Which County Characteristics Predict Medical Debt?

Fredric Blavin, Breno Braga, and Anuj Gangopadhyaya

June 2022

Medical debt is a critical challenge to Americans' financial stability and well-being. People with medical debt are likely to forgo needed medical care, have difficulty meeting other basic needs, and face an increased risk of bankruptcy (Dobkin et al. 2018; Hamel et al. 2016). Recent evidence shows that hospital systems are also becoming increasingly more aggressive in collecting payments by filing lawsuits and garnishing wages (Cooper, Han, and Mahoney 2021). Attention to the risk of being burdened with high medical bills, even among those with health insurance, is growing, as shown by the passing of the No Surprises Act in 2020. The act aims to provide federal protections for consumers against surprise out-of-network medical bills.[1]

Although medical debt has gained national focus, more attention to the geographic distribution and drivers of this debt is needed. This brief fills this gap by answering two research questions:

- In which counties are people more likely to have medical debt in collections?
- Which county-level socioeconomic and health conditions better predict medical debt?

Using unique Urban Institute credit bureau data from August 2021 on more than 10 million consumers, we find that **counties in Georgia, North Carolina, and Texas have the largest shares of adults unable to pay their medical bills on time**. In fact, among the 100 US counties with the highest levels of medical debt in collections, 34 are in Texas, 20 are in Georgia, and 12 are in North Carolina. These three states are among the 12 that have not adopted the Medicaid expansion under the Affordable Care Act (ACA). In fact, 79 of the 100 counties with the highest levels of medical debt were in states that have not expanded Medicaid under the ACA.

To better investigate how medical debt relates to a county's socioeconomic status and health, we combine our credit bureau data with a rich set of publicly available data at the county level. We find that **a county's prevalence of chronic conditions is the strongest predictor of medical debt in collections**. We also find that counties with high shares of uninsured, low-income, or Black populations have higher rates of medical debt in collections. However, **a county's racial composition and average income are weaker predictors of medical debt when accounting for differences in chronic condition prevalence and other health measures, suggesting that health status is a potential factor through which racial and ethnic differences in medical debt emerge.**

While identifying causal mechanisms is beyond the scope of this brief, we offer a few possible explanations for the strong correlation between a county's population health and medical debt in collections. First, counties where the population is in poor health have higher demand for medical care. Residents of these counties are more likely to use medical services and therefore more likely to have medical bills. Second, some reverse causation is also possible, where residents with medical debt could forgo needed health care and face additional barriers to paying for food, housing, and other basic needs, putting them at risk of worse health (Adams et al. 2021; Hamel et al. 2016; Rabin et al. 2020). Recent evidence shows that medical debt is a social determinant of health (Mendes de Leon and Griggs 2021).

This brief advances the literature on medical debt in a few ways. The literature mostly focuses on the effect of health insurance on medical debt (Caswell and Waidmann 2017; Finkelstein et al. 2012; Hu et al. 2018). Past work has also shown that financially knowledgeable people have a lower risk of past-due medical debt (Braga et al. 2017). The Consumer Financial Protection Bureau (CFPB) shows that Black and Hispanic[2] people, young adults, and people with low incomes are more likely to have medical debt (CFPB 2022). We investigate the associations of medical debt with a much broader set of local characteristics. The closest study to our work is a recently published paper by Kluender and colleagues (2021), which uses credit bureau data to measure medical debt nationally and by geographic region from January 2009 to June 2020. The study finds that medical debt is more concentrated in the South and among residents of low-income zip-codes. We advance this work by focusing on the share of adults with reported medical debt in collections in a county, correlating medical debt with a more comprehensive set of local characteristics and using more recent data (August 2021), which includes debt acquired during the COVID-19 pandemic.

BOX 1

Data and Definitions

We use credit bureau data consisting of a random four percent sample of depersonalized consumer data (more than 10 million consumers) from a major credit bureau. Consumers were chosen based on the last two digits of their personal identification numbers (assigned by the credit bureau for internal use). The information was collected in August 2021. All records were stripped of personally identifiable information, and no data on race and ethnicity, gender, or income were included.

The primary outcome we study is medical debt in collections, meaning medical debt that was sent to a third-party collector or assigned to a creditor's internal collections department. Debt in collections is normally at least 180 days past due, where the creditor acted to collect the unpaid debt but was unsuccessful. Collections are defined as medical if the original creditor is identified as a medical provider. Medical bills that are paid using a credit card but ultimately go to collections would not appear as medical debt in collections. Two important limitations of our data are that (1) we do not observe flows of new collections—only the stock (collection balance)—and (2) a debt appears at the time it was reported to the credit bureau as being in collections, not at the time the debt was initially incurred.

