Supplementary Online Content


eAppendix.
eReferences
eTable 1. Characteristics of Individuals in the Claims Data, 2003-2016.
eTable 2. Interaction Analyses
eFigure 1. Changes in Primary Care Physician Supply
eFigure 2. Changes in Specialist Physician Supply
eFigure 3. Instrumental Variable Regression Results for Cause-Specific Mortality
eFigure 4. Near-Far Matching Results for Cause-Specific Mortality
eFigure 5. County-Level Estimated Price Parities, 2015
eFigure 6. Distribution of the Instrumental Variable
eFigure 7. Regression Tree Analysis of Changes in Primary Care Physician Supply
eFigure 8. Regression Tree Analysis of Changes in Specialist Physician Supply

This supplementary material has been provided by the authors to give readers additional information about their work.
1 Mixed effects model specification

We estimated linear mixed effects models, fitting a model with both fixed-and random-effects terms from which the profiled deviance (negative twice the profiled log-likelihood) was evaluated using constrained optimization to provide the parameter estimates. The main mixed effects model was written to enable correlated random intercepts and slopes with respect to primary care physician supply grouped by area (county, in the main specification), and random intercept with fixed mean with respect to time, enabling among-area variations in associations between primary care physician supply and the outcome variable in each regression. Note that we do not assume independence of intercepts and slopes, to preserve invariance to additive shifts in the continuous predictor. The standard form of a linear mixed effects model is:

$$y = X\beta + Zb + \varepsilon,$$

in which \(y\) is the \(n\)-by-1 outcome vector, \(n\) is the number of observations (one for each area at each time period), \(X\) is an \(n\)-by-\(p\) fixed-effects design matrix, \(p\) is the number of covariates with fixed effects terms (all covariates in main text Table 1, including primary care supply), \(\beta\) is a \(p\)-by-1 fixed-effects vector, \(Z\) is an \(n\)-by-\(q\) random-effects design matrix, \(q\) is the number of covariates with random effects terms (two terms in our model: primary care physician supply and time), \(b\) is a \(q\)-by-1 random-effects vector, and \(\varepsilon\) is the independent \(n\)-by-1 observation error vector. The random-effects vector \(b\) and the error vector \(\varepsilon\) are defined with normal (Gaussian) prior distributions:

$$b \sim N(0, \sigma^2 D(\theta)),$$

$$\varepsilon \sim N(0, \sigma^2 I),$$

where \(\sigma^2\) is the error variance, \(D\) is a symmetric positive semidefinite matrix defined by variance vector \(\theta\), and \(I\) is the \(n\)-by-\(n\) identity matrix.
Across time periods $i$ (where $i = 1, 2, 3$ in the main specification) and areas $m$ (where $m = 1, 2, ..., M$ counties in the main specification), the model with a predictor variable $x$ for which the random-effects term is defined as $z$ is:

\[ y_{im} = (\beta_{00} + \beta_{10}x_{im}) + (b_{0m} + b_{1m}z_{im}) + \varepsilon_{im}, \]

in which the first parenthetical grouping reflects the fixed effects terms and the second reflects the random effects terms. The intercept and slope component elements $b_{0m}$ and $b_{1m}$ are related through a multivariate normal symmetric and positive semidefinite matrix:

\[ \{b_{0m}, b_{1m}\} \sim N(0, \Sigma), \]

\[ \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}. \]

As some counties changed boundaries over the study period, Health Resource and Service Administration guidelines were used to identify carry-overs from one designated county to newly-defined counties over the study period.  

In specifications with different area grouping variables (primary care service areas, hospital referral regions), where there are $R$ grouping variables and $m(r, i)$ indicates the level of the grouping variable $r$, the model is:

\[ y_i = x_i^T\beta + \sum_{r=1}^{R} z_{ir} b_{m(r, i)}^{(r)} + \varepsilon_i, \]

where $\beta$ is a $p$-by-1 fixed-effects vector, $b_{m(r, i)}^{(r)}$ is a $q(r)$-by-1 random-effects vector for the $r$-th grouping variable and level $m(r, i)$ and $\varepsilon_i$ is a 1-by-1 error term for observation $i$.

