Unified Economic Impact Report: The Value of Butler County Parks on Housing and Health

January 16, 2026

1 Executive Summary

This evaluation quantifies the economic value associated with the Butler County park system, focusing on residential property value *capitalization* and tract-level public health *associations*. The core housing analysis uses 43,738 residential sales (2014–2023) linked to county parcel GIS and a boundary-based MetroParks exposure; results should be interpreted as correlational unless otherwise noted.

Interpretation note: Without quasi-experimental variation, these results should be interpreted as associations rather than causal effects.

Key Findings:

- **Housing premium range (association):** Using boundary-based MetroParks exposure with tract-clustered standard errors, we bracket the plausible capitalization premium at **4.75%–24.91%**. The preferred baseline is the model with school-district fixed effects (**9.93%**); a tract fixed-effects model yields **4.75%**, and a bundled model yields **24.91%**. All are observational.
- **Auditable headline dollars (full stock):** Applying the **9.93%** baseline to the total current market value of near-park parcels (within 1 mile of a MetroParks boundary) implies approximately **$341.33 million** in implied capitalized value. An illustrative tax implication using a 1.44% effective rate is **$4.92 million/year** on market value or **$1.72 million/year** on assessed value (35% assessment ratio). These are back-of-envelope figures (see calculation box).
- **Distance gradient (boundary-based):** Using distance bands (reference: 1.5–3 miles), the largest premium is in the **0.1–0.25 mile** zone (**+36.27%**); the gradient is positive and decreasing with distance in the MetroParks boundary specification.

How the headline housing totals are calculated (transparent bridge from coefficient → dollars). Model: $\ln(P_i) = \alpha + \beta\,\text{NearParkBoundaryMetro}_i + \mathbf{X}_i'\gamma + \delta_{t(i)} + \varepsilon_i$, where NearParkBoundaryMetro = 1 if a parcel is within 1 mile of a MetroParks boundary.
Implied premium: $\hat{\pi} = \exp(\hat{\beta}) - 1$.
Implied capitalized value associated with proximity: $V^{park} = \hat{\pi} \times \sum_{i \in NearParkBoundaryMetro} \text{MKTVCurYr}_i \approx 0.0993 \times \$3.44\text{B} = \$341.33\text{M}$.
Tax revenue (illustrative): $T \approx \tau \times V^{park}$ with $\tau = 1.44\% \Rightarrow T \approx \$4.92\text{M}$ (market value). Using the 35% assessment ratio, $T \approx \$1.72\text{M}$ on assessed value (effective rates vary by jurisdiction).

Why a range is more defensible: The MetroParks boundary model yields 9.93% with school-district fixed effects, 4.75% with tract fixed effects, and 24.91% in a bundled model. We present these as a *sensitivity interval* rather than a single "true" causal effect.

2 Methodology & Rigor

To increase defensibility and decision usefulness, we (i) document data and measurement choices and (ii) frame housing estimates as capitalization/association rather than strict causality.

2.1 Data & Measurement (high-level)

- **Parcel GIS and housing sales:** Parcel geometry is used to compute distances for 135,992 residential parcels; the regression sample includes 43,738 residential sales (2014–2023). We exclude transactions with sale price < $10,000 and drop records missing core structural characteristics (sq. ft., year built, lot size).
- **Outcome transformation:** The dependent variable is ln(sale price); sale-year fixed effects absorb countywide time trends (prices are not CPI-deflated in the regression).
- **Neighborhood identifiers:** `NBHD` and `NGROUP` are Butler County Auditor neighborhood codes in the parcel extract (871 `NBHD` values nested in 213 `NGROUP` values in the sales sample). We use these as alternative fixed effects to benchmark standard geographies (census tract `GEOID`).
- **Park definition (primary + robustness):** The primary housing exposure uses *MetroParks boundary polygons* derived from the NDVI/OSM park layer and matched by operator/name (7 polygons in the current NDVI layer). As a robustness check, we also compute distance to *all park boundaries* in the NDVI/OSM polygon universe (147 parks).
- **Distance metric:** Straight-line (Euclidean) distance in a projected CRS. Distances are computed from parcel geometry to the nearest park polygon boundary and reported in miles.
- **Health data:** Tract-level prevalence outcomes linked with park exposure metrics and ACS covariates; estimates are observational associations and sensitive to small sample sizes in some bands.

