Appendix

A1. OZ Overview

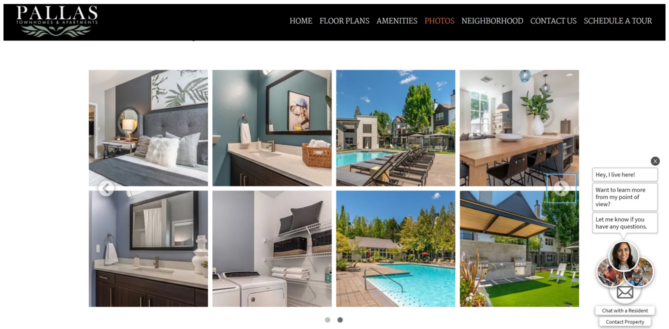


Figure A1 An OZ-based $186M luxury apartment complex acquired by MG Properties in 2019 via a joint venture with the Holland Partner Group and Invesco (source: https://pallas.mgproperties.com, date accessed: Jan. 2020).

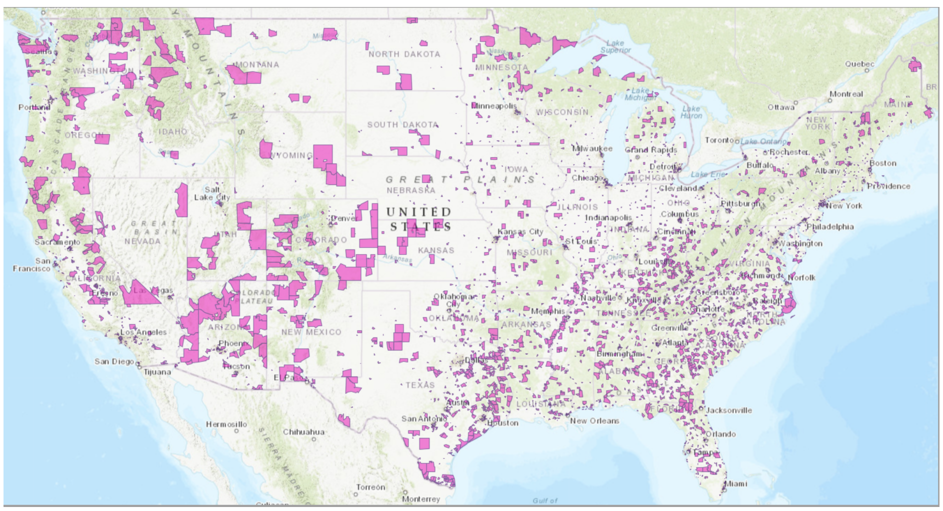


Figure A2 Map of continental U.S. opportunity zones (source: U.S. Department of the Housing and Urban Development).

A2. OZ Selection Process

	Price	Volume	Poverty rate	Unempl. rate	Income level
Price	1.000	0.046	-0.183	-0.191	0.386
Volume	0.046	1.000	-0.250	-0.103	0.268
Poverty rate	-0.183	-0.250	1.000	0.624	-0.729
Unempl. rate	-0.191	-0.103	0.624	1.000	-0.504
Income level	0.386	0.268	-0.729	-0.504	1.000

Table A1 Correlation matrix for qualified census tracts.

A2.1. Using all OZs and all NOZs

To complement our analysis using the sample including all OZs and all the QNS tracts, we estimate the model in Eq. (1) using the sample that includes all OZs and all NOZs (including all the QNS and NQ tracts) and present the results in Tables A2 and A3, respectively.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Price	-0.009***					0.002***
	(-18.968)					(5.548)
Volume		-0.029***				-0.007***
		(-11.179)				(-2.904)
Poverty rate			0.132***			0.035***
			(50.186)			(8.277)
Unempl. rate				0.230***		0.065***
				(43.349)		(9.416)
Income level					-0.097***	-0.067***
					(-41.939)	(-21.554)
Const	-1.631***	-2.156***	-4.454***	-5.002***	2.038***	-0.608***
	(-28.341)	(-38.327)	(-87.670)	(-76.629)	(20.762)	(-3.063)
No. Obs.	30,374	30,374	30,374	30,374	30,374	30,374
Pseudo R^2	0.041	0.013	0.207	0.145	0.242	0.264

Table A2 Logistic regression to examine how census tracts were selected as OZs.