Based on the consumer's residential zip-code, we estimate the *share with medical debt in collections for each county*—defined by the share of people in the county with a credit bureau record who have medical debt in collections. Medical debt statistics are not reported when they are based on fewer than 50 people.

We supplement our data with county-level socioeconomic and health status measures from different sources. We obtain data on racial and ethnic composition for each county from the 2020 Census. Data on average household incomes are from the American Community Survey (ACS). This information is only available for all counties as an average over five years; we use years 2015 through 2019. Data on the share of votes for Biden and Trump in the 2020 election come from the MIT Election Lab.

We use chronic condition measures among the Medicare population as a proxy for the county's health status as a whole. We obtain 2018 county-level data on Medicare spending per capita and the prevalence of chronic conditions (0–1, 2–3, 4–5, and 6+ conditions) among Medicare beneficiaries using the Multiple Chronic Conditions[a] dataset from CMS. The Multiple Chronic Conditions data provide state and county information on the number of chronic conditions among Original Medicare beneficiaries. The dataset includes information on condition prevalence, utilization of care, and spending organized by the four categories of chronic condition counts. These chronic conditions come from a select set of 21 diagnosed chronic conditions[b] in Medicare claims data.

Using the 2018 survey data from the BRFSS, we provide evidence that Medicare beneficiary health status serves as a good proxy for the health status of the nonelderly population at the state level. We calculate the average number of chronic conditions by state among respondents ages 65 and older and, separately, for adults ages 18 to 64. We find a strong association between the prevalence of chronic conditions in adults older and younger than age 65 by state (figure A.1), with a correlation coefficient of 0.85.

To further validate that our results are not solely based on the elderly population's health, we also estimate models controlling for the disability rate among the 18- to 64-year-old population using data from the 2015–19 ACS. The ACS disability measure[d] captures the presence of any one of the following six disability types: hearing difficulty, vision difficulty, cognitive difficulty, ambulatory difficulty, self-care difficulty, and independent living difficulty.

Although County Health Rankings[c] data include health measures among all adults, most of these measures (e.g., fair or poor health status and prevalence of chronic conditions) are imputed from state-level Behavioral Risk Factors Surveillance System (BRFSS) data, using the same county-level characteristics (e.g., age and race and ethnicity) in our regression model.

We obtain 2013–19 average county-level data on the percentage of live births with low birthweight from the National Center for Health Statistics natality files. High rates of low birthweight are correlated with preterm birth, which in turn is correlated with variation in maternal health and risk factors as well as access to pre-pregnancy health care.

Table A.1 summarizes the different sources and metrics we use in this analysis.

[a] "Multiple Chronic Conditions," Centers for Medicare and Medicaid Services (CMS), last updated 2018, https://data.cms.gov/medicare-chronic-conditions/multiple-chronic-conditions.
[b] "Chronic Conditions," CMS, last updated December 1, 2021, https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Chronic-Conditions/CC_Main.
[c] "County Health Rankings and Roadmaps," accessed May 9, 2022, https://www.countyhealthrankings.org/.
[d] "How Disability Data are Collected from The American Community Survey," US Census Bureau, last updated November 21, 2021, https://www.census.gov/topics/health/disability/guidance/data-collection-acs.html.

