The Broadband Initiatives Program (BIP) Impact on New Construction

Technical Report

Guy Leonel Siwe, Josh Goldstein, Neil Kattampallil,
Biocomplexity Institute, Social and Decision Analytics Division
Zhengyuan Zhu, Iowa State University
Xin Wang,, San Diego State University
John Pender, USDA Economic Research Service

February 1, 2024

Abstract

This study examines how the Broadband Initiatives Program (BIP) has affected new construction in rural areas. We analyze property sales data from CoreLogic at the census tract level to estimate the number of new constructions. By comparing the amount of new construction in tracts within the project area to those outside the area (within a 25-mile buffer zone), we evaluate the impact of BIP. We match tracts inside and outside the project area based on similar characteristics. Our analysis uses a combination of difference-in-differences, zero-inflated Poisson, and Hurdle models. Surprisingly, our findings reveal a notable decrease in the volume of new property sales in the BIP project area, which is contrary to what we expected.

Citation: SIWE G, Goldstein J, Kattampallil N, Zhu Z, Wang X, Pender J. (2024). The Broadband Initiatives Program (BIP) Impact on New Construction. Social and Decision Analytics Division, Biocomplexity Institute, University of Virginia. https://doi.org/10.18130/gznb-0v71

Source of Funding: This research was funded through a collaborative agreement with the USDA Economic Research Service #AWD-003263.

INTRODUCTION

Broadband Initiatives Program (BIP)

- Large investments are made by the Federal Government to address the rural-urban digital divide.
- The BIP goal is to facilitate broadband deployment in rural communities and was established in 2009 funded by the American Recovery and Reinvestment Act.
- $2.5 billion was allocated to the program in two funding rounds (FY 2009 and FY 2010) and the program was implemented by the USDA Rural Utilities Service (RUS).
- Eligibility and coverage (Pender et al., 2023):
  - Unserved areas (90% of households lacked access to broadband services at the minimum advertised speed of 768 kbps downstream and 200 kbps upstream).
  - Underserved areas (50% of households lack access at those speeds).
  - During the second round the criteria was relaxed to 5 Mbps (downstream + upstream).
- Projects were implemented in 2009 but approved by September 2010.