		Price	Volume	Poverty rate	Unempl. rate	Income level
All	Pseudo R^2	0.288	0.288	0.283	0.283	0.240
	LR stat.	2.465	0.015	66.860	68.698	647.117

Table A3 Significance of each component in explaining the selection of OZs.

A2.2. Using matched subsamples from PSM

To compare the matched subsamples, we present the histogram of the demographics variables in Figure A3 and the increase in price per square foot in OZs versus QNS tracts across states in Figure A4, based on the matched subsamples. Table EC5 in the e-companion reports the state-level descriptive statistics of the matched subsamples. These descriptive analyses and figures show that the matched subsamples have similar demographic characteristics, and the average price per square foot in OZs was higher relative to that in QNS tracts before the launch of the OZ program (4.4% higher as shown in Table EC5).

To further conduct a rigorous comparison between the matched subsamples in terms of residential real estate price, we estimate Eq. (1) using the matched subsamples, and present our results in Table A4. We observe that the coefficients corresponding to demographics variables are not statistically significant in all model specifications, confirming that the matched OZs and NOZs have similar demographics characteristics. However, the coefficient for the price is positive and statistically significant, thus supporting our finding that tracts designated as OZs had a higher average price relative to tracts with similar demographics that were not selected as OZs.

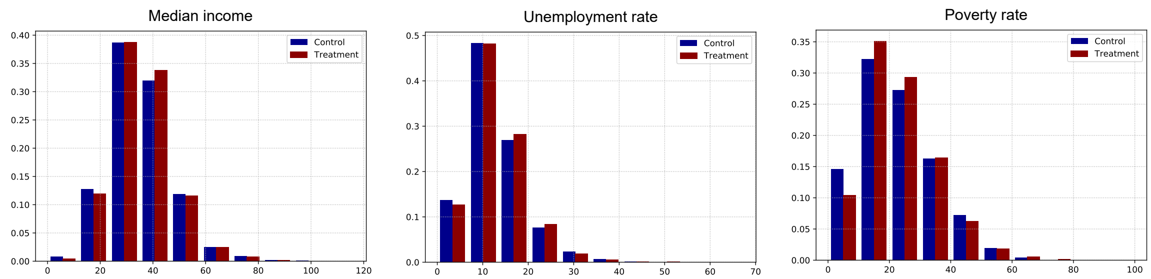


Figure A3 OZs versus NOZs using the matched dataset.

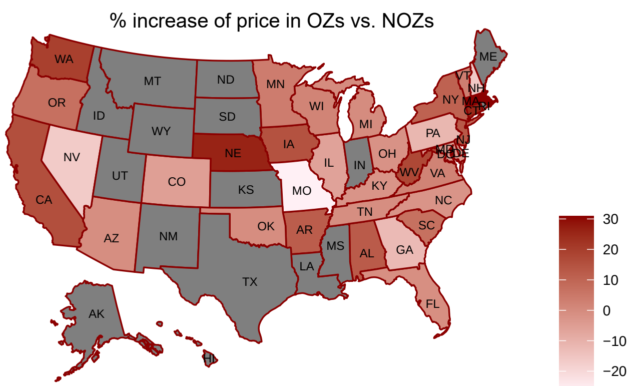


Figure A4 Percentage price increase in OZs vs. NOZs using the matched dataset (states that were not included appear in gray).

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Price	0.001*					0.001**
	(1.873)					(2.120)
Volume		0.000				0.000
		(1.098)				(1.202)
Poverty rate			0.006			0.007
			(1.498)			(1.113)
Unempl. rate				0.009		0.007
				(1.203)		(0.753)
Income level					-0.001	-0.000
					(-0.431)	(-0.038)
Const	-0.108	-0.070	-0.112	-0.110	0.056	-0.408
	(-1.585)	(-0.955)	(-1.347)	(-1.119)	(0.415)	(-1.465)
No. Obs.	3,026	3,026	3,026	3,026	3,026	3,026
Pseudo R^2	0.001	0.000	0.001	0.000	0.000	0.002

Table A4 Logistic regression to examine the OZ selection process. The OZ and NOZ tracts are matched based on key demographic variables.