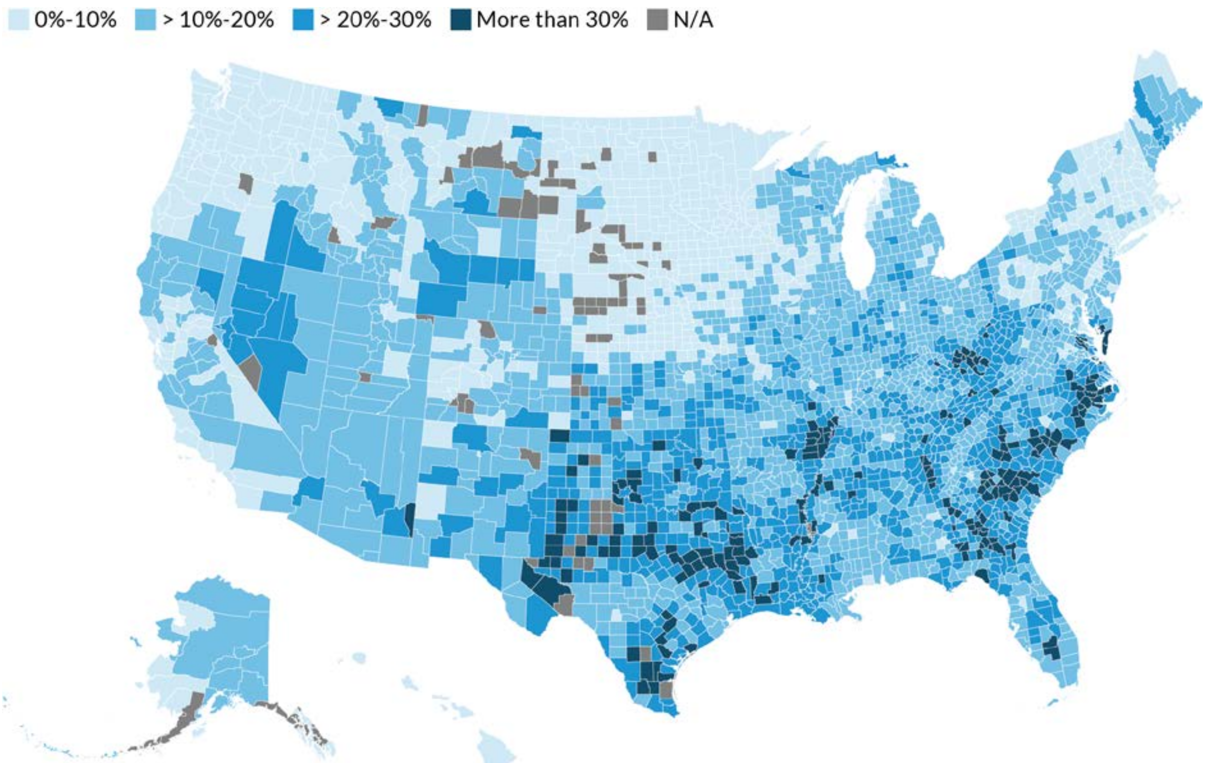
The Distribution of Medical Debt across the US

How Is Medical Debt Geographically Distributed across the US, and Which Counties Have the Highest Concentrations of Medical Debt in Collections?

Consistent with past work (CFPB 2022; Karpman and Caswell 2017; Kluender et al. 2021),[3] we find that counties with significant medical debt in collections are not evenly distributed across the US and are heavily concentrated in the South. Figure 1 shows that the share of US consumers (with a credit file) with medical debt in collections varied significantly by county in August 2021. The darkest areas are counties where 30 percent or more consumers have a report of medical debt in collections. In the lightest areas, less than 10 percent of residents have medical debt in collections.

FIGURE 1

Percentage of Consumers with Medical Debt in Collections, August 2021



Source: Urban Institute Analysis of August 2021 credit bureau data.
Note: N/A = not available because the sample size is too small.

We find that 99 out of the 100 counties with the largest shares of adults unable to pay their medical bills are located in the South—including 34 in Texas, 20 in Georgia, 12 in North Carolina, and 11 in South Carolina.[4] Seventy nine out of these 100 counties with highest levels of medical debt are in states that have not expanded Medicaid under the ACA, and 12 are in states that expanded Medicaid more than five years after the major ACA coverage provisions officially went into effect.

On average, more than 36 percent of consumers in these counties have a medical debt in collections recorded in their credit report compared with only 14 percent nationally (table 1). The populations of these 100 counties are more likely to be without health insurance (14.8 percent uninsured versus 8.8 percent nationally) and more likely to have six or more chronic conditions (20.5 percent versus 17.7 nationally). Finally, these counties have lower incomes on average ($57,825 versus $88,607 nationally) and a higher share of non-Hispanic Black residents (23.6 percent versus 12.1 percent nationally).

Table 1 also provides socioeconomic and health information on the ten counties with the highest shares of adults with medical debt in collections. Key highlights include the following:

- **The top 10 counties with the highest rates of medical debt also have high shares of people without health insurance coverage**. For example, more than 18 percent of the populations in Brooks County, Georgia, and Haskell County, Nolan County, and Pecos County in Texas do not have health insurance coverage compared with 9 percent of the whole country.
- **All top-10 counties also have an average household income lower than the national average**. For example, the average income in Haskell County, Texas, is about $49,000 while the average American household earns almost $89,000.
- **Most top-10 counties with high rates of medical debt also have higher shares of Hispanic or Black people compared with the national average**. More than 25 percent of the populations in four counties (Harmon, Haskell, Nolan, and Pecos) are Hispanic compared with 19 percent nationally, while more than 30 percent of the populations in the other six counties (Anson, Brooks, Greene, Lenoir, McDuffie, and Warren) are non-Hispanic Black compared with 12 percent nationally.
- **Counties with high rates of medical debt in collections also typically have higher chronic condition prevalence than the national average**. More than 24 percent of the Medicare population in Nolan County, Texas, have six or more chronic conditions, compared with about 18 percent of the entire Medicare population in the US.

