP median area (of 12 acres)

ELECTRICITY MARKETS & POLICY



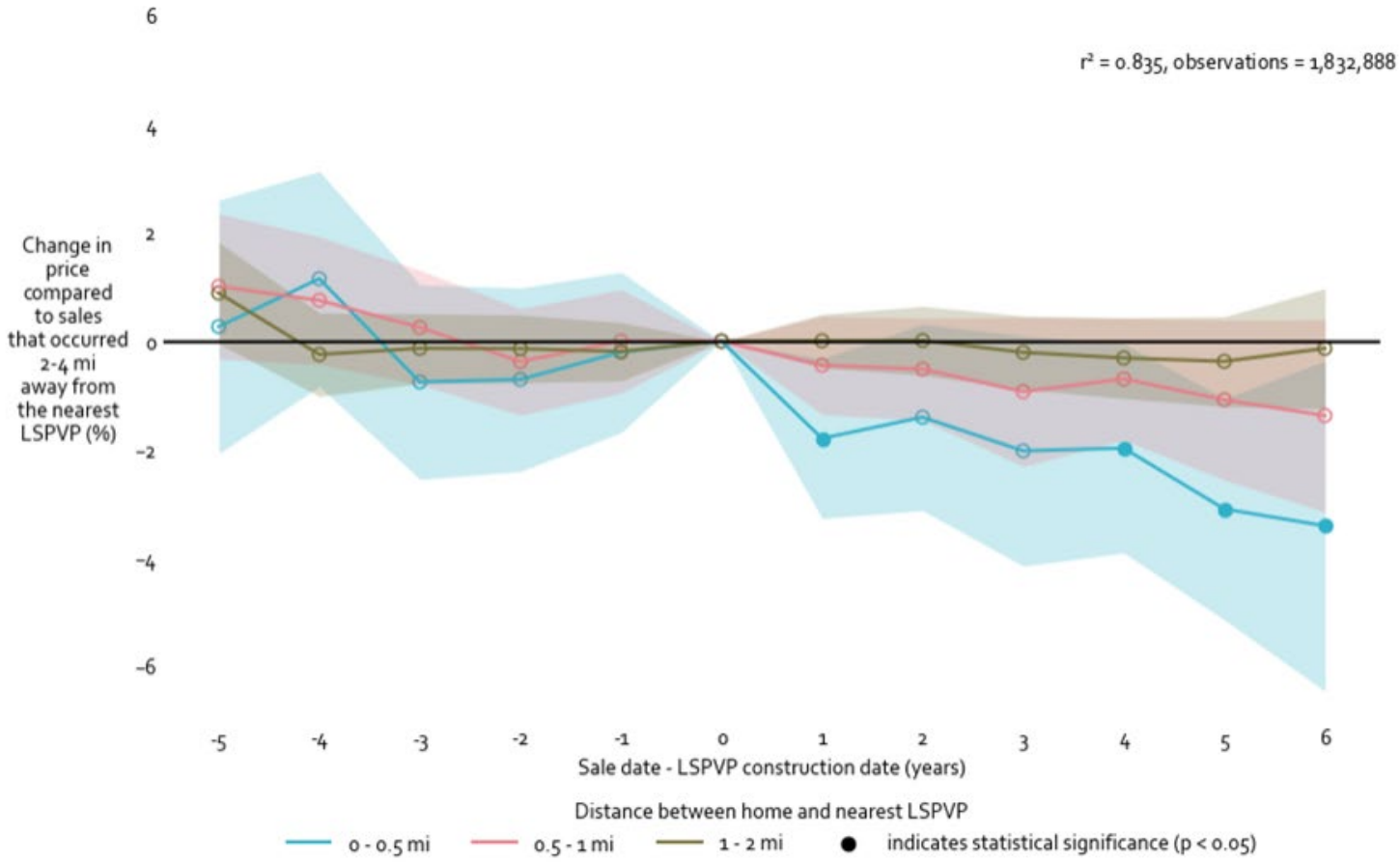
Base model shows an average 1.5% reduction in house prices for homes within 0.5 miles of a LSPVP that transacted post-LSPVP construction

Distance between home and LSPVP	% difference in house prices compared to homes 2-4 mi away
0-0.5 mi	**-1.54%** (0.63%)**
0.5-1 mi	**-0.82%** (0.413%)**
1-2 mi	-0.0841% (0.226%)
Observations	1,832,888
R^2	0.835

- **Bolded estimates**, at 0-0.5 mi and 0.5-1 mi, are statistically significant at the 5 percent level or better
- We additionally find an average 2.3% reduction in home prices within 0.25 mi of a LSPVP (full results in Appendix)
- In both models, the estimated treatment effects for homes located 1 to 2 miles from a LSPVP are quite small in magnitude and statistically insignificant, suggesting that the impact of LSPVPs on home values fades relatively quickly with distance from a LSPVP
- As expected, effects are monotonically ordered from closest distances to further away
- Robustness checks (full results in Appendix) provide largely consistent coefficients with base model