Research Questions

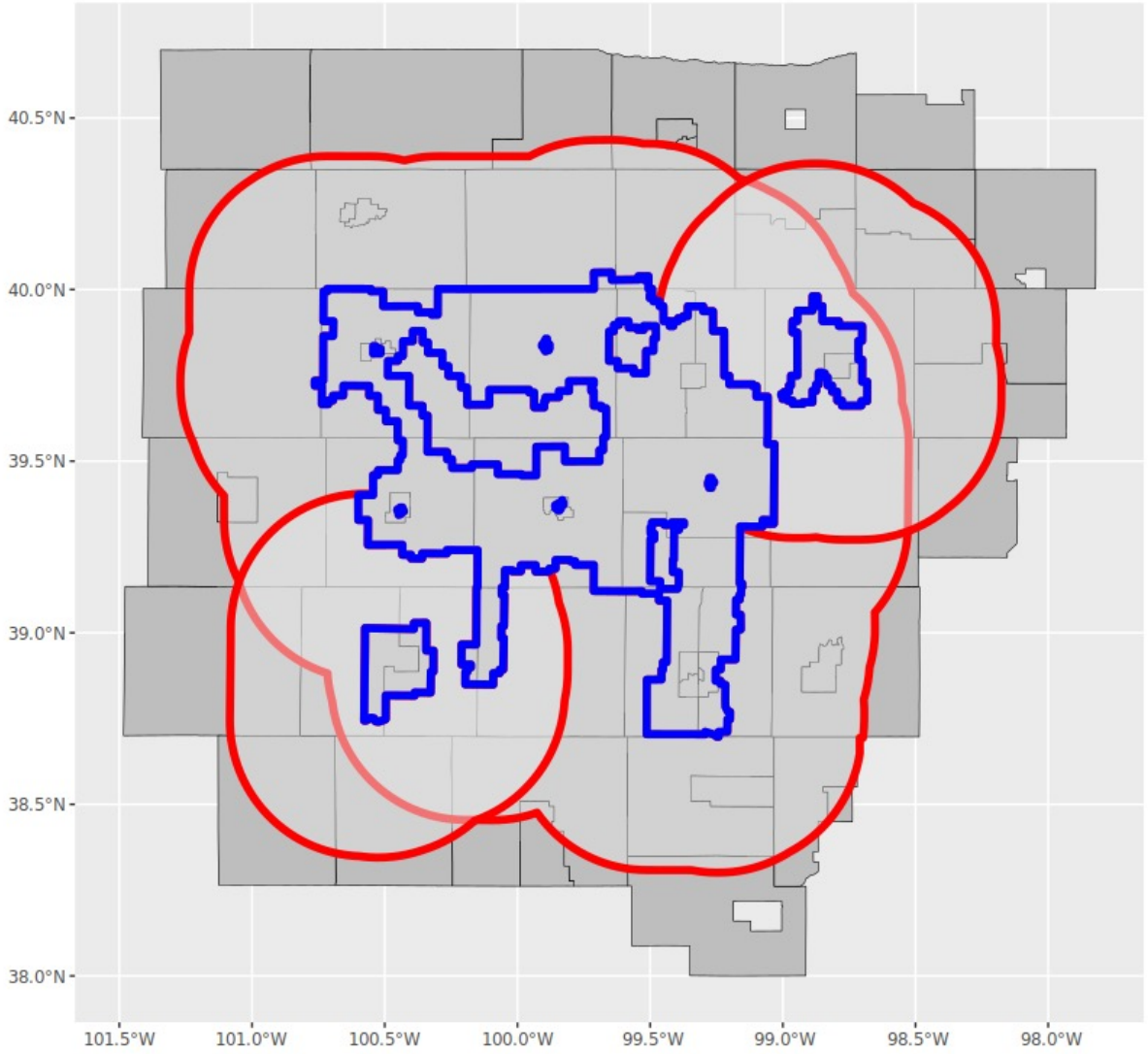
- Our goal is to investigate the economic benefits of the BIP program. More precisely, we are investigating how the BIP program impacted the supply of new properties in the housing market.
- We hypothesized that BIP has increased the supply of properties in rural areas in expectation of population growth.
    - Phumsith and al. (2009) and Tim Marema (2015) found that broadband accessibility has a positive impact on population growth in rural areas.
    - Following an anticipation of population growth, home builders would increase their supply of houses.
- We use a difference in difference framework to assess the impact of the BIP program on the volume of new property sales.

DATA

- We used three main data sources for this analysis:
  - BIP projects shapefile: GIS shapefiles describing the geography boundary of each project awarded. 236 BIP projects were covered.
  - CoreLogic data: property sale transactions and property tax assessments.
    - Sale transactions report information on sales price, sales date and the transaction type.
    - Property tax data include property characteristics such as location, size of the parcel, year house built, etc.
  - American Community Survey (ACS) data: Census tract characteristics
  - Units for analysis: 2010 census tracts
  - CoreLogic variables: the type of sales which are: resale, new construction, short sale, etc.
    - Focus on sales with transaction type coded 3 which corresponds to ***New construction***.
    - Assumption: CoreLogic records all new construction and sales