A2.3. Using an unfiltered sample

We study the OZ selection process by focusing on census tracts with at least ten transactions per quarter during the pre-treatment period in Section 4, for the reasons described in Section 3.2. In this section, we replicate the analysis using the data of qualified census tracts without filtering out the census tracts with fewer than ten transactions per quarter during the pre-treatment period for robustness check. Specifically, Table A5 corresponds to Table 1. Figures A5 and A6 correspond to

Figures A3 and A4, respectively, and Table EC6 corresponds to EC5. The robustness tests show similar qualitative results as in Section 4.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Price	-0.002***					0.001***
	(-11.130)					(5.573)
Volume		-0.031***				-0.007***
		(-19.321)				(-4.294)
Poverty rate			0.060***			0.021***
			(48.564)			(10.514)
Unempl. rate				0.090***		0.020***
				(37.545)		(6.627)
Income level					-0.062***	-0.040***
					(-47.516)	(-19.845)
Const	-1.229***	-1.132***	-2.621***	-2.494***	0.960***	-0.591***
	(-54.132)	(-56.658)	(-86.634)	(-75.568)	(19.510)	(-5.453)
No. Obs.	32,583	32,583	32,583	32,583	32,583	32,583
Pseudo R^2	0.004	0.014	0.079	0.045	0.086	0.096

Table A5 Logistic regression to examine how qualified census tracts were selected as opportunity zones (without filtering out the tracts with fewer than ten transactions per quarter during the pre-treatment period).

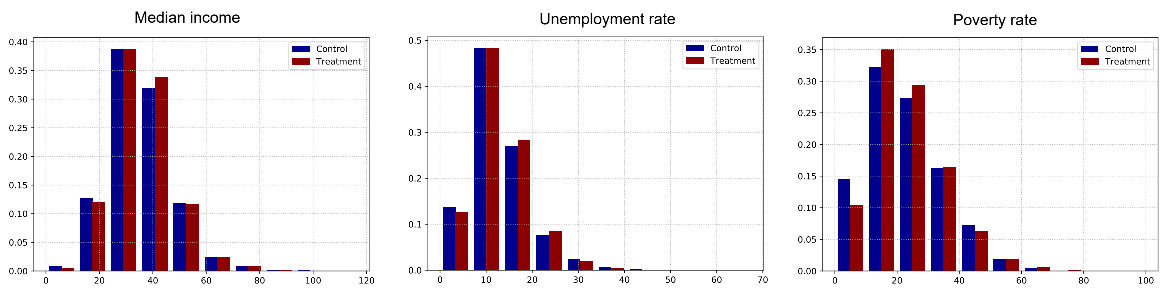


Figure A5 OZs versus NOZs using matched census tracts (without filtering out the tracts with fewer than ten transactions per quarter during the pre-treatment period).

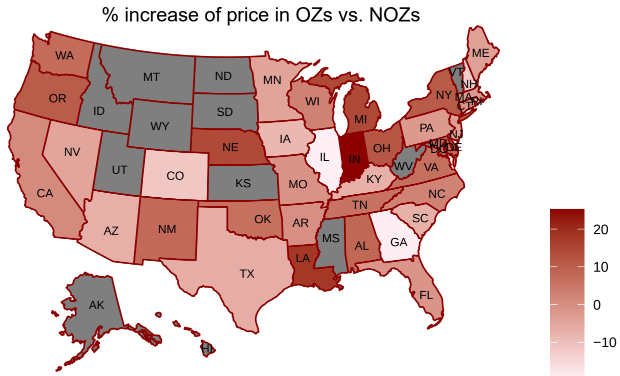


Figure A6 Percentage increase in real estate prices in OZs versus NOZs across U.S. states using the matched dataset (without filtering out the tracts with fewer than ten transactions per quarter during the pre-treatment period). States that were not included appear in gray.

A3. Impact of the OZ Program

A3.1. Propensity Score Matching

First, we describe the steps to implement the PSM below:

1. Aggregate our original dataset at the census-tract level and split it into two groups, treatment (i.e., designated OZs) and control (i.e., all NOZs or only QSN tracts).

2. Run a logistic regression specified below on the binary variable $Treatment_i$, which is equal to 1 if observation i was selected as an OZ and 0 otherwise, using explanatory variables including price, volume, income level, poverty rate, unemployment rate, population, density, and area:

$$Treatment_i \sim Logit\Big(\beta_1 \cdot Price + \beta_2 \cdot Vol + \beta_3 \cdot Inc + \beta_4 \cdot Pov + \beta_5 \cdot Unempl + \beta_6 \cdot Popul + \beta_7 \cdot Dens + \beta_7 \cdot Area\Big).$$

3. Obtain a predicted propensity score for each census tract, that is, the probability that census tract i is selected as an OZ.