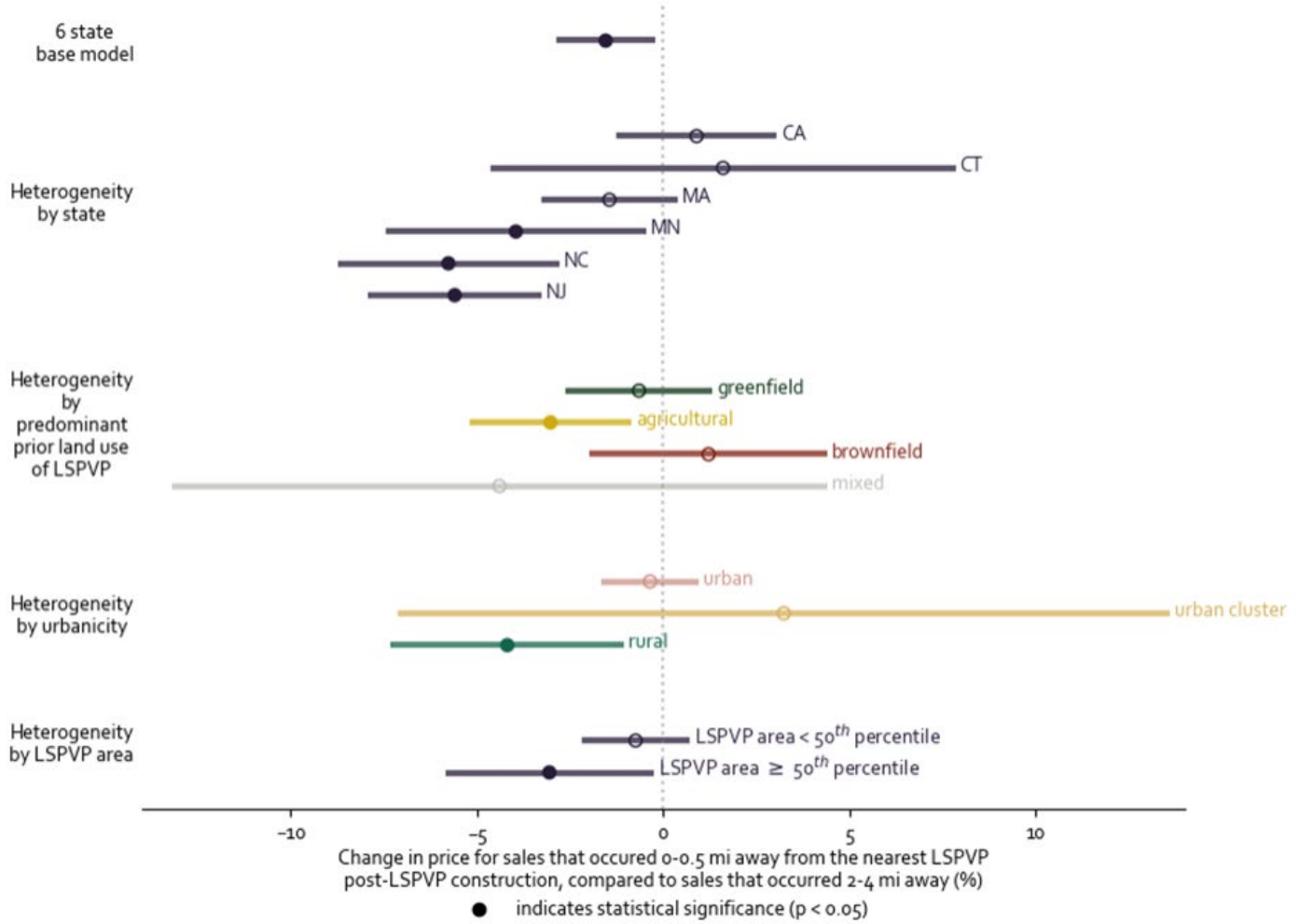
*Note: dependent variable is the logarithm of house prices. Standard errors are clustered at the project cohort level and are in parentheses. Significance levels: *** p< 0.01, ** p<0.05, * p<0.1*

Event study results show no evidence of differential trending pre-LSPVP construction for homes closer to LSPVPs vs homes further away



Average effect of proximity to LSPVP by year of sale relative to year of LSPVP construction; shaded area represents 95% confidence interval; x-axis label represents lower bound of year range (e.g. -5 refers to all transactions that occurred [-5, -4) years before the construction date of the nearest LSPVP)

Heterogeneity analysis results show that home price impacts are only measurable in certain states, where prior land use is agricultural, for homes in rural areas, and for larger LSPVPs



Results from base model as well as each heterogeneity analysis, showing average effect of LSPVP construction and proximity for homes 0-0.5 mi away from nearest LSPVP. Range of change in price represents the 95th percent confidence interval.

Takeaways, limitations, and future work

Key conclusions

- **The effects of large-scale solar projects on home sale prices clearly exist in our data**
  - We observe reductions in sales prices within 0.5 miles of an LSPVP in MN, NC, and NJ, and also between 0.5 and 1 mile in those states.
- **But those results are not consistent across all states**
  - We do not see reductions in sales prices within 0.5 miles of an LSPVP in CA, CT, or MA nor for sales prices beyond 0.5 miles
- **We see variation in results depending on how the data are divided, but mostly only for MN, NC and NJ**
  - **Project size:**
    - Across the full dataset (all 6 states) only larger projects (>12 acres) appear to be correlated with a loss in house prices within 0.5 miles (compared to 2-4 miles away)
    - But this analysis only applies to relatively small projects (90% are <35 acres/8 MW), so “large” is relative to the median of 12 acres.
    - And when focused only on MN, NC and NJ as a group, smaller projects are also correlated with a loss.*
    - No effect is found the group of CA, CT, and MA for either size.*

** not shown in the journal paper*

Key conclusions (Continued)

- **We see variation in results depending on how the data are divided, but mostly only for MN, NC and NJ**
  - **Prior Land Use:**
    - Across the full dataset (all 6 states) projects built on previous agricultural land are correlated with a loss in house prices within 0.5 miles (compared to 2-4 miles away)
    - And when focused on just MN, NC and NJ as a group, projects built on greenfield locations are also correlated with a loss.*
    - No effect is found for the group of CA, CT, and MA for any land type.*
  - **Urbanicity:**
    - Across the full dataset (all 6 states) homes in urban block groups are correlated with a loss in house prices (-4%) within 0.5 miles (compared to 2-4 miles away)
    - But almost all “rural” locations in our data are abutting urban or urban fringe areas.
    - And when focused on just MN, NC and NJ as a group, that “rural” difference in house price is larger (-7%) than it is across the full sample.*
    - No effect is found for the group of CA, CT, and MA for any population density category.*

** not shown in the journal paper*

Limitations

- **This analysis gives us a sense for the "what" but does not tell us the "why"**
  - We have identified correlations but do not know the causations of property value impacts
- **Our dataset is the largest assembled to-date but is not comprehensive in geographic or temporal scope**
  - Projects built between 2007 and 2020 are included; 80% were constructed between 2011 and 2018 (p10-p90). Our results might not apply to the most recent or future projects.
  - Only 6 states are included; therefore the results would not necessarily apply outside the sample area
- **The dataset is centered on relatively small projects in relatively urban areas**
  - Because we are focused on projects near homes that have sold, we concentrate our analysis on projects near urban areas that tend to be smaller in terms of acres occupied and MW.
  - Our results should not be applied to larger projects, e.g., those > 18 MW (p95), and, of course projects built far from homes.
- **Site design and neighboring community attitudes are not considered in this analysis**
  - Our study did not consider site design, setbacks or landscaping features, or the attitudes or sentiment of neighboring communities, home sellers, or buyers
- **We consider only one aspect of the economic impacts of LSPVPs: property values**
  - Energy development also shapes local tax revenue and employment, which have consistently been found to result in positive benefits (Brunner et al., 2021; Brunner and Schwegman, 2022a, 2022b)