Defining the Treatment and Control Areas

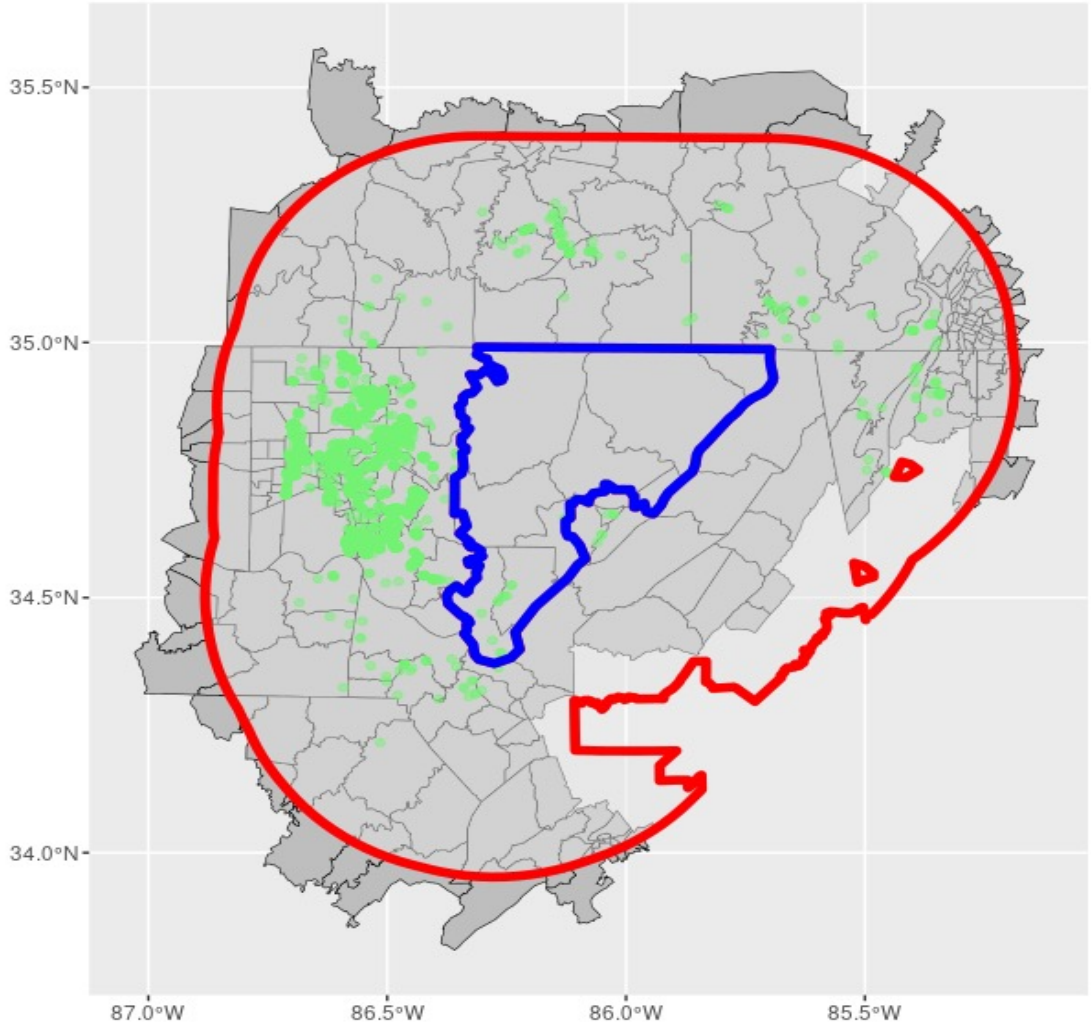
- The BIP project areas (*in blue*) are the treatment areas. Census tracts inside the project area are considered as treated observation.
  - For 2010 census tracts overlapping with the project area, we include census tracts covered at 75% by the BIP project in the treatment area.
  - There are 616 tracts in the treatment area. There are 239 BIP projects in this area.
- Control area definition:
  - Introduce a buffer of 25 miles from a program area to define the control area (*in red*).
  - We include all census tracts fully or partially included (at least 75%) between the treatment boundary (*in blue*) and the control boundary (*in red*).
  - There are 23,556 tracts in the control area.



Project KS1103-A40. The blue line represents the boundary of the project area under BIP. The red line represents the boundary of the control area. Gray polygons correspond to our unit of analysis representing the 2010 census tracts.

DATA TREATMENT

- We geolocated new single-family home property sales (green dots), between 2005 and 2017, from CoreLogic into the Treatment (inside blue area) and Control areas (between the blue and red) and computed the volume of new property sales by year.
- Properties were georeferenced to census tracts using latitude-longitude information.
- The volume of new property sales was counted by census tracts over the years. Treatment and control areas without new construction were excluded from our analysis (see the example of ***project AL1105-B39***).
- There are 166 BIP projects (69% of the initial sample of projects), with a total of 7281 census tracts that have new home property sales.
    - 279 census tracts are included inside the treatment area.



Project AL1105-B39. The blue line represents the boundary of the project area under BIP. The red line represents the boundary of the control area. Gray polygons correspond to our unit of analysis representing the 2010 census tracts. Green points represent new single-home family homes both in the treated and controlled areas.