4. Implement a matching algorithm to match the census tracts from the control and treatment groups. We tested several matching algorithms (e.g., nearest neighbor, caliper, genetic matching) and obtained consistent results. For conciseness, we present the results using the nearest neighbor matching.

Next, we compare the key observed characteristics between the treatment and control groups before and after matching. Figure A7 shows the histogram of several major covariates of the treatment and control groups for unmatched data (top panel) and matched data (bottom panel), corresponding to qualified census tracts. As shown in Figure A7, the covariates are considerably more balanced after the matching procedure, and as an example, the average price per square foot during the pre-treatment period is \$108.98 (resp. \$109.8) for treatment (resp. control) for the matched data.

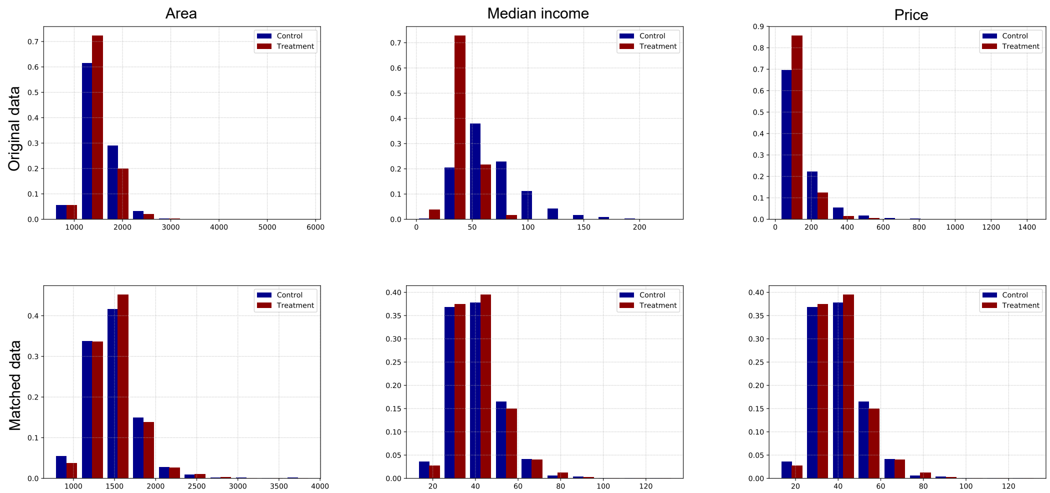


Figure A7 **Original unmatched data (top panels) versus matched data (bottom panels) based on only qualified census tracts.**

A3.2. Parallel Trends Validation

The parallel trends assumption is an essential condition for the validity of the DID approach. We present a formal test of the parallel trends assumption by estimating the pre-treatment trends on a yearly basis using Eq. (A1):

$$Y_{it} = \lambda_i + \sigma_t + \boldsymbol{X_{it}} \cdot \boldsymbol{\beta} + \sum_{k=2011}^{2017} \beta_k \cdot Treatment_i \cdot year_{kt} + \epsilon_{it}. \tag{A1}$$

Here, $year_{kt}$ is a categorical variable indicating the year of the time period t (we use 2010 as our baseline). The coefficients β_{2011} to β_{2017} identify year-by-year pre-treatment differences between price (or volume) of treatment and control groups before the launch of the OZ program.

In Table A6, we report the estimation results using the matched data based on only qualified tracts in columns (1) and (2) and using the matched data based on all tracts in columns (3) and (4). These results provide strong support for the parallel trends assumption, since the vast majority of the coefficients are statistically insignificant. Consistently, Figure A8 shows that the treatment and control groups exhibit parallel trends before the launch of the OZ program.

	Model (1) Qualified CTs Price	Model (2) Qualified CTs Volume	Model (3) All CTs Price	Model (4) All CTs Volume
Treatment × $year_{2011}$	-2.705	-0.113	-1.880	0.050
	(-1.556)	(-0.532)	(-1.009)	(-0.256)
Treatment × $year_{2012}$	-1.281	0.054	0.450	-0.070
	(-0.670)	(-0.266)	(-0.228)	(-0.331)
Treatment × $year_{2013}$	-0.303	-0.086	1.060	-0.080
	(-0.161)	(-0.388)	(-0.553)	(-0.387)
Treatment × $year_{2014}$	-4.256**	-0.077	-1.660	0.000
	(-2.379)	(-0.387)	(-0.912)	(-0.008)
Treatment × $year_{2015}$	0.061	0.112	3.190	0.010
	(-0.030)	(-0.543)	(-1.594)	(-0.038)
Treatment × $year_{2016}$	0.767	0.113	3.966*	-0.080
	(-0.396)	(-0.524)	(-1.955)	(-0.385)
Treatment × $year_{2017}$	0.999	0.427*	2.490	0.040
	(-0.487)	(-1.894)	(-1.128)	(-0.177)
No. Obs.	94,524	94,524	92,428	92,428
R^2	0.005	0.003	0.005	0.002
Controls	Yes	Yes	Yes	Yes
CT FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table A6 **Parallel trends assumption test using the matched data.**