Continuous variables were log-transformed, centered and scaled by two standard deviations, which allows coefficients for continuous covariates to reflect the change in the outcome variable given a change in the independent variable from its mean minus one standard deviation to its mean plus one standard
deviation on the logged scale (provided in main text Table 2), correcting for right skew and enabling fair comparison of magnitudes among regression coefficients.\(^3\)

Because it is known that the null distributions of parameter estimates from mixed models are not \(t\) distributed for finite sample sizes, nor are the null distributions of differences in scaled deviances \(F\) distributed, degrees of freedom and associated \(P\) values are poorly approximated.\(^4\) Hence, following criticism of excess focus on \(P\) values in the medical literature,\(^5\) we reported 95% confidence intervals rather than \(P\) values.

We also used variance inflation factors (VIFs) and the Farrar-Glauber test to identify the potential for the precision of estimates to be decreased by multicollinearity;\(^6\) the tests indicated that the factors with problematic levels of collinearity (having VIFs>4) were: elderly and Medicare population, uninsured <65 years old and unemployed, and income and poverty. Hence, the covariates with the highest VIFs (elderly, uninsured, and income) were removed to leave their collinear variables with lower VIFs (Medicare, unemployed, and poverty), which enabled the model to have all covariates with VIF<3 in the final specification.

We additionally ran a version of our main model specification while omitting pediatricians from the definition of the primary care measure, as our outcome measures were primarily driven by mortality among adults. We found the association between primary care physician density and life-expectancy remained robust to the omission of pediatricians, with an increase of 10 primary care physicians per 100,000 people associated with a 50.9-day (95% CI: 29.4, 72.3) increase in life expectancy, net of controls (versus 51.5 days in the main specification, main text Table 2).

2 Instrumental variable model specification

An instrumental variable analysis produces an estimate of the “local average treatment effect”, which is typically larger than the overall “average
treatment effect” given by ordinary least-squares regression. The local average treatment effect reflects the association between PCP supply and each outcome among counties where physician supply was influenced by the instrumental variable, analogous to a per-protocol analysis of adherent participants in a trial (i.e., adhering to the instrument’s influence); the primary mixed model analysis is, by contrast, analogous to a more generalizable intent-to-treat analysis.

We estimated two-stage least squares regression models, fitting a model using the instrument of real net present value of the Public Service Loan Forgiveness (PSLF) program for a median indebted U.S. medical school graduate over the study period, defined as a proportion of annual county price parity. The real net present value of federal subsidies is often used as an instrument in economic assessments. We calculated the real net present value of PSLF for each county and year, using the inflation-adjusted PSLF value for the median indebted medical school graduate who has $192,000 in qualified loans, entering a job with a $140,000 starting salary, as a proportion of the county’s price parity, which reflects the purchasing ability of a dollar in the county relative to the national median.

At a conceptual level, the instrumental variable regressions were conducted under several assumptions inherent to instrumental variables analysis: (i) that the instrument in the regressions was treated as uncorrelated with unmeasured characteristics of areas that were not already captured in the instrument definition or in the other covariates included in the regressions, including future mortality (exogeneity); (ii) that the instrument was treated a reliable predictor of primary care density (had a meaningful effect, tested in the first-stage regression defined below); (iii) that the relationship between density and mortality in one area was being studied without interference from the relationship in other areas, except for explicit nested relationship structures (stable unit treatment value); (iv) that the effect of the instrument on primary care density was assumed to be equal to or greater than the effect on supply that would otherwise occur if the instrument’s value were zero (monotonicity); and (v)
any effect of the instrument on the outcome was treated as mediated by primary care density, conditional on the other observed covariates included in the regressions (exclusion restriction).