MATCHING

- To account for the similarity in the unit of analysis, we match treated census tracts to the controlled census tracts within a project area based on confounding (similar) characteristics.
- We identify a set of confounding characteristics and match those census tracts based on those characteristics. The data on those characteristics were collected pre-BIP program from the ACS-5 years estimates in 2010 to match treated and controlled census tracts pre-program.[1]
- List of census tract characteristics:

  - Population density
  - Percentage of owned occupied housing units (homeownership rate)
  - Percentage of rented vacant housing units (sales vacancy rate)
  - Percentage of occupied housing units (occupancy rate)
  - Percentage of the population renters
  - Percentage of recent homes built between 2005-2010 (pre-program)
  - Poverty rate
  - Percentage of minority racial population (Black+Hispanic+Asian)
  - Median Income
  - Percentage of family household
  - Labor force participation rate
  - Foreign population (in percent)
  - Population aged 65 and more (in percent)
  - Percentage of family household
  - New property sales before 2005-2009 (we put half of the weight on this variable during the matching of census tracts)

[1] By using the ACS-5 years estimates from 2010, we covered the BIP pre-program going from 2005-2010.

MATCHING

- We use the ***Mahalanobis distance*** for a one-to-one matching between treated and controlled census tracts.

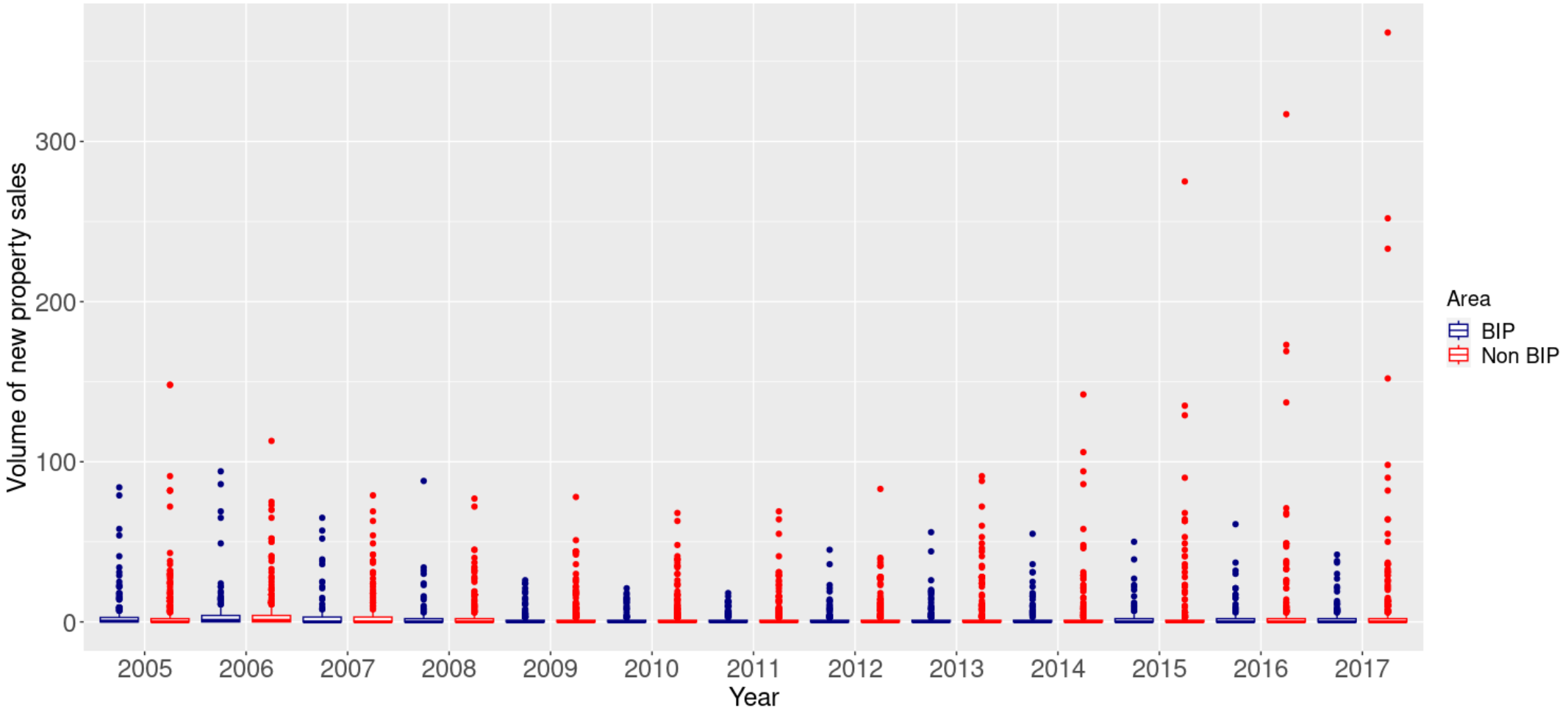
$$M(X_i; X_j) = \sqrt{(X_i - X_j)' S^{-1} (X_i - X_j)}$$

Where X_i are the vector of census tracts i characteristics and

$S^{-1} = I$ the weighted matrix of census tract characteristics.