Related and Possible Future Work

- **The limitations suggest two major areas for future work:**
  - **More research attention is needed on the economic impacts of LSPVPs, broadly understood to encompass dimensions such as tax revenue, ownership structures, or employment**. Added research on the local economic impacts of LSPVPs can position our findings on the average adverse impact of LSPVP development on home prices in a broader context of economic benefits and burdens due to LSPVP development
  - **More research is needed to understand the heterogeneity that we observe with respect to larger, agricultural, and rural LSPVPs in the MN, NJ and NC contexts**. Here, surveys, qualitative research, mixed-methods, and case study-based approaches may indicate how neighbors of LSPVPs engage differently with their nearby solar installation based on its size, land use, or the urbanicity of their home
- **A number of complementary DOE-funded research efforts are already ongoing:**
  - Survey of solar project neighbors
  - Developing resources for local governments that are considering hosting a LSPVP project
  - Estimating employment and income effects of LSPVP
  - Advancing innovative siting solutions such as co-located with agriculture
  - Studying decision-making processes and flows in communities hosting LSPVP
  - Creating GIS polygons for all ground-mounted > 1 MW_{DC} projects in the US

ELECTRICITY MARKETS & POLICY

Contacts

Ben Hoen: bhoen@lbl.gov; **Salma Elmallah**: salmae@lbl.gov

For more information

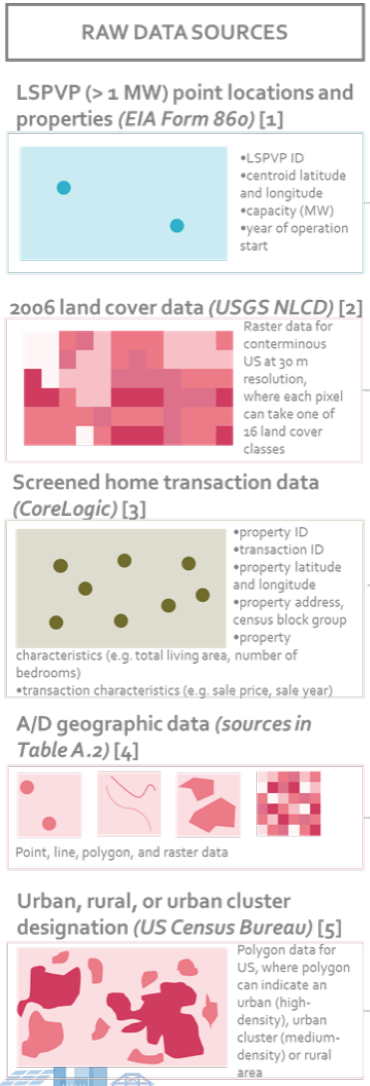
Report landing page: https://emp.lbl.gov/publications/shedding-light-large-scale-solar
Download publications from the Electricity Markets & Policy: https://emp.lbl.gov/publications
Sign up for our email list: https://emp.lbl.gov/mailing-list
Follow the Electricity Markets & Policy on Twitter: @BerkeleyLabEMP

Acknowledgements

This work was funded by the U.S. Department of Energy Solar Energy Technologies Office. under Contract No. DE-AC02-05CH11231. We thank several LBNL colleagues including Joseph Rand and Grace Brittan for collaborating on code writing and data cleaning, Mark Bolinger and Greta Bolinger for assistance with PV polygon drawing, Terry Chan for his information technology assistance and expertise, Naïm Darghouth for assistance with geocoding validation, and Sydney Forrester for help identifying datasets. We are grateful to Corey Lang, Ryan Wiser, and David Schwegman for their constructive comments on earlier drafts of this work. We are especially grateful to the U.S. Department of Energy (DOE) Solar Energy Technologies Office, specifically Ammar Qusaibaty, Dave Rench-McCauley, and Michele Boyd for their patient support of this work as it slowly made its way from conception to the final analysis with many iterations in between. This work is funded under DOE Contract No. DE-AC02-05CH11231. We are also grateful to the two anonymous reviewers, whose suggestions strengthened this paper. For our figures, we used color palettes from the collection “Wes Anderson Palettes” developed by Karthik Ram (https://github.com/karthik/wesanderson).

Appendices

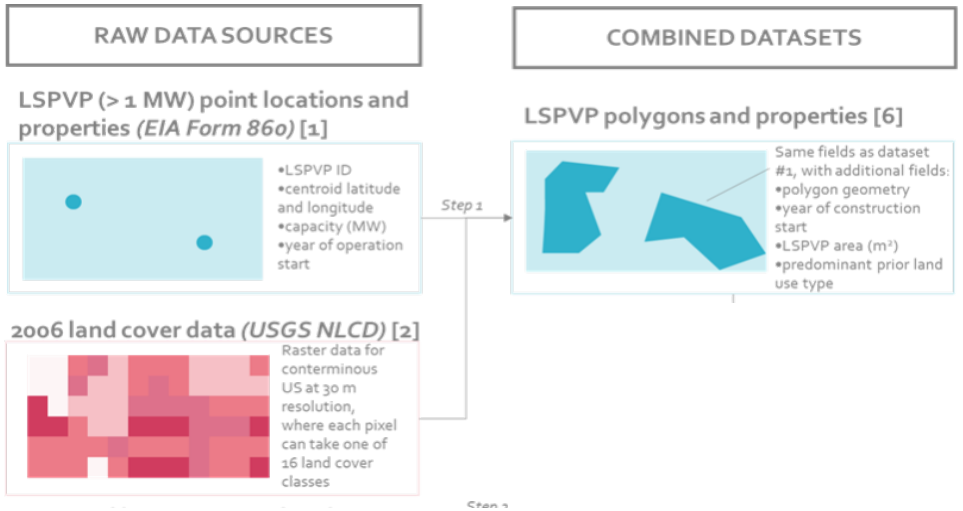
Overview of data sources and dataset preparation

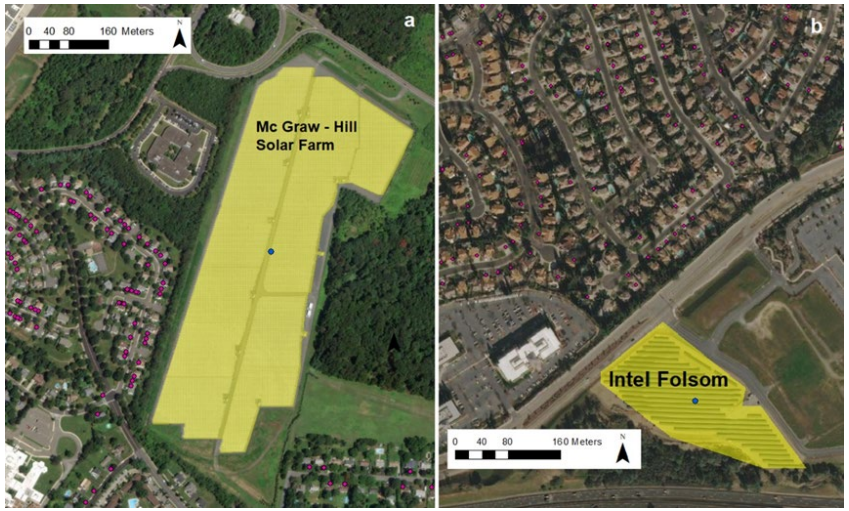


Key data sources:

- LSPVP point locations and properties (EIA Form 860): latitude-longitude data on solar plants, their installed capacities (in megawatts, MW), and their operation start date; we kept only solar plants within the study area with an installed capacity over 1 MW, and eliminated rooftop installations
- Land cover data (USGS National Land Cover Database): used to determine land use prior to LSPVP development
- Screened home transaction data: CoreLogic data (provided with a non-disclosure agreement) on home transactions and property characteristics; filtered for relevant and complete records
- Amenity and disamenity (A/D) data: several landscape characteristics that could positively or negatively impact the price of a home (e.g. flood zone designation, proximity to a landfill)
- Urban, rural, or urban cluster designation (US Census Bureau): a metric based on population density, where urban areas are the most dense, followed by urban clusters, then rural areas, used to determine the urbanicity of a home

Overview of data sources and dataset preparation cont'd

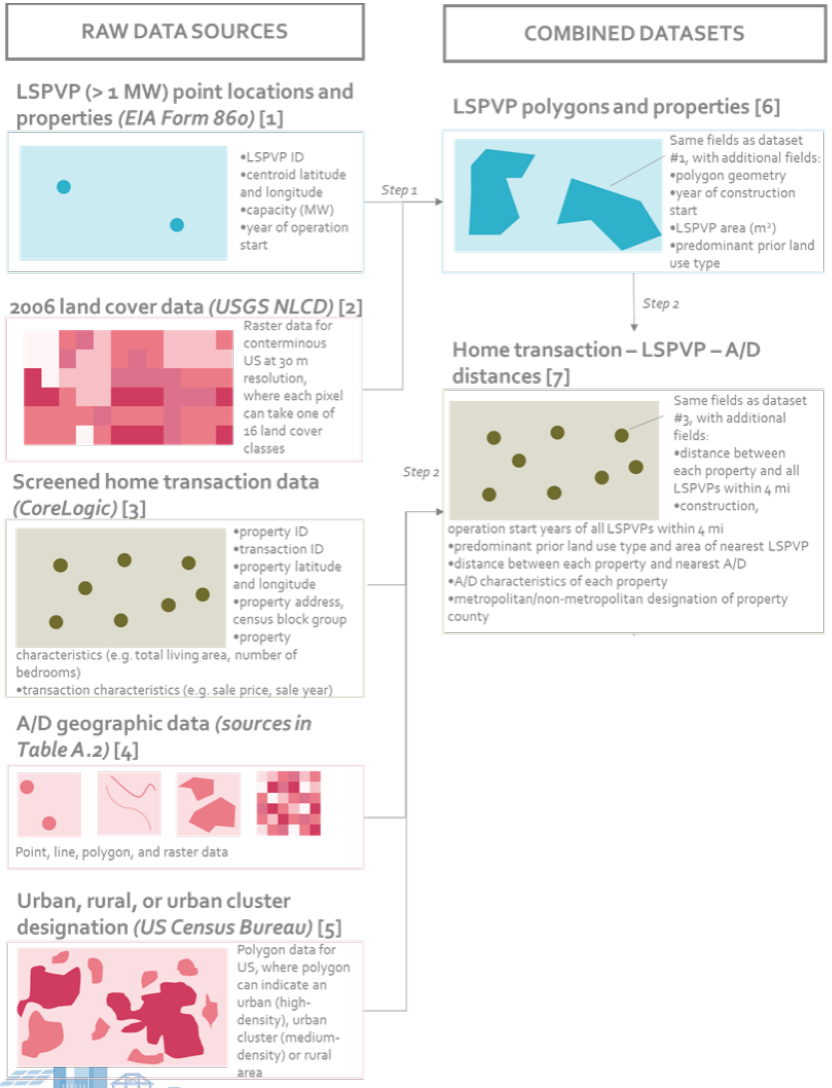




Determining LSPVP footprints and characterizing land use and construction start:

- Verify point locations of LSPVPs in EIA data, using satellite imagery; revise project centroids when necessary
- Manually draw polygons around LSPVP boundaries based on satellite imagery
- Calculate a construction start date for each LSPVP (assume that construction start date = operation start date - 1 year)
- Determine predominant prior land use of each LSPVP: agricultural, brownfield, greenfield, or mixed

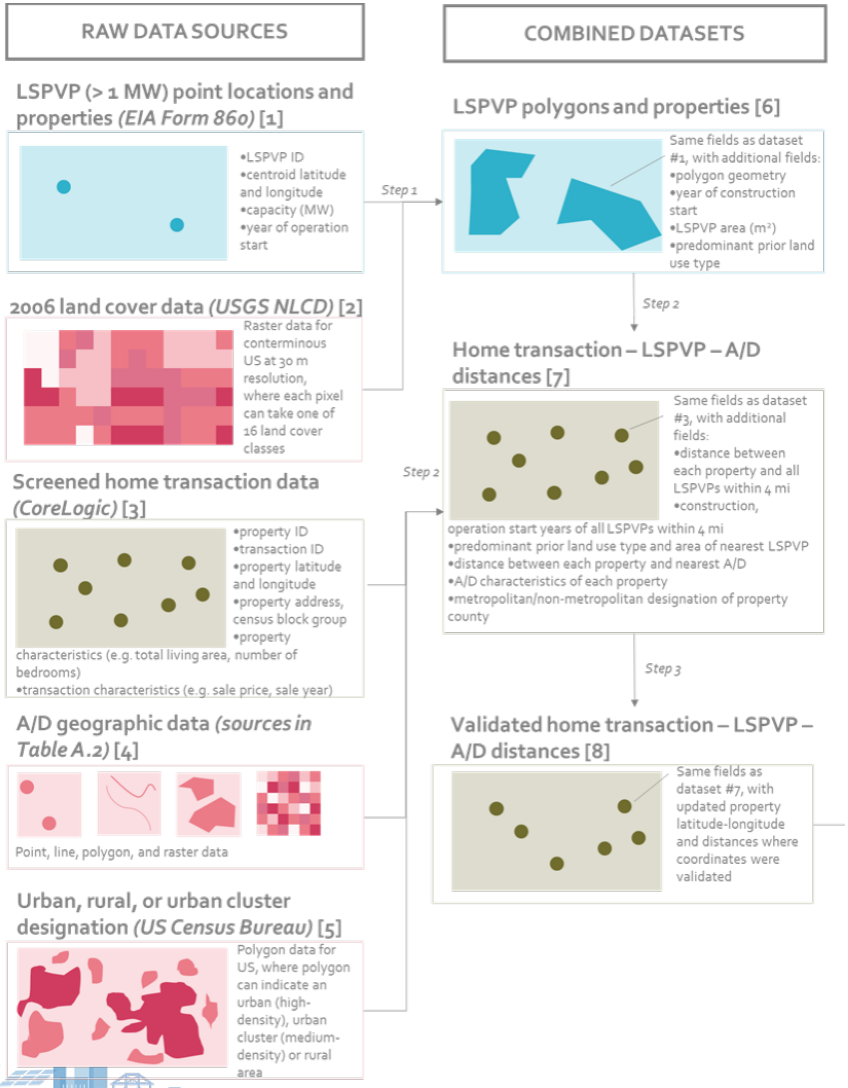
Overview of data sources and dataset preparation cont'd