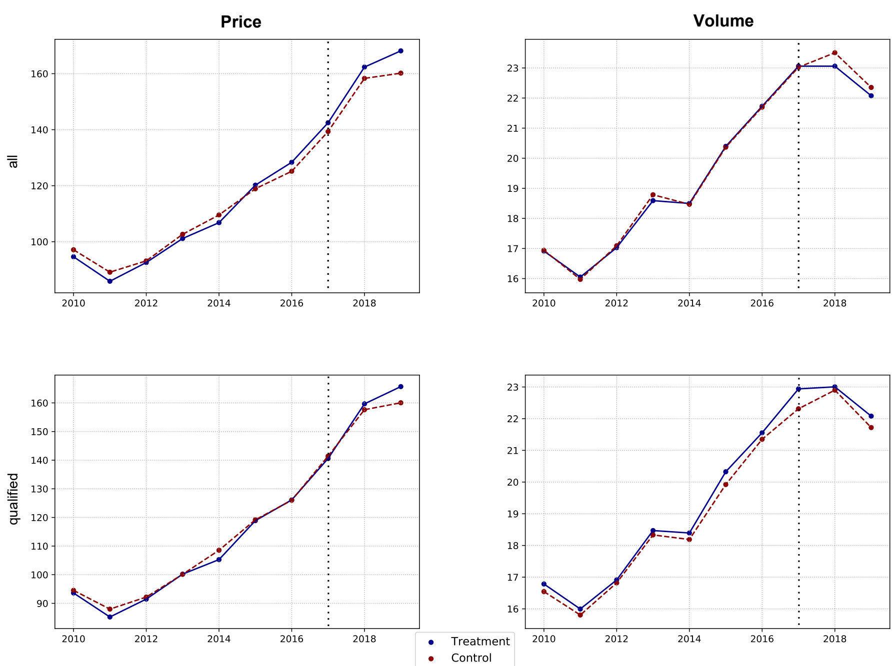


Figure A8 **Parallel-trend assumption verification using the matched data.**

A3.3. Impact of the OZ Program: Spillover-Robust DID for Volume

Table A7 shows the results when the control units are in three equal sized segments by distance and the dependent variable is volume. First, the estimates show that the average treatment effect for volume remains insignificant for the least proximate NOZs as shown in Column (3). Second, the estimated coefficient becomes significantly negative after removing segment 1, and hence we infer that the average volume should be lower in segment 1 from comparing Column (1) and (2). This indicates that the spillover effect of the OZ program should be negative. One possible explanation is that as ample investment interest and funds flow to the proximate NOZs, the property owners in those NOZs may also have strategically postponed selling their properties in the anticipation of continuous increase of demand, hence resulting in a negative spillover effect on the transaction volume. Thus, we find that our estimate of the average treatment effect in Table 3 using the standard DID estimation is an unbiased estimate.

	(1) Full Sample	(2) Segment 1 Removed	(3) Segments 1 and 2 Removed
ATT	-0.065	-0.422	-0.202
	(-0.505)	(-1.354)	(-0.643)
No. Obs.	117,282	110,938	93,210
R^2	0.003	0.021	0.021
CT FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Table A7 Effect of the OZ program on volume using matched data from spillover-robust estimation.

A3.4. Impact of the OZ Program: Census Tract Heterogeneity

In this section, we investigate whether the impact of the OZ program is heterogeneous based on pre-existing demographics or real estate characteristics in a census tract. Beyond price, we consider five other characteristics when defining $High_i$ in the estimation model as shown in Eq. (3), and we report the results in Table A8.