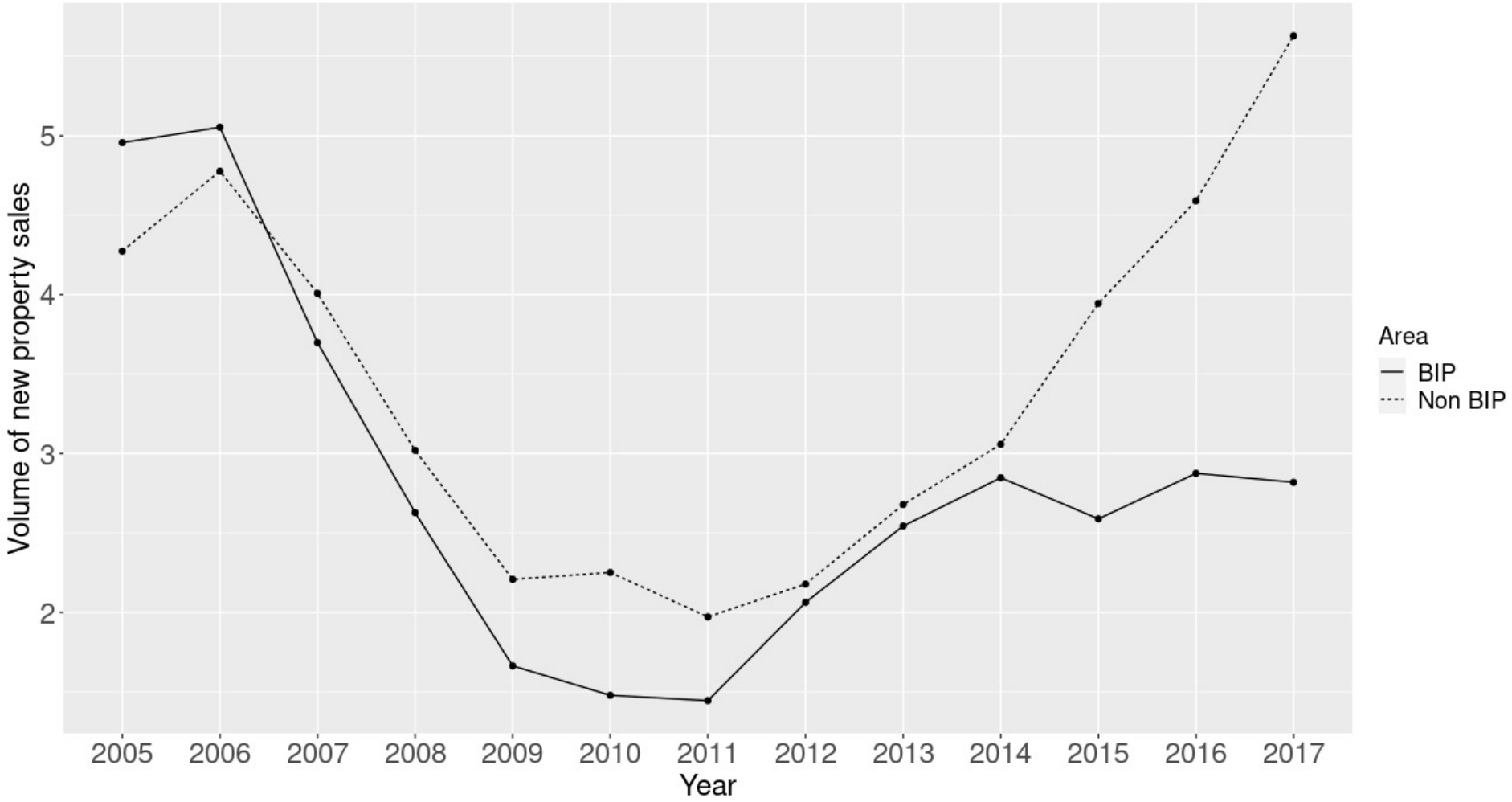
- After matching, the final sample drops to 157 BIP projects with a total of 483 census tracts.
- Out of those 483 census tracts, 188 were included in treatment areas.