Linking homes, home and region characteristics, and LSPVPs:

- For each home in the screened CoreLogic dataset, calculate the geodesic distance to the polygon boundary of the nearest LSPVP and separately to all A/D locations
- Determine underlying A/D characteristics, where appropriate, such as flood zone status and road/airport sound levels
- Determine the urbanicity of each home's location

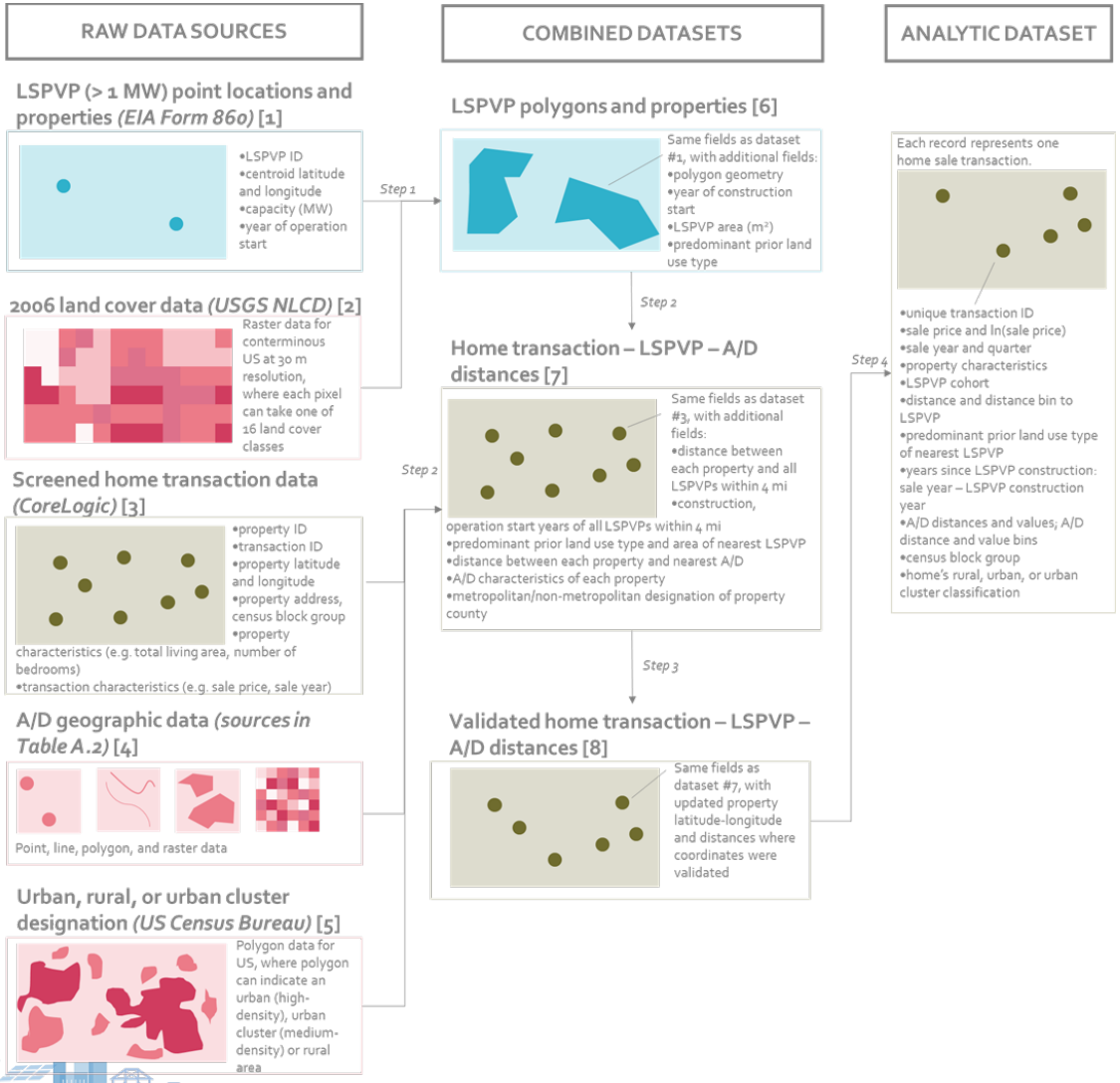
Overview of data sources and dataset preparation cont'd



Validating home transactions based on proximity to LSPVPs:

- Validate the coordinates of select homes sited near LSPVPs or A/Ds using the Google Geocoding API
- Where Google Geocoding returns a high precision indicator for coordinates, replace CoreLogic coordinates with Google coordinates
- Otherwise, drop any home transactions where CoreLogic and Google Geocoding reported inconsistent coordinates

Overview of data sources and dataset preparation cont'd



Assembling the analytic dataset:

- Retain only the home transactions that are suitable for use in the final analysis by eliminating (1) properties that host a LSPVP (i.e. their coordinates fall within the boundaries of a LSPVP polygon), (2) properties that are over four miles away from a LSPVP, and (3) properties that transacted over 6 years before or after the operation start date of a LSPVP
- Calculate three sets of key values used in the analysis: the transaction's project cohort, LSPVP distance bin, and years since LSPVP construction

Analytic dataset: 1.8 million transactions near 1,522 different LSPVPs are distributed among distance bins and years since construction