TABLE 1

Counties with the Highest Share of Consumers with Medical Debt in Collections as of August 2021 and the Counties' Characteristics

County	State	Pop.	% with medical debt in Collections	% Uninsured	Avg. Income	% Hispanic	% Black non-Hispanic	% 6+ CCP
Warren	GA	5,215	50.5	13.0	$53,077	1.0	58.4	20.3
Greene	NC	20,451	46.0	16.6	$53,007	14.4	35.2	17.3
Lenoir	NC	55,122	44.7	12.5	$56,708	7.9	40.0	20.3
McDuffie	GA	21,632	43.1	12.1	$55,341	3.7	40.0	19.7
Anson	NC	22,055	41.6	11.1	$52,077	3.0	44.6	19.5
Nolan	TX	14,738	40.9	19.0	$64,120	36.3	4.2	24.5
Pecos	TX	15,193	40.8	18.1	$68,797	71.4	3.3	16.4
Brooks	GA	16,301	40.7	18.1	$60,621	5.9	34.9	23.7
Haskell	TX	5,416	40.6	20.8	$49,230	25.4	3.3	17.2
Harmon	OK	2,488	40.3	15.2	$65,261	29.7	6.0	22.8
Average top 10			42.9	15.7	$57,824	19.9	27.0	20.2
Average top 100			36.9	14.8	$57,825	19.2	23.6	20.5
US			13.9	8.8	$88,607	18.7	12.1	17.7

Sources: Urban Institute Analysis of August 2021 credit bureau data combined with county-level characteristics (see table A.1 for additional details).

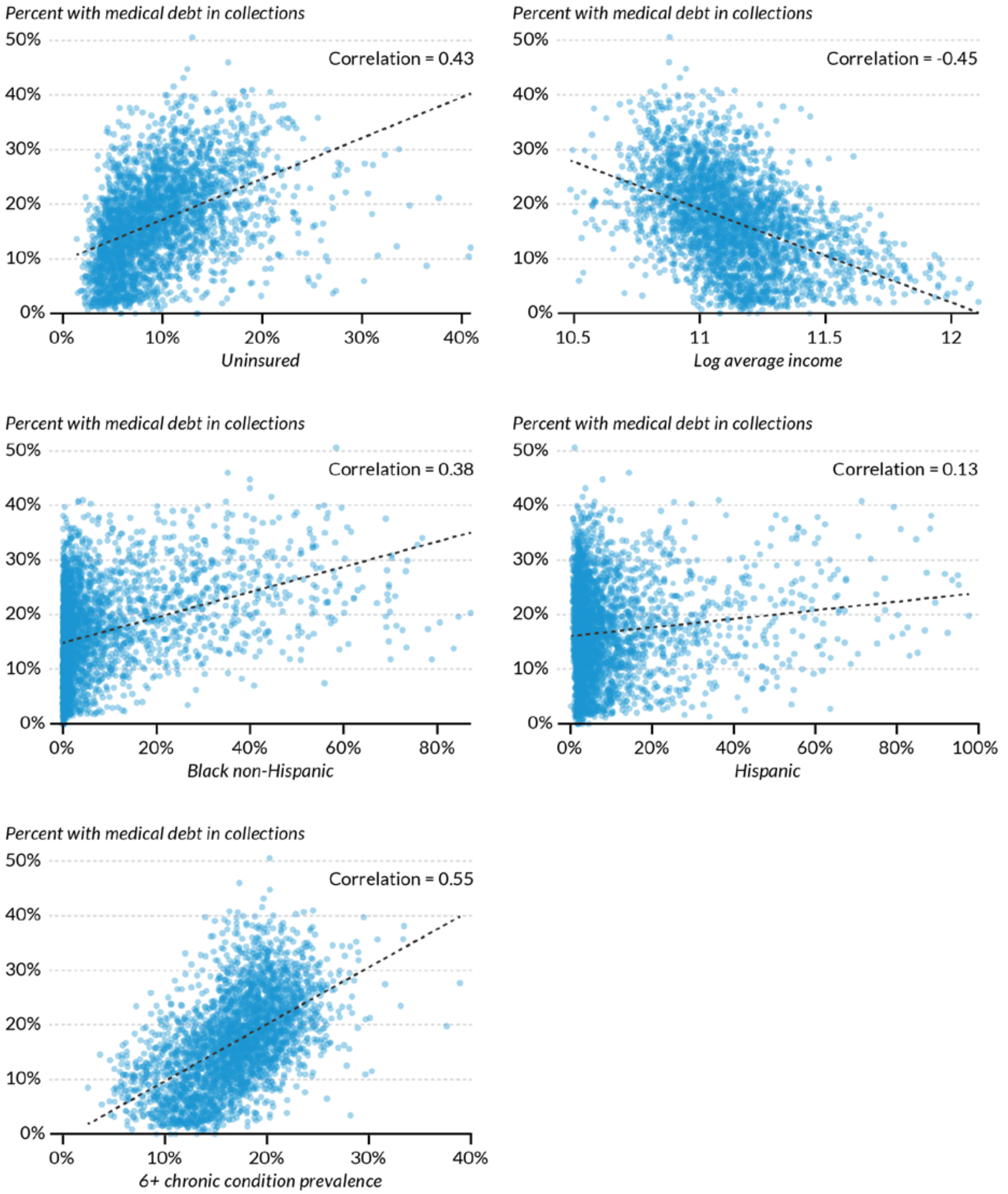
Notes: Pop. = population, CCP = chronic condition prevalence.