Moderating CT Characteristic	(1) Price	(2) Volume	(3) Median Income	(4) Area	(5) Density	(6) Population
Low (base)	-11.845***	4.890***	-0.030	14.619***	-11.270***	1.530
	(-7.576)	(2.812)	(-0.019)	(7.746)	(-7.039)	(0.900)
High	32.266***	-0.990	8.957***	-20.358***	32.504***	5.661***
	(15.649)	(-0.479)	(4.314)	(-9.808)	(15.615)	(2.728)
No. Obs.	117,282	117,282	117,282	117,282	117,282	117,282
CT FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Table A8 Heterogeneous effect of the OZ program on price using the matched data based on qualified census tracts.

Table A8 shows that the effect of the OZ program on price is stronger in higher-end (i.e., pricier) census tracts as demonstrated in Column (1). However, the extent to which the OZ program affects prices does not depend on the volume of transactions in census tracts. We also find that real estate investors are more interested in buying properties located in neighborhoods with higher medium income. Furthermore, we find that the impact of OZ program on price is stronger for census tracts with real estate properties of smaller square footage, and for denser and more populated census tracts. Overall, our analyses show that there exists heterogeneity in how the OZ program affects the price in a census tract based on its pre-existing characteristics.

A3.5. Impact of the OZ Program: Generalized Synthetic Control

In this section, we use the generalized synthetic control method proposed by (Xu, 2017) to evaluate the average treatment effect over the years post the launch of the OZ program. A major benefit of this approach, relative to using a DID with matching, is that besides observed confounders, it also accounts for unobserved time-varying confounders that may have an impact on the outcome of interest. As detailed in (Xu, 2017), the method computes counterfactuals for each treated unit using control group information based on a linear interactive fixed effects model that incorporates unit-specific intercepts (referred to as factor loadings) interacted with time-varying coefficients (referred to as factors). The estimation of the average treatment effect involves three steps: (i) estimation of the time-varying coefficients from an interactive fixed effect model using only the control group data; (ii) estimation of unit-specific intercepts for each treated unit; (iii) computation of treated counterfactuals using estimates from steps (i) and (ii), and finally the calculation of the average treatment effect from the difference between the actual outcome and the estimated counterfactual for all treated units. We note that the estimation requires that both the factors and the factor loadings are vectors of dimension r, and if not specified, the method uses cross-validation for the inference of the optimal value of r.

We use the R package gsynth to implement the generalized synthetic control method.[1] We list the details of our implementation below:

1. We consider two dependent variables, price and volume.

2. We consider the following variables in the vector of time-varying observed covariates: density, population, area, median income level, unemployment rate, price (when volume is the DV), and volume (when price is the DV).

3. We use a period of eight years before the OZ program launch and two years after.

4. We choose to use a cross-validation to find the optimal number of factors (i.e., optimal r) to include in the model.

5. We choose to use a bootstrapping process with 200 runs (as by default) to generate the standard errors for the average treatment effect estimates.

We present the results obtained by using the generalized synthetic control method on the matched data based on the qualified census tracts. We find similar qualitative results when we use the matched data based on all census tracts and the unmatched data (these results are omitted for conciseness). Table A9 shows that the average treatment effect of the OZ program on price becomes

[1] Xu Y, Liu L, Xu MY (2021) Package 'gsynth'.

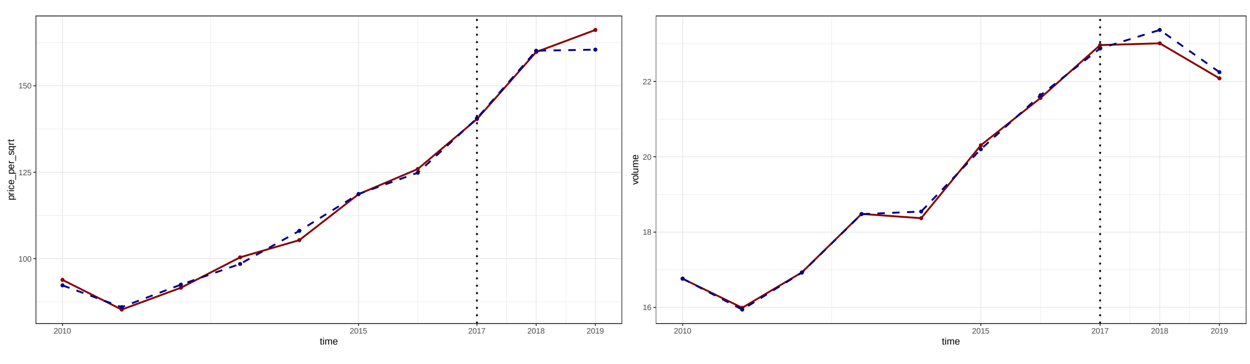
salient over time, and is significantly positive in 2019. This is consistent with the estimation result from Eq. (4) shown in Table 5. However, we do not observe a significant effect on the volume, which is once again consistent with the estimation result from Eq. (4) shown in Table A10. Figure A9 plots the yearly average of the outcome variable (price or volume) for the treatment group and the estimated counterfactual of the treatment group generated from the generalized synthetic control method. The vertical dashed line indicates the time post which the OZ program was launched. Overall, we find that the results from the generalized synthetic control method is in harmony with the DID model.