Sale date-LSPVP construction date (years)	Number of transactions 0-0.25 mi from LSPVP	Number of transactions 0.25-0.5 mi from LSPVP	Number of transactions 0.5-1 mi from LSPVP	Number of transactions 1-2 mi from LSPVP	Number of transactions 2-4 mi from LSPVP	Total
-5 to -4	481	1,723	7,521	30,517	79,074	119,316
-4 to -3	564	1,973	8,636	33,296	83,735	128,204
-3 to -2	681	2,381	10,113	37,738	90,373	141,286
-2 to -1	808	2,614	11,611	41,721	98,742	155,496
-1 to 0	884	3,032	12,685	45,476	107,230	169,307
0 to 1	1,006	3,283	13,570	49,018	114,579	181,456
1 to 2	991	3,502	14,106	51,041	119,670	189,310
2 to 3	1,021	3,403	13,830	49,552	118,589	186,395
3 to 4	955	3,183	12,732	45,468	110,887	173,225
4 to 5	802	2,762	11,167	39,714	95,421	149,866
5 to 6	771	2,546	9,612	34,298	83,560	130,787
6 to 7	641	2,123	7,997	28,741	71,741	111,243
Total	9,605	32,525	133,580	486,580	1,173,601	1,835,891

Project cohort, distance bin, and number of years since construction

A home transaction's **project cohort** refers to the unique ID of the LSPVP that is nearest to a home transaction within 4 miles, and for which the operation start date occurred up to 6 years before or after a LSPVP began construction. If a given transaction belonged to more than one cohort, we retained only the nearest project cohort for that transaction.

For each transacted home, a **distance bin** was determined that indicated the distance between that home and its nearest LSPVP. The distance between the transacted home and the nearest LSPVP was binned into 4 categories: [0 mi, 0.5 mi), [0.5 mi, 1 mi), [1 mi, 2 mi), and [2 mi, 4 mi].

To calculate the **number of years since LSPVP construction**, we subtracted the LSPVP year of construction start from the sale year (recall that the construction start year is assumed to be the operation start year minus 1 year). The years since LSPVP construction were categorized into 1-year bins (i.e. a sale occurred [-5 years, -4 years), [-4 years, -3 years),…,[5 years, 6 years), [6 years, 7 years] since LSPVP construction).

Transaction screening criteria

Condition for retention	Rationale
Coordinate values are populated	Coordinates are needed to obtain distances between homes and LSPVP, amenities, and dis-amenities
Land area, year built, and home square footage are populated	Land area, year built, and home square footage are essential property characteristics to control for in analysis
Coordinates appear 20 times or less	Repeated, identical coordinates for multiple properties may indicate data quality issue
Property type is residential (including single family residence, condominium, duplex, apartment)	Analysis only considers homes (i.e. residential properties) sold in arms length transactions after the year 2000
Transaction is categorized as arms length	
Year of sale between 2000 and 2021	
Sale amount is greater than \$5000 or the 1st percentile of sale price (whichever value is higher) and less than the 99th percentile of sale amount values within a given state	Removing outliers from analysis
Sale amount per unit area of living space is greater than the 1st percentile and less than the 99th percentile of sale amount per unit area of living space values within a given state	
Land area is greater than the 1st percentile and less than the 99th percentile of land area values within a given state	
Property was built before 2020, and after the 1st percentile of values for year built within a given state	

Condition for retention cont'd	Rationale cont'd
Sale amount is greater than the mortgage amount, or mortgage amount is missing	Any other relationship (between sale amount & mortgage amount, land area & living space area, sale year & year built, set of variables representing land area) may indicate data quality issues
Land area is greater than living space area	
Age of property (sale year minus year built) is non-negative	
Both variables representing land area converge within 0.01 acres	
Deed is not categorized as foreclosure	Sale amount in a foreclosure may not accurately represent the value of a home
Sale occurred over one year after last recorded sale for that property	Removes potentially "flipped" homes, or homes that undergo a rapid renovation and are re-sold, from dataset; for those homes, characteristics in CoreLogic dataset may not be representative of characteristics after renovation
Property address was not determined from mail	Address determined from mail may reflect the address of an absentee owner, not of the physical property location

Amenity and dis-amenity (A/D) data sources

Amenity/dis-amenity	Data source	Data description	Reference
Aviation noise	U.S. Department of Transportation	Raster representing approximate average noise energy due to transportation noise sources over a 24-hour period at the receptor locations where noise is computed, expressed in decibels (dB)	(U.S. Department of Transportation, 2020)
Road noise			
Flood zones	U.S. Federal Emergency Management Agency	Categorizes areas by likelihood of flood, ranging from minimal risk to 26% chance of flooding over the life of a 30-year mortgage	(Federal Emergency Management Agency, 2021)
Municipal, industrial, and transfer landfills	U.S. Department of Homeland Security	Provides locations of active permitted municipal solid waste facilities and construction and demolition debris facilities.	(Department of Homeland Security, 2020)
State and national parks	Esri	Provides boundaries of parks and forests in the United States at the national, state, regional, and local level	(Esri, 2021)
Nuclear power generation facilities	National Institute of Health	Provides locations of U.S. commercial nuclear power plants	(Hochstein and Szczur, 2006)
Coal power generation facilities	U.S. Environmental Protection Agency	Facility data (as of 2017) where primary or secondary fuel type is coal-related (e.g., Coal, Coal Refuse, and Petroleum Coke).	(U.S. Environmental Protection Agency, 2021)
Coastline	Hitachi Velocity Suite	Locations of U.S. coastline, including bays, river outlets, and Great Lakes	(ABB Group, 2020)
Lakes		Locations of U.S. lakes, represented as polygons	
High-voltage lines		Transmission and distribution lines with a voltage of 100 V or greater, represented as polylines	

Determining predominant prior land use for each LSPVP

Process:

- Determine the distribution of prior land cover types by area for each LSPVP; each LSPVP polygon is composed of some proportion of the NLCD land cover classes shown in the right-most column of the table below (15 of the 16 possible NLCD classes showed up in our sample)
- Group and sum each LSPVP's distribution of NLCD classes as per the right-most column of the table below
- Assign each LSPVP the predominant prior land use type that constituted 50% or more of its land cover
- If no single predominant prior land use type accounted for 50% or more of an LSPVP's prior land cover by area, that LSPVP was assigned a predominant prior land use type of "mixed"