Net present value was adjusted for inflation for each study year using the Consumer Price Index, which for the median U.S. graduate with $192,000 in qualified loans entering a job with $140,000 in starting salary averaged $75,840 across the U.S. in 2015 Dollars. We projected the net present value accounting for the typical post-graduate three-year residency training of physicians going into general internal medicine, general family medicine or general pediatrics following medical school, after which a physician chooses their initial job, thus accounting for the period of loan payments covered by the PSLF. PSLF provides forgiveness for 120 monthly payments after October 1, 2007 under a qualifying repayment plan for the full amount due on monthly bills paid no later than 15 days after the due date, among persons employed full-time by a qualifying employer, not including persons with in-school status or under a grace period, deferment, or forbearance. Qualifying employers include government organizations at any level (federal, state, local, or tribal, including the U.S. military), not-for-profit organizations that are tax-exempt under Section 501(c)(3) of the Internal Revenue Code (which include Federally-Qualified Health Centers and their registered Look-Alikes), and other types of non-tax-exempt organizations who provide certain types of qualifying public services (including public health and primary health care services in a clinical setting).12

Real net present value was computed by dividing the net present value of the PSLF by estimated area price parity, which reflects the average costs of goods and services in an area divided by the national average across all such areas.13,14 The national average is set to a value of 100 so that an area’s price parity be interpreted as a percentage of the national average (e.g., an area with 14.1% higher costs than the national average has a price parity of 114.1). To derive the price parity, we used Bureau of Economic Analysis (BEA) estimates of price and expenditure levels of individual goods and services in 16 expenditure
classes (apparel, rents, and a goods class and a services class in each of the
categories of: education, food, housing excluding rents, medical, recreation,
transportation, and other), which are further subdivided into strata (e.g., “major
appliances”, under “goods”) and elementary level items (e.g., “refrigerators and
freezers”, under “major appliances”), and clusters (e.g., “refrigerators”, under
“refrigerators and freezers”). The prices for rents are obtained from the American
Community Survey, while the prices for other goods and services are estimated
from expanded Bureau of Labor Statistics (BLS) data obtained from product
sellers, as is done to construct Consumer Price Indices. The individual price
observations (~1 million observations per year) include hundreds of consumer
goods and services, often including multiple quotes for the same product from
multiple sellers. The geometric average of the prices for each type of good,
specific to outlet type and unique product, is then taken and linked to
expenditure weights designed to reflect the distribution of personal consumption
expenditures in a geographic area. The data are then allocated to counties, such
that the price parity methodology implicitly ignores within-county variations.
Finally, the data are subjected to hedonic regressions, which attempt to account
for variations in characteristics of goods and services provided, including
differences in packaging, unit size, and type of outlet from which they are sold, to
assemble an aggregate index of cost in each item stratum. Hedonic regressions
account for consumer preference variations by area (e.g., apples may be a
preferred fruit in one county, and oranges in another, so food area price parities
will account for variations in fruit preferences by location, rather than only
comparing apple prices across all areas). An outlier analysis was performed to
exclude extreme values, and missing data were imputed via chained equations
(<5% of counties with missing input data). We published these estimation details
and results previously. Estimates are mapped in SI Figure 5.

In the instrumental variables regressions, the first stage model was estimated as:
\[ d_{im} = \gamma_0 + \gamma_z \zeta_{im} + \sum_{p=1}^{P} \gamma_p x_{pim} + \mu_m + \eta_i + \varepsilon_{im}, \]

where \( d_{im} \) is the primary care physician density (the endogenous variable) in county \( m \) in year \( i \), \( \zeta_{im} \) is the instrumental variable in county \( m \) in year \( i \), \( x_{pim} \) are the \( P \) time-varying county-level covariates in main text Table 1 except for primary care physician density (all log-transformed, centered, and scaled by two standard deviations), and \( \mu_m \) and \( \eta_i \) are county- and time-fixed effects. The second stage model was estimated as:

\[ y_{im} = \beta_0 + \beta_d \widehat{d}_{im} + \sum_{p=1}^{P} \beta_p x_{pim} + \mu_m + \eta_i + \varepsilon_{im}, \]

where \( \widehat{d}_{im} \) is the predicted primary care physician density obtained from the first stage. While prior rules-of-thumb have suggested that the minimum first stage F statistic should be greater than 10, we note that if we wish Wald tests of nominal size 0.05 of hypotheses about the local average treatment effect \( \beta_d \) to have size <0.1, the first stage F statistic should be greater than 22.3. We observed a first stage F statistic of 25.6 for prediction of primary care physicians (versus 7.2 for specialists), calculated using an approach robust to heteroskedasticity and clustering. We further subjected the estimation to a conditional likelihood ratio test for \( \beta_d \) confidence that outperforms the Anderson-Rubin test and is robust to weak instruments.