	Price	Volume
Year_2018_ATT	-0.377	-0.355
	(3.094)	(0.231)
Year_2019_ATT	5.698**	-0.163
	(2.081)	(0.244)

Table A9 Generalized synthetic control: Yearly effect of the OZ program on price and volume using the matched data. (standard deviation in parentheses)

	Volume
Year_2018_ATT	-0.148
	(-0.920)
Year_2019_ATT	0.005
	(0.032)
No. Obs.	117,282
R^2	0.019
CT FE	Yes
Time FE	Yes
Controls	Yes

Table A10 Yearly effect of the OZ program on volume using the matched data.



(a) Price per sqft (b) Volume

Figure A9 Generalized synthetic control using the matched data based on qualified census tracts (solid curve is treatment group and dashed curve is synthetic control group).

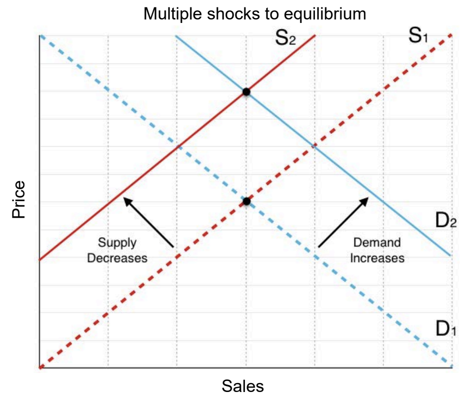


Figure A10 Illustration of supply and demand curve shocks.

A4. Fairness-Aware Optimization: Justification of Model Formulation

In this section, we provide detailed justifications of our choice of the objective function and constraints. We adopt the mean-variance framework for the objective function to accommodate the following two objectives. First, we favor an assignment that selects OZs that are the most disadvantaged. Thus, we include the means of median income and non-poverty rate to explicitly account for equity. Second, we favor an assignment that embeds equality. Our empirical analyses show that the government OZ assignment favored higher-end census tracts, and consequently, investors cherry-picked higher-end opportunity zones. Thus, to alleviate the fairness concerns raised by our empirical findings, we further include the variances of the median income and non-poverty rate to induce an assignment that delegates more homogeneous census tracts as OZs.

We note that the objective function that consists of both mean and variance was introduced by Markowitz (1968) for portfolio management, where it captures investors' goal to achieve the highest expected return along with a low-risk level (measured by the variance) via portfolio diversification. This approach has also been adopted in operations management (e.g., Van Mieghem 2003, Wei and Choi 2010) to model risk-averse agents making inventory decisions and managing supply chains. In contrast to the classical mean-variance framework where the goal is to maximize the mean while minimizing the variance, our objective is to minimize both the mean and the variance to account for equity and equality. Our choice of the mean-variance objective function also aligns with the criteria provided in Marsh and Schilling (1994), particularly appropriateness and analytical tractability.

Our MIP includes four major constraints captured in Eqs. (11)–(14). Constraint (11) ensures that the total number of OZs assigned by our model is equal to the number of actual OZs assigned by the governors. This constraint enables a fair comparison of the optimal solution obtained from our optimization relative to the government OZ designation. As described in Section 3.1, non-LICs could not account for more than 5% of the total number of OZs in the state, and constraint (12) represents a sufficient condition to satisfy this requirement. We also include constraints (13) and (14) into our MIP, which are designed to contribute to the fairness of the OZ assignment. Constraint (13) focuses on the real estate metric investment growth, which is the product of the price and transaction volume metrics previously defined. Specifically, constraint (13) ensures that the average investment growth in the census tracts selected to be OZs in a specific state does not exceed the first quartile of the investment growth of all the qualified census tracts in the state, during a four-year OZ pre-assignment period. This constraint adds to an equitable OZ assignment, as it can prevent the OZ program from majorly choosing the areas that have already drawn investors' attention or in a favored position in the real estate market as OZs. Consequently, we also expect this constraint to discourage the opportunistic investment behavior in the real estate market.