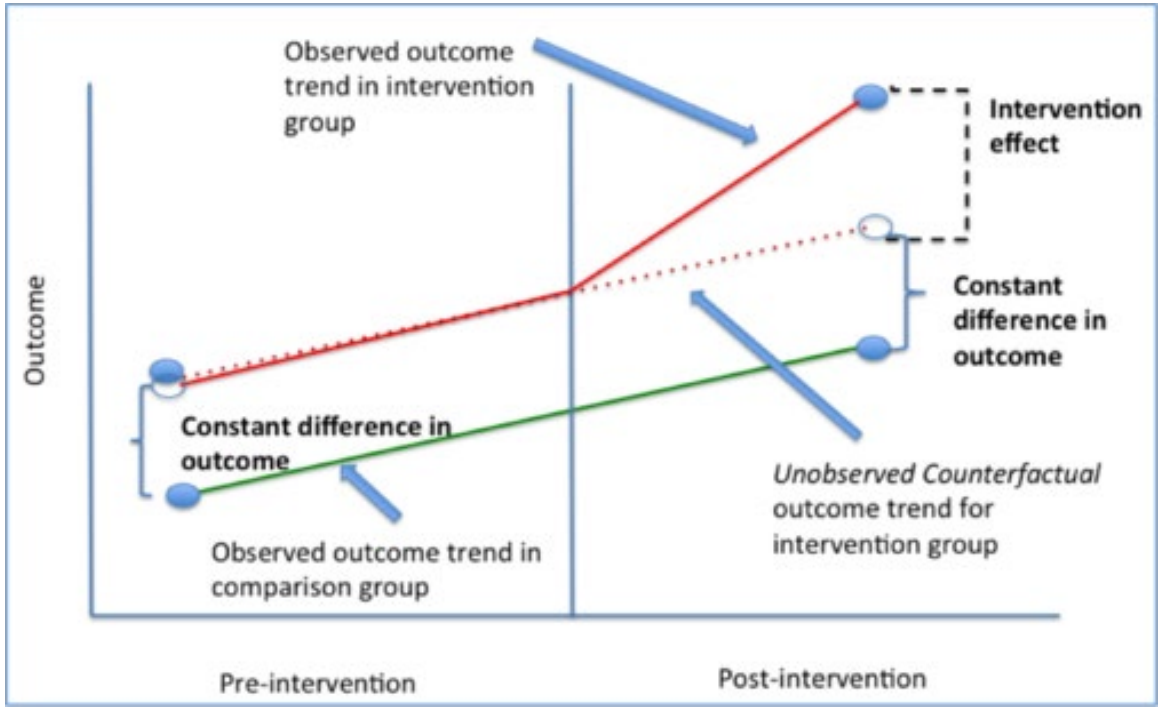
For instance, a solar installation on land that was, in 2006, 15% barren land, 25% cultivated crops, 25% herbaceous, and 35% hay/pasture, would be generalized as 60% agriculture and 40% greenfield, and would be given the predominant prior land use type of "agriculture". A solar installation on land that was, in 2006, 15% barren land, 25% developed, high intensity, 25% herbaceous, and 35% hay/pasture, would be generalized as 35% agriculture, 40% greenfield, and 25% brownfield, a would be assigned the predominant prior land use type of "mixed", because no single category amounted to greater than 50%.

Predominant prior land use type	**NLCD classes**
Agriculture	Cultivated Crops; Hay/Pasture
Brownfield	Developed, High Intensity; Developed, Low Intensity; Developed, Medium Intensity
Greenfield	Barren land; Deciduous forest; Developed, Open Space; Emergent Herbaceous Wetlands; Evergreen Forest; Herbaceous; Mixed Forest; Open Water; Shrub/Scrub; Woody Wetlands

Validating home coordinates: detailed process

- **Selecting properties to validate:** We selected properties that were < 0.5 miles from an LSPVP or A/D, within a flood zone with at least 1% chance of flooding, or within an area with road or aviation noise exceeding 55 dB. Of the properties that satisfied these conditions, only those with an area greater than 1 acre or those with missing or non-unique coordinates were validated.
- **Selecting homes to drop from dataset based on validation:** We dropped home transactions from our analysis if the difference between the coordinates provided by the Google Geocoding API and CoreLogic was greater than 2 times the distance between that home and its nearest PV plant or A/D. We additionally dropped any duplicate coordinates within 0.5 mi of a PV plant. Where the Google Geocoding API returned a “rooftop” precision indicator, we replaced the CoreLogic coordinates with Google coordinates; for those homes, we recalculated distances to LSPVPs and A/Ds using the process described in Step 2.

Methods: difference-in-difference estimation



Source: Columbia Public Health

A **difference-in-difference (DiD)** estimation is generally used to understand the impact of a "treatment" or intervention on a specific outcome by comparing the change in that outcome over time between a population that received the intervention (the intervention group) and one that did not (the comparison group).

In our case, the intervention is **the construction of an LSPVP,** and the outcome is **a home's sales price.** The intervention group are **homes proximal to an LSPVP** - up to 2 miles away - while the comparison group are **homes 2-4 miles away**; both the proximal and comparison groups have transacted within 6 years of the LSPVP construction date.

A DiD estimation allows us to assess how both *proximity to an LSPVP* and *construction of an LSPVP* impact sales price.

Base model

$$ln(P_{icajqt}) = \beta T_{iat} + X_i\alpha + \delta_{ac} + \lambda_{tc} + \rho_{qc} + \varphi_j + \varepsilon_{icajqt}$$

Natural log of sales price

Distance bin, interacted with an indication of whether the transaction occurred before or after LSPVP construction; transactions that occur 2-4 mi away from an LSPVP are the omitted category

Vector of individual home characteristics (e.g. home age, number of bedrooms)

Fixed effects (distance bin-by-cohort; transaction year-by-cohort; transaction quarter-by-cohort; census block group)

Random disturbance term

β: coefficients of primary interest; represent the DiD estimates of the effect of treatment (being close to an LSPVP post construction) on home prices for homes located [0 mi, 0.5 mi), [0.5 mi, 1 mi), and [1 mi, 2 mi) away from an LSPVP, respectively

Note: standard errors clustered at project cohort level

Robustness checks

We estimated three alternative specifications to investigate the robustness of the base model to the choice of spatial FEs, time FEs, and treatment and control categories:

1. We add a distance bin for homes located within 0.25 miles of a LSPVP; specifically, we augment the distance bins in (1) to include four (rather than three) indicators for homes located in the [0 mi, 0.25 mi) , [0.25 mi, 0.5 mi), [0.5 mi, 1 mi), and [1 mi, 2 mi) distance bins; the indicator equals 1 if a transaction occurred within that distance bin in the same year or after LSPVP construction started, and 0 otherwise. This specification allows us to investigate the presence of a home price effect at even smaller distances
2. We replace the year-by-project cohort and quarter-by-project cohort FEs in the base model by a single vector of quarter-by-year-by-project cohort FEs to allow for more granular trending of home values across quarters and years
3. We add the vector of A/D variables to account for any potential correlation between the A/D variables and the timing and location of a LSPVP that may bias our base model estimates