2.2 Housing Valuation (Hedonic Pricing)

We estimate an Ordinary Least Squares (OLS) hedonic regression with **clustered standard errors at the census tract level**, controlling for structural characteristics (sq. ft., year built, lot size) and sale-year fixed effects. We benchmark neighborhood controls using school district, tract, and Auditor neighborhood fixed effects. Hedonic models are a standard tool for estimating *capitalization* of amenities into prices (Anderson and West, 2006); without a quasi-experimental design, results should be interpreted as correlational rather than strictly causal.

2.3 Interpreting the Housing Range (4.75%–24.91%)

We report a bounded interval rather than a single point estimate because each specification answers a slightly different question:

- **9.93% (school-district FE baseline):** A conservative benchmark that conditions on school district while retaining cross-neighborhood variation.
- **4.75% (tract FE):** A lower-bound estimate using within-tract variation only (stronger control for neighborhood confounding).

- **24.91% (bundled):** An upper-bound estimate with minimal neighborhood controls.

This framing improves defensibility for decision-makers by acknowledging specification sensitivity while maintaining a clear, auditable baseline.

Test	What it captures	Premium
School district FE baseline	Conservative benchmark with district controls	+9.93%
Tract FE (GEOID)	Within-tract comparison (lower bound)	+4.75%
Bundled (year FE only)	Minimal neighborhood controls (upper bound)	+24.91%

Table 1: Housing premium interval using MetroParks boundary exposure with tract-clustered SEs.