Figure 1: Distribution of New Property Sales over Time



- Boxplots of new property sales distribution by project areas (BIP or non-BIP) and years.
- New property sales distribution is right skewed and highlights that a lot of census tracts have zero new property sales each year. This analysis is based on the 483 census tracts, noted in slide 10.

Figure 2: Pre-Program Parallel Trends Validation



- The pre-program parallel trend assumptions seems to be verified, except in 2010 (the year of BIP project approval) where the difference starts to increase.
- Note that the pre-program declining trends in the volume of new property sales in and outside the BIP area may be driven by the subprime crisis.

MODEL

- We investigate the impact of the BIP program on the volume of new property sales.
- Due to the high number of census tracts with no new property sales, we use a ***zero-inflated Poisson*** and ***Hurdle model*** combined with a difference in difference framework.
- This framework estimates two models:
  - *A count model:* to evaluate how the program has changed the volume of new property sales for census tracts where new property sales were recorded.
  - *A zero-inflated/zero Hurdle model:* to evaluate how the program has changed the likelihood of observing census tracts with no property sales.
- ***Difference between the zero-inflated and Hurdle models:*** Those two models differ in the interpretation of the zero count.
  - The hurdle model assumes the zero count to be only structural (*meaning tracts with always zero new property sales that may be due to the land structure*)
  - The zero-inflated model considers that the zero count can additionally result from sampling (*included tracts with partially non-zero new property sales at some point of time*).[2]
- For the zero-count model, we use the log-log link due to the high proportion of census tracts with no new property sales.

[2] A great resource is the paper of Mei-chen Hu et al. (2011), *Zero-inflated and Hurdle Models of Count Data with Extra Zeros: Examples from an HIV-Risk Reduction Intervention Trial*, Am J Drug Alcohol Abuse. Author manuscript; available in PMC 2012 Sep 1.

Zero-inflated Poisson Estimations

- The table illustrates the estimated effect of BIP on the volume of new property sales from the count and zero-inflation model. The reference year is 2005. The parallel trend assumption is verified (except in 2009).
- From the zero-inflation model, the BIP implementation in 2010 has reduced the likelihood of observing census tracts with new property sales between 2011-2012. However, those changes weren't significant.
- From the count model, the BIP program significantly reduces the volume of new property sales in areas where new property sales were observed. For example, in 2011, the program reduced by 50% the volume of new property sales inside the project area relative to the control area. **The result is counter to expectations that broadband deployment would increase new construction.**

	Count model (truncated Poisson with log link)			Zero-inflation model (binomial with cloglog link)		
	Estimate	Std. Error	z value	Estimate	Std. Error	z value
bip:2006	0.0193	0.0582	0.3312	0.1481	0.1867	0.7931
bip:2007	-0.0889	0.0624	-1.4255	0.1871	0.1820	1.0281
bip:2008	-0.1304	0.0683	-1.9084	0.1992	0.1768	1.1265
bip:2009	**-0.2117**	**0.0781**	**-2.7105**	0.2447	0.1718	1.4247
bip:2010	**-0.4234**	**0.0814**	**-5.2005**	0.1807	0.1731	1.0439
bip:2011	**-0.5059**	**0.0861**	**-5.8776**	0.0069	0.1765	0.0392
bip:2012	**-0.2550**	**0.0765**	**-3.3316**	-0.0124	0.1782	-0.0699
bip:2013	-0.0521	0.0709	-0.7346	0.1863	0.1783	1.0451
bip:2014	**-0.2133**	**0.0688**	**-3.0991**	0.0274	0.1818	0.1505
bip:2015	**-0.4170**	**0.0691**	**-6.0315**	0.1913	0.1789	1.0691
bip:2016	**-0.4289**	**0.0665**	**-6.4455**	0.2363	0.1802	1.3111
bip:2017	**-0.5689**	**0.0662**	**-8.5895**	0.3435	0.1809	1.8988

*BIP and Year fixed effects were included. We also use the log-log link to account for the imbalance of tracts with and without new property sales. Absolute z-value higher than 2 indicated significant effect.