Event study model

A parallel trends assumption - that home prices near LSPVP sites were not already significantly different from home prices further away before the LSPVP was constructed - underlies the base model. We use the event study model to test this assumption.

$$\ln(P_{icdjqt}) = \sum_{j=-5}^{7} T_{j,idt}\gamma_j + \boldsymbol{X}_i\kappa + \delta_{dc} + \lambda_{tc} + \rho_{qc} + \varphi_j + \eta_{icdjqt}$$

Natural log of sales price

Lead and lag indicators, indicating the number of years before or after LSPVP construction, for each distance bin

Vector of individual home characteristics (e.g. home age, number of bedrooms)

Fixed effects (distance bin-by-cohort; transaction year-by-cohort; transaction quarter-by-cohort; census block group)

Random disturbance term

γ: coefficients of primary interest; represent the DiD estimates of the effect of treatment (being close to an LSPVP in a given year before or after that LSPVP is constructed) on home prices for homes located [0 mi, 0.5 mi), [0.5 mi, 1 mi), and [1 mi, 2 mi) away from an LSPVP compared to homes 2-4 mi away

Note: standard errors clustered at project cohort level

Base model results with robustness checks

*Dependent variable: the logarithm of house prices. Standard errors are clustered at the project cohort level and are in parentheses. Significance levels: *** p< 0.01, ** p<0.05, * p<0.1*

	Base model	Including 0-0.25 mi distance bin	Including quarter-year-project cohort FEs	Including amenities and disamenities vector
Distance between home and LSPVP: [0 mi, 0.25 mi)		-0.0226*** (0.00767)		
Distance between home and LSPVP: [0.25 mi, 0.5 mi)		-0.0133** (0.00641)		
Distance between home and LSPVP: [0 mi, 0.5 mi)	**-0.0154**** **(0.00630)**		-0.0171*** (0.00642)	-0.0170*** (0.00589)
Distance between home and LSPVP: [0.5 mi, 1 mi)	**-0.00820**** **(0.00413)**	-0.00820** (0.00413)	-0.00941** (0.00424)	-0.00987** (0.00403)
Distance between home and LSPVP: [1 mi, 2 mi)	**-0.000841** **(0.00226)**	-0.000841 (0.00226)	-0.00179 (0.00234)	-0.00131 (0.00225)
Home characteristics	✓	✓	✓	✓
Distance-project cohort FEs	✓	✓	✓	✓
Sale year-project cohort FEs	✓	✓		✓
Sale quarter-project cohort FEs	✓	✓		✓
Census block group FEs	✓	✓	✓	✓
Sale year-sale quarter-project cohort FEs			✓	
Amenities and disamenities				✓
Observations	**1,832,888**	1,832,888	1,826,915	1,778,533
R^2	**0.835**	0.835	0.839	0.835

Heterogeneity analysis results: state

*Dependent variable: the logarithm of house prices. Standard errors are clustered at the project cohort level and are in parentheses. Significance levels: *** p< 0.01, ** p<0.05, * p<0.1*						
	CA	CT	MA	MN	NC	NJ
Distance between home and LSPVP: [0 mi, 0.5 mi)	0.00899 (0.0106)	0.0161 (0.0314)	-0.0144 (0.00892)	-0.0395** (0.0174)	-0.0576*** (0.0148)	-0.0559*** (0.0114)
Distance between home and LSPVP: [0.5 mi, 1 mi)	0.000849 (0.00696)	0.0234 (0.0150)	-0.00933** (0.00469)	-0.0209** (0.00932)	-0.0473*** (0.0118)	-0.0135* (0.00698)
Distance between home and LSPVP: [1 mi, 2 mi)	0.00296 (0.00384)	0.0186** (0.00786)	-0.00190 (0.00319)	-0.0108* (0.00625)	-0.0117** (0.00570)	-0.00487 (0.00331)
Observations	931,735	34,135	291,403	74,905	203,005	297,677
R^2	0.881	0.774	0.777	0.708	0.735	0.751

Heterogeneity analysis results: prior LSPVP land use

*Dependent variable: the logarithm of house prices. Standard errors are clustered at the project cohort level and are in parentheses. Significance levels: *** p< 0.01, ** p<0.05, * p<0.1*				
	Greenfield	Agricultural	Brownfield	Mixed
Distance between home and LSPVP: [0 mi, 0.5 mi)	-0.00646 (0.00960)	-0.0302*** (0.0107)	0.0122 (0.0159)	-0.0439 (0.0445)
Distance between home and LSPVP: [0.5 mi, 1 mi)	-0.000991 (0.00480)	-0.0202*** (0.00629)	-0.00909 (0.0170)	-0.00679 (0.0342)
Distance between home and LSPVP: [1 mi, 2 mi)	0.000836 (0.00248)	-0.00408 (0.00498)	-0.00483 (0.00739)	-0.000377 (0.0191)
Observations	1,074,492	577,769	147,951	12,987
R^2	0.843	0.833	0.860	0.828

Heterogeneity analysis results: urbanicity

*Dependent variable: the logarithm of house prices. Standard errors are clustered at the project cohort level and are in parentheses. Significance levels: *** p< 0.01, ** p<0.05, * p<0.1*			
	Rural	Urban cluster	Urban
Distance between home and LSPVP: [0 mi, 0.5 mi)	-0.0418*** (0.0156)	0.0324 (0.0524)	-0.00350 (0.00619)
Distance between home and LSPVP: [0.5 mi, 1 mi)	-0.0201* (0.0119)	0.0221 (0.0316)	-0.00342 (0.00437)
Distance between home and LSPVP: [1 mi, 2 mi)	0.00775 (0.00613)	-0.00597 (0.00896)	0.00137 (0.00222)
Observations	151,792	79,279	1,592,715
R^2	0.803	0.785	0.845

Heterogeneity analysis results: LSPVP size

*Dependent variable: the logarithm of house prices. Standard errors are clustered at the project cohort level and are in parentheses. Significance levels: *** p< 0.01, ** p<0.05, * p<0.1*		
	LSPVP area < 50th percentile of area (75,138 m^2)	**LSPVP area ≥ 50th percentile of area (75,138 m^2)**
Distance between home and LSPVP: [0 mi, 0.5 mi)	-0.00737 (0.00694)	-0.0305** (0.0138)
Distance between home and LSPVP: [0.5 mi, 1 mi)	-0.00483 (0.00521)	-0.0166** (0.00684)
Distance between home and LSPVP: [1 mi, 2 mi)	0.00225 (0.00287)	-0.00841** (0.00344)
Observations	1,291,762	537,189
R^2	0.841	0.833