Moreover, to favor an equitable assignment from a social justice perspective, we include constraint (14) to ensure that the proportion of African American residents in the assigned OZs in a specific state is no less than the first quartile of the proportion of African American residents across all qualified census tracts in the state, using a four-year OZ pre-assignment period. This constraint can facilitate the OZ program's goal to serve distressed communities, as our data show that African Americans are disproportionately more represented in the census tracts with high poverty level. We highlight that adding constraints attributed to race in optimization problems is not new. For example, this type of constraints was used to mitigate the algorithmic bias in different business applications and public policies (Geyik et al., 2019).

A5. Fairness-Aware Optimization: MILP Reformulation

In this section, we provide another representation of the fairness-aware optimization problem based on a MILP formulation.

$$
\begin{aligned}
\textbf{MIQP:} \quad & \min_{\boldsymbol{y},\boldsymbol{x},\boldsymbol{p},\boldsymbol{q},\boldsymbol{\mu},\boldsymbol{\sigma}^2} \ \alpha_1\mu_{MI} + \alpha_2\hat{\sigma}^2_{MI} + \beta_1\mu_{NPR} + \beta_2\hat{\sigma}^2_{NPR}, \qquad \text{(A2)}\\
& \textbf{where} \quad \mu_{MI} = \Big(\sum_{i\in I} \mathit{Income}_i \cdot x_i\Big)/k,\\
& \qquad \hat{\sigma}^2_{MI} = \sum_{i\in I}\big(\mathit{Income}_i^2 x_i - 2\mathit{Income}_i\cdot\mu_{MI}\cdot x_i + p_i\big)/k,\\
& \qquad \mu_{NPR} = \Big(\sum_{i\in I} \mathit{NonPoverty}_i \cdot x_i\Big)/k,\\
& \qquad \hat{\sigma}^2_{NPR} = \sum_{i\in I}\big(\mathit{NonPoverty}_i^2 x_i - 2\mathit{NonPoverty}_i\cdot\mu_{NPR}\cdot x_i + q_i\big)/k,\\
& \textbf{subject to:} \quad x_i \in \{0,1\}, \quad \forall i\in I,\\
& \qquad y_{ij}\in\{0,1\}, \quad \forall i\in I, \forall j\in I,\\
& \qquad \sum_{i\in I} x_i = k,\\
& \qquad \sum_{i\in I_c} x_i \le k_c,
\end{aligned}
$$

$$
\begin{aligned}
(\sum_{i \in I} IG_i x_i)/k &\leq IG\_Quartile, \\
(\sum_{i \in I} BP_i x_i)/k &\geq BP\_Quartile, \\
p_i &\leq M \cdot x_i, \quad \forall i \in I, \\
p_i &\leq \sum_{i \in I} \sum_{j \in I} Income_i Income_j \cdot y_{ij}/k^2 + M \cdot (1 - x_i), \quad \forall i \in I, \\
p_i &\geq 0, \quad \forall i \in I, \\
p_i &\geq \sum_{i \in I} \sum_{j \in I} Income_i Income_j \cdot y_{ij}/k^2 - M \cdot (1 - x_i), \quad \forall i \in I, \\
q_i &\leq M \cdot x_i, \quad \forall i \in I, \\
q_i &\leq \sum_{i \in I} \sum_{j \in I} NonPoverty_i NonPoverty_j \cdot y_{ij}/k^2 + M \cdot (1 - x_i), \quad \forall i \in I, \\
q_i &\geq 0, \quad \forall i \in I, \\
q_i &\geq \sum_{i \in I} \sum_{j \in I} NonPoverty_i NonPoverty_j \cdot y_{ij}/k^2 + M \cdot (1 - x_i), \quad \forall i \in I, \\
y_{ij} &\leq x_i, \quad \forall i \in I, \forall j \in I, \\
y_{ij} &\leq x_j, \quad \forall i \in I, \forall j \in I, \\
y_{ij} &\geq x_i + x_j - 1, \quad \forall i \in I, \forall j \in I,
\end{aligned}
$$

where the last three inequalities ensure that $y_{ij} = x_i \cdot x_j$ for all $i, j \in I$.