Hurdle Model Estimations

- The table illustrates the estimated effect of BIP on the volume of new property sales from the count and Hurdle model. The reference year is 2005. The parallel trend assumption is verified (except in 2009).
- From the zero hurdle model, the BIP implementation in 2010 has increased the likelihood of observing census tracts with new property sales only between 2011-2012. However, those changes weren't significant.
- From the count model, the BIP program significantly reduces the volume of new property sales in areas where new property sales were observed. Similarly to the zero-inflated model, in 2011, the program reduced by 50% the volume of new property sales inside the project area relative to the control area. **The result is counter to expectations that broadband deployment would increase new construction.**

	Count model (truncated Poisson with log link)			Zero hurdle model (binomial with cloglog link)		
	Estimate	Std. Error	z value	Estimate	Std. Error	z value
bip:2006	0.0190	0.0582	0.3260	-0.1556	0.1843	-0.8442
bip:2007	-0.0892	0.0624	-1.4294	-0.1963	0.1880	-1.0439
bip:2008	-0.1307	0.0683	-1.9125	-0.2166	0.1942	-1.1156
bip:2009	**-0.2121**	**0.0781**	**-2.7167**	-0.2922	0.2048	-1.4272
bip:2010	**-0.4238**	**0.0814**	**-5.2051**	-0.2144	0.1992	-1.0763
bip:2011	**-0.5063**	**0.0861**	**-5.8819**	0.0192	0.2075	0.0923
bip:2012	**-0.2554**	**0.0765**	**-3.3369**	0.0460	0.2023	0.2275
bip:2013	-0.0524	0.0709	-0.7392	-0.2028	0.2002	-1.0128
bip:2014	**-0.2135**	**0.0688**	**-3.1029**	-0.0178	0.1961	-0.0910
bip:2015	**-0.4173**	**0.0691**	**-6.0354**	-0.2102	0.2013	-1.0440
bip:2016	**-0.4291**	**0.0665**	**-6.4492**	-0.2582	0.1969	-1.3110
bip:2017	**-0.5692**	**0.0662**	**-8.5933**	-0.3738	0.1960	-1.9072

*BIP and Year fixed effects were included. We also use the log-log link to account for the imbalance of tracts with and without new property sales

Conclusion and Interpretation

Comments :

- The model results indicate that the BIP program reduces the supply of new property in rural areas. **The result is counter to expectations that broadband deployment would increase new construction.**
- First, there is not enough variability in the volume of new property sales observed across census tracts in the final sample. As shown in Figure 2, the average number of new property sales per year for tracts inside or outside the project area is between 1 to 5. Thus, a 50% reduction in the volume of new property due to the BIP only suggests a reduction of new property sales by 1 or 2.
- Second, the negative effect of the BIP may be driven by some equilibrium in the housing market. With an increase in housing prices following the broadband (Hannah and al., 2022), people have reduced their demand for single-family homes, and then new properties.[3] Therefore, the supply of new property may have fallen even if the home value has increased.
- Additional investigation needs to be done, especially with data where new properties are observed (not only new properties that are sold)

[3] Given that most of the people in rural area have a low income according to the Census

BIBLIOGRAPHY

1. Mei-chen Hu et al. (2011), *Zero-inflated and Hurdle Models of Count Data with Extra Zeros: Examples from an HIV-Risk Reduction Intervention Trial*, Am J Drug Alcohol Abuse. Author manuscript; available in PMC 2012 Sep 1.
2. Phumsith and al. (2009), Does Broadband Access Impact Population Growth in Rural America?, working paper
3. Pender, J., Goldstein, J., Mahoney-Nair, D., & Charankevich, H. (2023). Three USDA rural broadband programs: Areas and populations served (Report No. EIB-258). U.S. Department of Agriculture, Economic Research Service. https://doi.org/10.32747/2023.8134353.ers
4. Tim Marema (2015), Broadband Linked to Population Growth.