3 Near-far specification

To assess the robustness of the instrumental variable analysis described in the previous section, we conducted near-far matching, a newer form of instrumental variable analysis that can improve the strength of an instrument and further reduce bias. Near-far matching mimics a matched-pair randomized trial, generating weights to match counties to be as similar as possible (“near”) in their observed characteristics, but as different as possible
(“far”) in their value of the instrumental variable. The intuition is that near-far matching helps augment the influence of an instrument; the instrument will help make one county more attractive to a primary care physician than its partner county that is otherwise similar.

We performed near-far instrumental variable analysis by 1:1 matching pairs of counties that are as similar as possible among all covariates except primary care physician supply listed in Table 1, while simultaneously being as different as possible in values of the instrumental variable specified in the prior section. Near-far matching can strengthen an instrumental variable and reduce bias in the effect size estimates, while providing benefits of matching on observed covariates—namely, reduced model dependency, nonparametric adjustment for measured confounders, and lowered mean squared error in the estimated effects of the predictor variable. By selecting a matched sample and applying an instrumental variable analysis, however, the near-far approach privileges strengthening inference in terms of reducing bias and mean squared error, while reducing generalizability in the assessment by removing unmatched counties from the national sample.

As we have detailed previously, near-far matches are controlled by two key parameters: the percent sinks (the percentage of sample to be lost as unsuitable matches due to inadequate common support for inference), and the cut-point of differentiation for the instrumental variable that specifies the difference in instrument values in the pair match below which strong penalties are enforced. We empirically searched for the percent sinks and cut-point that maximized the F statistic measuring association between the instrumental variable and primary care physician density, i.e., the percent sinks and cut-point that maximally strengthened the instrument.

Specifically, a distance matrix was constructed between every pair of counties, yielding a 2N-by-2N distance matrix that reflected both the similarity of counties on the measured covariates and differences in the instrumental variable, using rank-based Mahalanobis distance to limit the impact of any
To get the closest covariate balance between the two groups, and at the same time achieve maximal separation in the values of the instrumental variable, some counties were not matched to other counties but to \( e \) sinks, where each sink is at zero discrepancy to each county and at infinite discrepancy to all other sinks, yielding a \((2N + e)\)-by-\((2N + e)\) discrepancy matrix. As detailed previously by Baiocchi et al.,\textsuperscript{21} an optimal match will pair \( e \) counties to the \( e \) sinks to minimize the total of the remaining discrepancies within \( N - e/2 \) pairs of \( 2N - e \) counties (analogous to choosing inclusion/exclusion criteria in a randomized trial).

We used simulated annealing\textsuperscript{30} to find the percent sinks and cut-point maximizing the partial F statistic from the first stage regression of primary care physician density on the instrumental variable and the above-specified measured confounders, applying a nonbipartite matching algorithm to divide the \( 2N \) individuals into nonoverlapping pairs of counties to minimize the sum of the discrepancies within the \( N \) pairs.\textsuperscript{31}

A total of 2,228 matched pairs of counties were obtained through the near-far matching process (29\% sinks) at a cut-point of the instrumental variable of 25\% of the instrument’s median value (SI Figure 6), producing post-match absolute standardized differences across all covariates of <0.2 and an adjusted first stage F statistic of 154.8.