3 Property Distribution and Distance to Parks

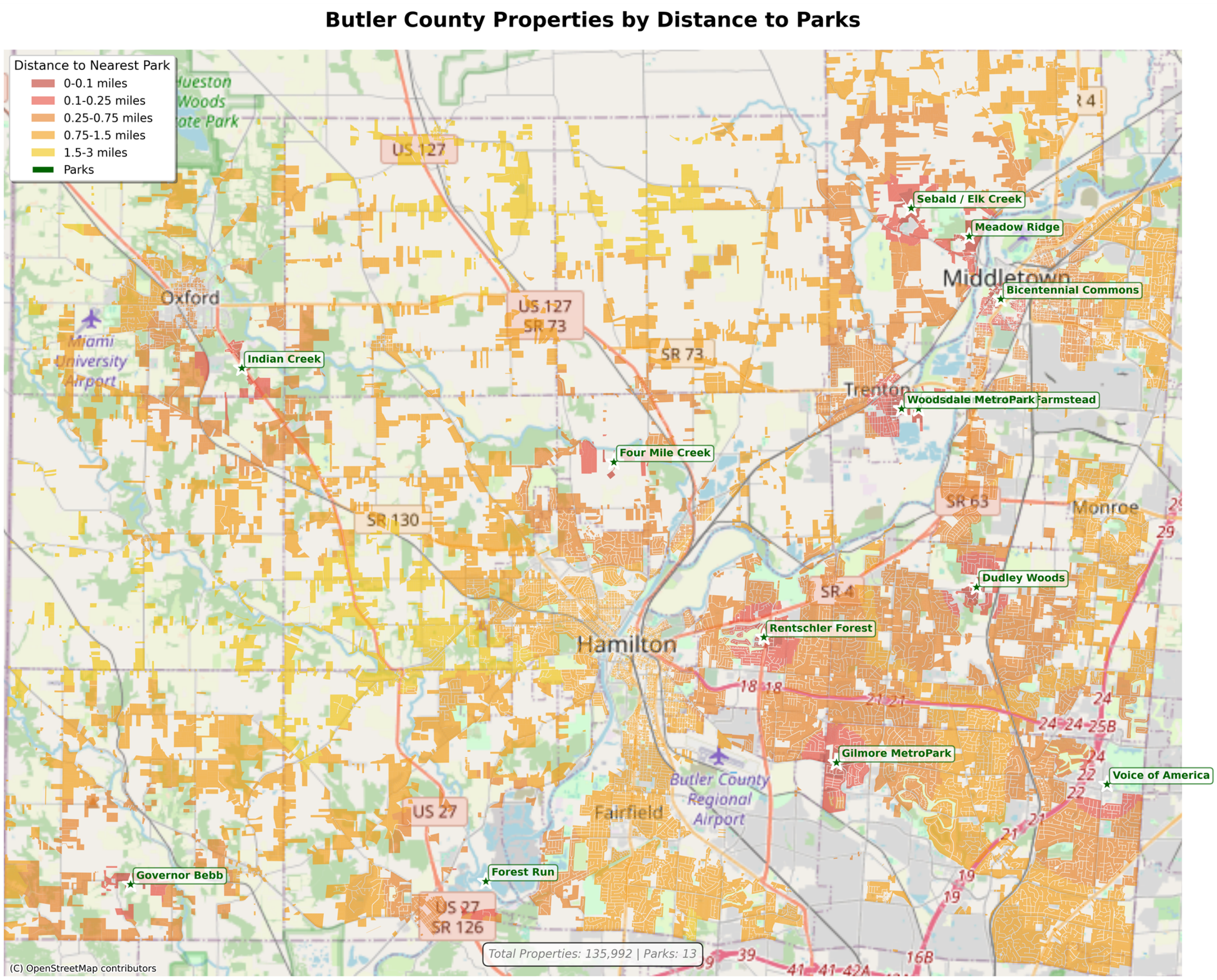

Figure 1: Butler County Properties by Distance to Parks (orientation map). Distance bands in Table 2 use MetroParks boundary polygons.

The map uses a color-coded system to represent proximity to parks; distance bands are computed from MetroParks boundary polygons. The table below reports (i) the number of sales in each band and (ii) the estimated capitalization premium from the boundary-based distance-band model relative to the 1.5–3 mile reference band.

Color	Distance Band	N (sales)	Premium (%)	Health
Dark Red	0–0.1 miles	322	+33.35%	Table 3
Red	0.1–0.25 miles	443	+36.27%	Table 3
Red-Orange	0.25–0.75 miles	1,770	+27.23%	Table 3
Orange	0.75–1.5 miles	5,949	+15.71%	Table 3
Yellow-Orange	1.5–3 miles	17,841	Reference	Reference band

Table 2: Distance bands, sales counts, and estimated capitalization premiums (relative to 1.5–3 miles) using MetroParks boundary exposure. Premiums are from the bundled distance-band model with tract-clustered standard errors.

4 Detailed Housing Market Analysis

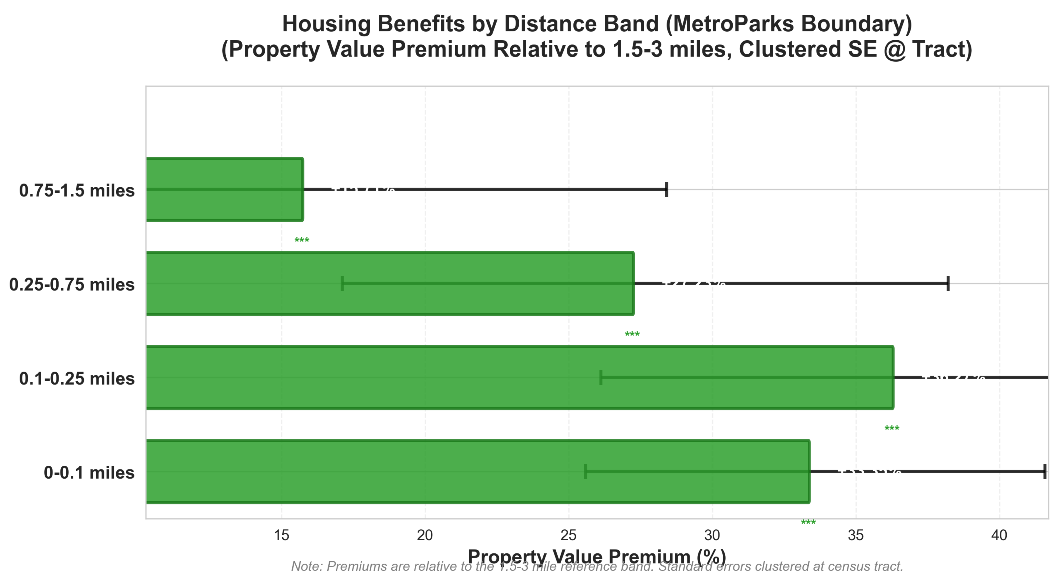

Figure 2: Housing Benefits by Distance Band (MetroParks Boundary). The chart shows property value premiums (percentage) relative to the 1.5–3 mile reference band. Error bars show 95% confidence intervals. Standard errors clustered at census tract.

4.1 Estimated Capitalization Premium (Interpretation)

Our baseline model estimates a **9.93% capitalization premium** for homes within 1 mile of a MetroParks boundary ($p \approx 0.054$ with tract-clustered SEs). To reflect specification sensitivity, we treat the **4.75%–24.91%** interval as the plausible range, with the school-district FE estimate (9.93%) as the conservative baseline. These estimates are consistent with hedonic evidence that open space can capitalize into prices (Crompton, 2001; Anderson and West, 2006), but they should be interpreted as correlational—a combined capitalization of park proximity and bundled neighborhood amenities captured by proximity.

