Geographic Dispersion in the Marginal Returns to U.S. Healthcare Spending: Evidence from County-level Estimates

Tara Templin*, Shuyi Song†, Emily K. Johnson‡, Joseph Dieleman§

January 30, 2026

Extended Abstract

In this study we assess the marginal benefits of an additional dollar of disease-specific health spending vary across U.S. counties and clinical conditions, and what this dispersion implies for the long-standing puzzle that U.S. health care appears welfare-enhancing on average yet wasteful at the margin. Building a county–year panel that links age- and sex-standardized mortality for ischemic heart disease (IHD), type 2 diabetes (T2DM), and breast cancer to real, all-payer, per-capita spending for the same conditions, we deliver causal, condition-specific estimates of the local average treatment effect (LATE) of spending on mortality and map its cross-county heterogeneity. Our framework distinguishes movements along a concave local production function from outward shifts in the national technological frontier, allowing us to parse how much of the mortality decline reflects incremental spending versus common innovations.

Our data cover 3,110 counties observed annually and combine two sources: (i) the IHME Disease Expenditure (DEX) database, which assembles and reconciles more than forty billion claims across Medicare, Medicaid, and large commercial payers to construct disease-attributable, all-payer spending by county of patient residence; and (ii) IHME's U.S. Health Disparities estimates of age- and sex-standardized, cause-specific death rates derived from complete vital registration, with small-area methods used to address sparse cells. We analyze a balanced panel from 2010–2019 and incorporate rich, annually measured covariates (education, racial/ethnic composition, insurance coverage, and median household income).

Identification addresses the central endogeneity that counties spend more where morbidity is higher. Our instrument is per-capita spending on 'other musculoskeletal (MSK) disorders'—a residual category of non-fatal soft-tissue and connective-tissue conditions that absorbs billing and utilization shocks but has no direct biological pathway to deaths from IHD, T2DM, or breast cancer. Intuitively, MSK shocks move revenues, staffing, and equipment in integrated delivery systems, creating budgetary comovement across service lines without directly affecting cardiometabolic or oncologic mortality. We estimate two-stage least squares (2SLS) with county fixed effects and year fixed effects, so identification comes from within-county year-to-year deviations in focal-disease spending induced by MSK shocks. Placebo outcomes (accidental-injury mortality), risk-environment checks (obesity), and provider-supply checks (physicians per capita) support the exclusion restriction; lead–lag tests and geography-of-care checks address

dynamics and SUTVA. First-stage Kleibergen–Paap F-statistics exceed conventional thresholds, and instrumented estimates are robust to clustering and alternative instruments.

Three sets of results emerge. First, marginal returns are positive but modest, rejecting a strictly 'flat-of-the-curve' frontier in complier counties while falling well short of large average gains documented in the literature. Specifically, elasticities of mortality with respect to own-disease spending are −0.055 for IHD, −0.090 for T2DM, and −0.049 for breast cancer (all $p<0.01$). In other words, a 1% increase in per-capita spending for the same disease is associated with a 0.05–0.09% decline in the corresponding death rate. Fixed-effects OLS slopes are substantially closer to zero, consistent with attenuation from measurement error and negative selection (sicker counties spend more), while the IV estimates remain economically small—indicating that selection alone cannot explain the paradox of high average but low marginal returns. Distributed-lag specifications show that effects materialize within the calendar year, and cumulative multi-year responses lie within one standard error of the contemporaneous effect.

Second, heterogeneity is large and systematic. Using generalized random forests and person-centered IV treatment effects, we recover the distribution of county-level marginal products and its predictors. Educational attainment (share with a bachelor's degree) is the dominant splitter of heterogeneity, followed by insurance coverage and income, with the share having only a high-school diploma contributing the least. Maps of county-level effects reveal pronounced spatial clustering: higher marginal products are more common across much of the Southeast and parts of Middle Texas, whereas lower returns are more prevalent in the Upper Midwest, Mountain West, and pockets of the Northeast. Overall, IHD exhibits the widest dispersion (including a nontrivial left tail with near-zero or negative returns), while T2DM and breast cancer cluster more tightly around small positive gains. These patterns are predictive, not causal with respect to the covariates, and they identify where extra dollars buy more health.

Third, we provide a welfare interpretation by linking estimated marginal products of spending to values of statistical life (VSL)–consistent willingness-to-pay to form a marginal social benefit to marginal social cost ratio. Applying this framework to our county-level elasticities implies pockets of both excess and insufficiency, even within the same state. Combined with a simple decomposition of the observed mortality decline into frontier progress versus movement along the local curve, our estimates suggest that only a minority of the 2010–2019 mortality gains can be attributed to within-county spending increments, with the remainder reflecting common technological improvements (e.g., therapy innovation and diffusion) that shift the frontier outward for all counties.

This research contributes in three ways. Methodologically, we introduce a disease-specific IV that isolates real resource variation and is amenable to small-area welfare mapping; we merge this with modern causal-forest machinery to recover treatment-effect heterogeneity at scale. Empirically, we supply the first county-level, condition-specific 2SLS elasticities that reconcile positive—but limited—marginal returns with large historical average benefits, offering a unified lens on the 'flat-of-the-curve' vs. innovation-driven frontier debate. For policy, our maps identify

high-return ‘under-funded’ pockets where additional dollars likely pass the VSL test and low-return ‘over-funded’ pockets where reallocation could reduce spending with little mortality cost. Taken together, the evidence favors a re-targeting margin—rather than across-the-board cuts or increases—so that incremental health dollars are deployed benefit-cost ratio is highest, while frontier-shifting investments (e.g., clinical innovation, diffusion of proven therapies, data systems) continue to drive most aggregate health gains.

JEL: I10, I11, I12, H51, C21

* Corresponding Author. Department of Health Policy and Management, University of North Carolina at Chapel Hill; Carolina Population Center. Email: ttemplin@unc.edu.
† Department of Health Policy and Management, University of North Carolina at Chapel Hill; Carolina Population Center.
‡ Danish Centre for Health Economics, University of Southern Denmark.
§ Department of Health Metrics Sciences, University of Washington; Institute for Health Metrics and Evaluation.