An increase of 10 primary care physicians per 100,000 people was associated with an 84.8-day (95\% CI: 15.9, 153.8) increase in life expectancy in the near-far analysis (as compared to an 88.9-day increase in the main instrumental variable specification described in the previous section). The near-far analysis also found positive associations between primary care physician supply and improved cardiovascular, cancer, and respiratory mortality (SI Figure 4).
4 Individual-level analysis specification

We estimated the change in restricted mean survival time (RMST) conditional on changes in area-level physician density. The RMST was estimated using the approach previously derived by Tian et al. in which the RMST is defined as the area under the curve of the survival function up to a time $\tau$ (10 years in our analysis):

$$\mu_\tau = \int_0^\tau S(t)dt,$$

where $\mu_\tau$ is interpreted as the mean survival time for individuals followed for up to $\tau$ time (subject to censoring, as detailed below), and $S(t)$ is the time-to-death survival function. We estimate $S(t)$ using the Kaplan-Meier estimator. We preferred the RMST to the median survival time as an outcome metric, as the latter becomes degenerate if there is substantial censoring or if the mortality rate is low, such that the Kaplan-Meier curve does not reach 0.5. The RMST calculation also avoids the proportional hazards assumption upon which Cox models depend. Tian et al. consider the regression model:

$$g\{E(Y|Z, X)\} = \alpha + \beta Z + \gamma'X,$$

where $g(\cdot)$ indicates a smooth and strictly increasing link function, $Y$ is the estimated restricted mean survival time, $Z$ indicates the primary care density, and $X$ are the additional covariates. They show that we can estimate the restricted mean survival time under this regression model by adjusting for censoring using an inverse probability censoring weighted estimating function of $\beta$ defined by:

$$S_n(\beta) = n^{-1} \sum_{i=1}^n \frac{\tilde{\Delta}_i}{G(Y_i)} X_i \{Y_i - g^{-1}(\beta'X_i)\},$$

where for $i$ individuals $1,\ldots, n$, $\tilde{\Delta}_i = I(Y_i \leq C)$, $C$ is the censoring time, and $\hat{G}(\cdot)$ is the Kaplan Meier estimator of the censoring time. We estimated the RMST
conditional on primary care density using the identity link and conditioning on all Table 1 covariates, fixed effects for county and year, and the individual-level covariates of year of birth and sex (see SI Table 1 for participant characteristics). All covariates were log-transformed, centered and scaled by two standard deviations.

In a subgroup analysis, we additionally analyzed individuals who moved between areas \((n = 664,443)\).\(^{35}\) We isolated the analysis to those with two zip codes during their period of enrollment, and calculated the exposure \(Z\) and all covariates \(X\) as their changes in values weighted by the duration of exposure in each area \((t_2\) for the second locale and \(t_1\) for the first). Hence, the regression model became:

\[
g\{E(Y|Z', X')\} = \alpha + \beta(Z_2t_2 - Z_1t_1) + \gamma'(X_2t_2 - X_1t_1).
\]

Among those who moved from one zip code to another, survival time increased by 61.7 days (95% CI: 61.1, 62.5) per decade of exposure to 10 more primary care physicians per 100,000.

5 Regression tree analyses

As a supplementary analysis, we investigated predictors of gain or loss in primary care and specialist physician supply among counties using recursive partitioning (regression tree models).\(^{36}\) Recursive partitioning is an approach to explaining variance in an outcome, in this case defined as the change in density of physicians per 100,000 persons within a county between 2005 and 2015. The partitioning algorithm seeks the covariate that can divide the total sample of counties into two subgroups with high between-group variance and low within-group variance in the outcome. The algorithm then finds a second covariate to divide each of those two subgroups (independently), and so on. We implemented the recursive partitioning approach on the outcome against the change in each covariate in Table 1 of the main text between 2005 and 2015, and “pruned” the
resulting tree to reduce the risk of overfitting the data by selecting a complexity parameter associated with the lowest cross-validated error.\textsuperscript{36}

Regression tree analyses revealed analogous combinations of county-level factors predictive of gains or losses in primary care versus specialty physician supply, with decreases in specialist density and hospital beds being predictive of losses in primary care physicians, and decreases in primary care physician density and hospital beds being predictive of losses in specialty physicians (SI Figures 7 and 8).