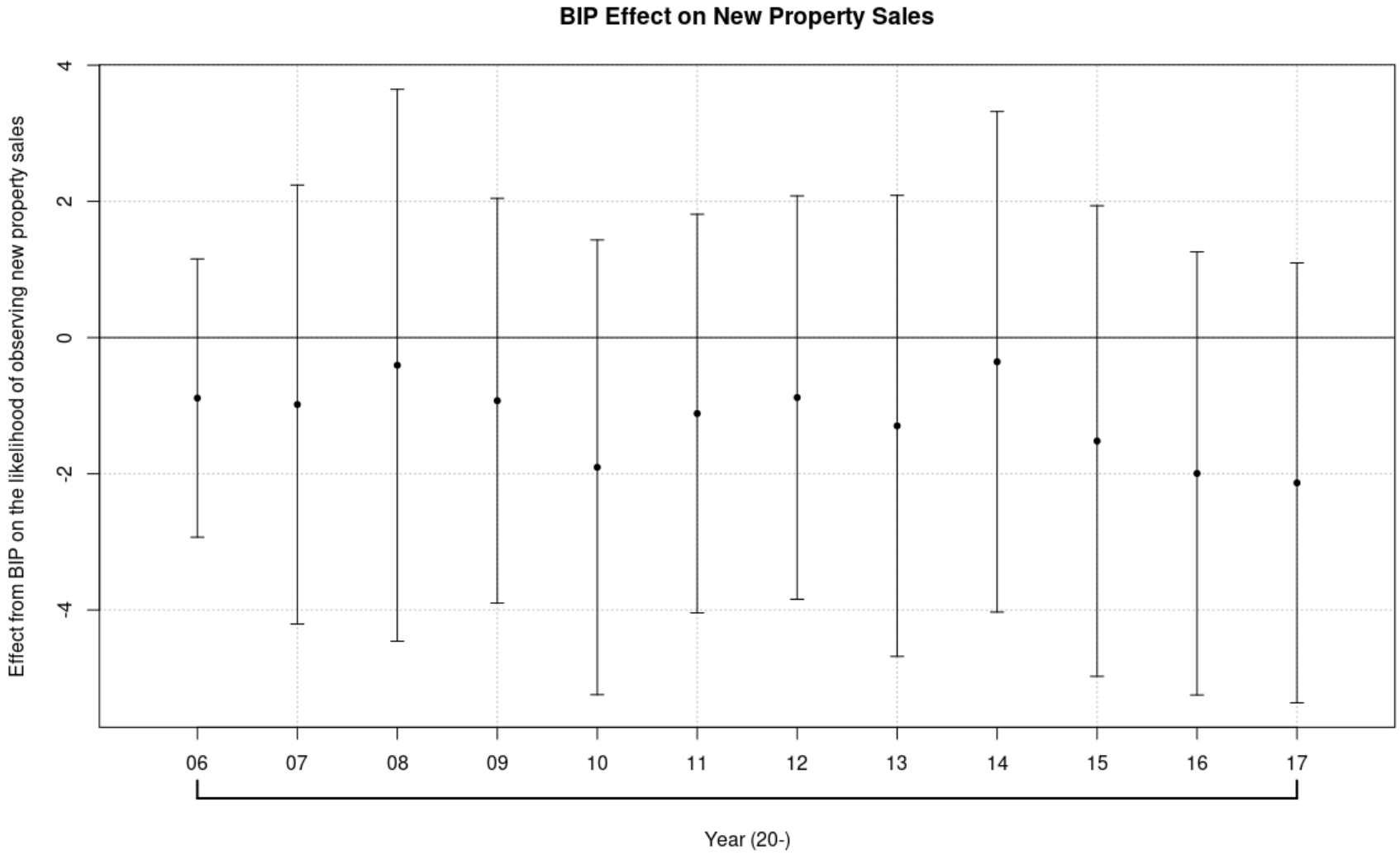
Appendix

- We investigate the BIP impact on the likelihood that a new property is observed in the sales data.
- In this scenario, the unit of observation is a property unit (single-family homes). The outcome variable *(Y_t)* indicates if a property is a new construction or not. By changing the unit of observation, we increase the sample size and reduce the zero count in the volume of new construction. However, the distribution remains imbalanced because a large number of properties are not classified as new.
- The model specification is:

$$Y_t = \alpha_0 + \alpha_1 Year + \alpha_3 BIP + \alpha_4 Year * BIP + FE_{project;\, tracts} + \epsilon$$

where Y_t = 1 if new property sales and 0 if not

Appendix: BIP Effect on the Likelihood of New Property Sales



- There is no significant and positive effect from the BIP.
- The marginal effect is still negative after the program implementation.