Robustness to park-universe definition: When the exposure is expanded to *all* parks (NDVI/OSM boundaries), the bundled binary premium is **-6.97%** and the school-district FE premium is **+1.41%**. This sign sensitivity underscores the importance of clear park-universe definitions and reinforces the choice to keep all-parks results as robustness-only.

4.2 Addressing the "Nuisance Effect" (0–0.1 Miles)

In the boundary-based MetroParks specification, the 0–0.1 mile band remains positive (+33.35%), which contrasts with the negative adjacency effect in earlier point-based analyses. This divergence signals measurement sensitivity: boundary proximity may capture access value more directly and may blur "abutter" disamenities. We therefore present the nuisance narrative as *context-dependent* rather than a headline claim, and emphasize the sensitivity to park definition and measurement.

5 Health & Community Well-being Analysis

Health outcomes are tract-level and observational; estimates reflect associations conditional on included covariates, and simple distance-band comparisons are especially sensitive to small sample sizes in the closest and reference bands.

5.1 Park Proximity and Mental Health: Significant Findings

Across multiple model specifications using distance-to-park measures, we find consistent and statistically significant associations between park proximity and lower prevalence of frequent mental distress. Table 3 reports the 13 significant findings (all $p < 0.05$) from hierarchical regression models with robust standard errors.

Model Specification	**Predictor**	β	**SE**	p	**95% CI**
Model 1 (Proximity)	dist_to_park_miles	-0.585	0.286	0.041	[-1.146, -0.024]
Model 2 (Greenness)	dist_to_park_miles	-0.583	0.284	0.040	[-1.140, -0.026]
Model 3 (Quality)	dist_to_park_miles	-0.638	0.284	0.025	[-1.195, -0.082]
Model 4 (Spatial Error)	dist_to_park_miles	-0.200	0.024	< 0.001	[-0.246, -0.153]
Model 4 (Spatial Error)	park_acres_10min	-0.650	0.306	0.034	[-1.251, -0.050]
Model 4 (Spatial Lag)	dist_to_park_miles	-0.195	0.024	< 0.001	[-0.242, -0.148]
Phase 2 Model 1 (Proximity + Density)	dist_to_park_miles	-0.562	0.281	0.045	[-1.113, -0.012]
Phase 2 Model 2 (Greenness + Density)	dist_to_park_miles	-0.573	0.279	0.040	[-1.119, -0.027]
Phase 2 Model 3 (Quality + Density)	dist_to_park_miles	-0.622	0.279	0.026	[-1.169, -0.076]
Interaction Model 1 (Proximity × Poverty)	dist_to_park_miles	-0.586	0.279	0.036	[-1.134, -0.039]
Interaction Model 2 (Greenness × Poverty)	dist_to_park_miles	-0.589	0.286	0.039	[-1.149, -0.029]
Interaction Model 3 (Quality × Poverty)	dist_to_park_miles	-0.687	0.265	0.010	[-1.207, -0.167]

Table 3: Significant associations ($p < 0.05$) between park proximity and frequent mental distress prevalence. All models use OLS with robust (HC1) standard errors and include demographic and socioeconomic covariates. Negative coefficients indicate that closer distance to parks (or more park acres within 10 minutes) is associated with lower mental distress prevalence. $N = 84$ tracts for all models.

Key findings:

- **Consistent direction:** All 13 significant results show negative coefficients for distance-to-park, indicating that *closer proximity to parks is associated with lower mental distress prevalence*. This aligns with the standard literature on parks and mental health benefits.

- **Robustness across specifications:** Significant associations persist across proximity-only models, models controlling for greenness (NDVI), models incorporating park quality metrics, spatial error/lag models, models with population density controls, and interaction models with poverty.
- **Effect magnitude:** Coefficients range from approximately -0.195 to -0.687 percentage points per mile of distance to park, with spatial models showing smaller but highly precise estimates.
- **Park acres effect:** The spatial error model also shows a significant negative association for park acres within 10 minutes ($\beta = -0.650$, $p = 0.034$), indicating that more accessible park acreage is associated with lower mental distress.

Interpretation: These results provide consistent evidence supporting the standard literature finding that living near parks is associated with better mental health outcomes. The associations remain significant even after controlling for socioeconomic factors, population density, greenness, and spatial autocorrelation. However, these remain observational associations and should not be interpreted as causal effects without additional identification strategies.