6 Interaction analyses

As an additional supplementary analysis, interaction terms to identify heterogeneous associations were added to the main mixed effects model specifications to identify whether the association between primary care density and outcomes systematically differed by urban versus rural counties, or by the county’s level of poverty, Black population fraction, or Hispanic population fraction. The interaction analyses suggested stronger associations between primary care physician supply and improved life expectancy in urban counties with larger minority populations; cancer and respiratory mortality also disproportionately improved with primary care supply in higher-poverty areas (SI Table 2).
References


22. Basu S, Meghani A, Siddiqi A. Evaluating the Health Impact of Large-Scale


33. Tian L, Zhao L, Wei LJ. Predicting the restricted mean event time with the subject’s baseline covariates in survival analysis. *Biostatistics.* 2014;15(2):222-233.
34. Kim D, Uno H. Restricted mean survival time as a measure to interpret clinical trial results. *JAMA Cardiol.* 2017.


SI Table 1: Characteristics of individuals in the claims data, 2003-2016.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall studied population, ( N )</td>
<td>1,505,554</td>
</tr>
<tr>
<td>Year of birth, mean (range)</td>
<td>1944 (1927, 2016)</td>
</tr>
<tr>
<td>Sex, % female</td>
<td>44.1</td>
</tr>
<tr>
<td>Distinct residential zip codes, No.</td>
<td>18,057 (of 29,791 possible)</td>
</tr>
<tr>
<td>States, No.</td>
<td>50 + Washington D.C. + Puerto Rico</td>
</tr>
<tr>
<td>Residence in metro area, %</td>
<td>90.0</td>
</tr>
<tr>
<td>Days of enrollment, mean (range)</td>
<td>952.5 (27.0, 5,113.0)</td>
</tr>
<tr>
<td>Crude mortality rate, per 1,000</td>
<td>22.3</td>
</tr>
<tr>
<td>Movers between two areas, ( N )</td>
<td>664,443</td>
</tr>
<tr>
<td>Year of birth</td>
<td>1944 (1927, 2016)</td>
</tr>
<tr>
<td>Sex, % female</td>
<td>43.8</td>
</tr>
<tr>
<td>Distinct residential zip codes, No.</td>
<td>15,412 (of 29,791 possible)</td>
</tr>
<tr>
<td>States, No.</td>
<td>50 + Washington D.C. + Puerto Rico</td>
</tr>
<tr>
<td>Residence in metro area, % (initial)</td>
<td>75.2</td>
</tr>
<tr>
<td>Residence in metro area, % (final)</td>
<td>75.6</td>
</tr>
<tr>
<td>Days of enrollment, mean (range)</td>
<td>901.2 (27.0, 5,113.0)</td>
</tr>
<tr>
<td>Crude mortality rate, per 1,000</td>
<td>22.3</td>
</tr>
</tbody>
</table>
SI Table 2: Interaction analyses.

Results of mixed effects regressions relating primary care physician density and county-level covariates to county-level age-standardized life expectancy (N=3,142 U.S. counties), 2005-2015. The table displays the coefficient of the interaction term between the density of primary care physicians and the covariate of interest in the leftmost column. All continuous variables are log-transformed, and the coefficients can be interpreted as the disproportionate association between the outcome variable and a change in primary care density from 1 standard deviation minus the mean to 1 standard deviation plus the mean (from 27.3 physicians to 33.7 physicians per 100,000 people) given a change in the covariate of interest from 1 standard deviation below to 1 standard deviation above the mean (e.g., 1.4% to 7.9% Black population in the county).
<table>
<thead>
<tr>
<th>Covariate of interest</th>
<th>Coefficient of interaction term between primary care physician density and covariate of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Life expectancy (days)</td>
</tr>
<tr>
<td>Metro area, change to non-metro area</td>
<td>-94.6 (-133.1, -58.0)</td>
</tr>
<tr>
<td>Population in poverty, change from mean minus 1 SD (13.4% poverty) to mean plus 1 SD (16.4% poverty)</td>
<td>18.5 (-7.5, 44.5)</td>
</tr>
<tr>
<td>Population Black, change from mean minus 1 SD (1.4% Black) to mean plus 1 SD (7.9% Black)</td>
<td>71.2 (37.6, 104.7)</td>
</tr>
<tr>
<td>Population Hispanic, change from mean minus 1 SD (2.8% Hispanic) to mean plus 1 SD (7.8% Hispanic)</td>
<td>70.4 (40.2, 100.5)</td>
</tr>
</tbody>
</table>
SI Figure 1: Changes in primary care physician supply.

Changes in (A) density (physicians per 100,000 population) and (B) absolute supply (physician number) by county-level features (N=3,142 U.S. counties), 2005-2015. Low versus high poverty, Black race percentage, and Hispanic ethnicity percentage were defined as less or greater than the national median, which was 16% of the population below the national poverty threshold for low vs. high poverty, 9% of the population of Black race for low vs. high Black race, and 8% of the population of Hispanic ethnicity for low vs. high Hispanic ethnicity. Boxes reflect the interquartile range, horizontal lines reflect the median, and whiskers extend to the extremes.

(A)
SI Figure 2: Changes in specialist physician supply.

Changes in (A) density (physicians per 100,000 population) and (B) absolute supply (physician number) by county-level features (N=3,142 U.S. counties), 2005-2015. Low versus high poverty, Black race percentage, and Hispanic ethnicity percentage were defined as less or greater than the national median, which was 16% of the population below the national poverty threshold for low vs. high poverty, 9% of the population of Black race for low vs. high Black race, and 8% of the population of Hispanic ethnicity for low vs. high Hispanic ethnicity. Boxes reflect the interquartile range, horizontal lines reflect the median, and whiskers extend to the extremes.
(B)

[Diagram showing changes in specialists' physicians (2005 to 2015) for various categories: Urban, Rural, Low poverty, High poverty, Low Black %, High Black %, Low Hispanic %, High Hispanic %].
SI Figure 3: Instrumental variable regression results for cause-specific mortality.

Changes in cause-specific mortality associated with an increase in 10 primary care physicians per 100,000 people, N=3,142 US counties, 2005-2015.

Instrumental variables regressions

- **Cardiovascular**: -84.4 [-135.1, -33.7]
- **Cancer**: -18.3 [-30.1, -6.5]
- **Respiratory**: -36.6 [-49.4, -23.7]
- **Infectious**: 14.1 [-4.8, 33.0]
- **Substance/injury**: 4.9 [-4.7, 14.6]
SI Figure 4: Near-far matching results for cause-specific mortality.

Changes in cause-specific mortality associated with an increase in 10 primary care physicians per 100,000 people, N=3,142 US counties, 2005-2015.

Near-far matching results

<table>
<thead>
<tr>
<th>Cause</th>
<th>ΔMortality per million</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular</td>
<td>-53.0 [-100.2, -5.9]</td>
</tr>
<tr>
<td>Cancer</td>
<td>-56.2 [-79.3, -33.2]</td>
</tr>
<tr>
<td>Respiratory</td>
<td>-12.3 [-23.5, -1.0]</td>
</tr>
<tr>
<td>Infectious</td>
<td>15.6 [-7.4, 38.6]</td>
</tr>
<tr>
<td>Substance/injury</td>
<td>25.6 [-16.6, 67.8]</td>
</tr>
</tbody>
</table>

© 2019 American Medical Association. All rights reserved.
SI Figure 5: County-level estimated price parities, 2015.
SI Figure 6: Distribution of the instrumental variable.

Log-transformed value of the instrument across counties, 2015.
SI Figure 7: Regression tree analysis of changes in primary care physician supply.

Results illustrate the predictive variables that best separate variation in changes in primary care physician supply between 2005 and 2015 at the county level. All predictive variables refer to changes in county-level covariate values between 2005 and 2015 as defined in main text Table 1.
SI Figure 8: Regression tree analysis of changes in specialist physician supply.

Results illustrate the predictive variables that best separate variation in changes in specialist supply between 2005 and 2015 at the county level. All predictive variables refer to changes in county-level covariate values between 2005 and 2015 as defined in main text Table 1.