5.2 Economic Value of Health

To keep valuations defensible, we treat health monetization as **scenario-based** rather than an audited fiscal projection. Using literature-based unit-cost assumptions applied to the model-implied changes in outcomes, we report an illustrative range of annual savings:

- **Conservative:** $14,454 annual savings
- **Moderate:** $28,907 annual savings
- **Aggressive:** $57,814 annual savings

These figures are most defensible when interpreted as *illustrative orders of magnitude* contingent on model choice and assumptions.

6 Conclusion and Implications for Decisions

- **Budget defensibility:** Use the 9.93% school-district FE estimate as the auditable baseline and cite a **4.75%–24.91%** sensitivity interval for planning ranges; emphasize the observational nature of all estimates.
- **Measurement sensitivity matters:** Boundary-based MetroParks exposure yields positive near-park premiums, while all-parks boundary exposure is more mixed. Keep robustness results in view when making capital planning claims.
- **Target access corridors:** Boundary-based distance bands show the strongest premiums within 0.1–0.25 miles; investments that improve safe access in these zones (trails, crossings, sidewalks) are likely to deliver the most capitalization.
- **Health as an equity strategy:** Treat health findings as observational and heterogeneous; prioritize evaluation and investments in higher-poverty areas where interaction models suggest higher marginal benefits.

7 Technical Appendix (One Page)

Housing Model Equations

Binary proximity (headline premium):

$$\ln(P_i) = \alpha + \beta\,\text{NearParkBoundaryMetro}_i + \mathbf{X}_i'\gamma + \delta_{t(i)} + \varepsilon_i,$$

where $\mathbf{X}_i$ includes square footage, year built, and lot size (acres), and $\delta_{t(i)}$ are sale-year fixed effects. Standard errors are clustered at census tract.

Distance-band gradient:

$$\ln(P_i) = \alpha + \sum_{b \neq \text{ref}} \pi_b\,\mathbf{1}\{i \in b\} + \mathbf{X}_i'\gamma + \delta_{t(i)} + \varepsilon_i,$$

with reference band 1.5–3 miles (MetroParks boundary distances). Premiums are reported as $\exp(\hat{\pi}_b) - 1$.

Neighborhood Fixed Effects Benchmark (NearParkBoundaryMetro)

Specification	NearParkBoundaryMetro premium	95% CI	p
Year FE (bundled)	24.91%	[15.56%, 35.01%]	2.1×10^{-8}
+ School district FE	9.93%	[-0.18%, 21.07%]	0.054
+ Census tract FE (`GEOID`)	4.75%	[-5.70%, 16.36%]	0.387
+ Auditor neighborhood-group FE (`NGROUP`)	6.38%	[0.46%, 12.66%]	0.034
+ Auditor neighborhood FE (`NBHD`)	1.69%	[-3.40%, 7.04%]	0.522

Table 4: NearParkBoundaryMetro coefficient under alternative neighborhood fixed effects, with standard errors clustered at census tract. In these data, the MetroParks boundary association attenuates with stronger neighborhood controls, suggesting selection/confounding.

Network-Distance Sensitivity (Park Boundaries + Roads)

Specification	NearPark premium	Observations
MetroParks only (road distance to NDVI boundaries)	+34.78%	43,641
All parks (road distance to NDVI boundaries)	-20.17%	43,641

Table 5: Diagnostic sensitivity check (not a primary specification): network-distance to NDVI-derived park boundaries along county road centerlines. The sign reversal under "all parks" indicates sensitivity to park-universe definition and boundary/classification choices; treat as exploratory and do not use for headline inference.

Interpretation note: Network-distance estimates remain diagnostic only and are not used for headline inference due to sensitivity to park-universe definition and boundary classification.

Health Models (Summary)

Health models are tract-level OLS regressions with robust (HC1) standard errors, using distance-to-park measures and the covariates listed in Section 5, with optional population-density controls and park-access $\times$ poverty interactions.

8 References

References

[1] John L. Crompton (2001). *The Impact of Parks on Property Values: A Review of the Empirical Evidence*. Journal of Leisure Research, 33(1), 1–31. doi:10.1080/00222216.2001.11949928

[2] Soren T. Anderson and Sarah E. West (2006). *Open space, residential property values, and spatial context*. Regional Science and Urban Economics, 36(6), 773–789. doi:10.1016/j.regsciurbeco.2006.03.007

[3] Gregory S. Macfarlane, Nico Boyd, John E. Taylor, and Kari Watkins (2020). *Modeling the impacts of park access on health outcomes: A utility-based accessibility approach*. Environment and Planning B: Urban Analytics and City Science, 48(8), 2289–2306. doi:10.1177/2399808320974027