How Does Medical Debt in Collections Relate to County Characteristics?

In figure 2, we use scatterplots to expand on the analysis to systematically assess the correlation between the percentage of people with medical debt in collections (y-axes) and other county characteristics (x-axes). Each dot represents a county and we include a line representing the best linear fit. We find the strongest correlation (0.55) between medical debt in collections and the share of the Medicare population with six or more chronic conditions. In other words, counties with high medical debt have a higher share of patients with multiple chronic conditions (e.g., diabetes) and costly medical needs. We find a moderate positive correlation of 0.43 between the share of people with medical debt in collections and the share of the uninsured population. There is also a negative correlation between medical debt in collections and (logarithm of) average income (-0.45) and a positive correlation between medical debt in collections and the share of the non-Hispanic Black population (0.38). Overall, racial inequities in income, wealth, and access to health care—the result of historical and structural factors—likely play a role in higher rates of medical debt by having negative consequences on population health and making it more difficult to pay for care. The correlation between medical debt in collections and the share of the Hispanic population is positive and small (0.13).

FIGURE 2

The Estimated Correlations between Consumer Medical Debt in Collections and County Characteristics



Sources: Urban Institute analysis of August 2021 credit bureau data combined with county-level characteristics (see table A.1 for additional details).

Notes: Each dot represents one county observation. The red line is a linear fit of the variables at the county level. Correlations are also calculated at the county level.

Regression Results

We use a regression model to disentangle the associations between medical debt in collections and county-level socioeconomic and health characteristics (table 2).[5] Model (1) is our baseline model. In models (2) and (4) we add health condition measures at the county level, which are not typically included as explanatory variables in medical debt modeling. In models (3) and (4) we control for state fixed effects, comparing the medical debt of counties within the same state.

TABLE 2

Estimated Relationship between Consumer Medical Debt in Collections and County Characteristics

	(1)	(2)	(3)	(4)
% Uninsured	0.479***	0.517***	0.119***	0.164***
	(0.032)	(0.029)	(0.033)	(0.032)
Log Medicare spending per capita	11.709***	-0.970	4.951***	-0.273
	(0.915)	(1.109)	(0.888)	(1.035)
Log average income	-12.023***	-7.019***	-9.969***	-7.177***
	(0.670)	(0.664)	(0.595)	(0.626)
Race and ethnicity				
Omitted: % non-Hispanic white				
% Hispanic	0.024**	0.035***	-0.008	-0.023*
	(0.011)	(0.010)	(0.012)	(0.012)
% non-Hispanic Black	0.110***	-0.002	0.082***	0.035***
	(0.012)	(0.014)	(0.012)	(0.013)
% non-Hispanic AIAN	-0.255***	-0.168***	-0.048**	-0.043**
	(0.022)	(0.020)	(0.020)	(0.019)
% non-Hispanic AAPI	-0.206***	-0.272***	-0.170***	-0.198***
	(0.048)	(0.043)	(0.045)	(0.043)
% non-Hispanic two or more races	0.670***	0.588***	0.088	0.052
	(0.078)	(0.071)	(0.084)	(0.082)
Share Biden minus share Trump	-0.009*	-0.012**	-0.024***	-0.024***
	(0.005)	(0.005)	(0.005)	(0.005)
Percent ages 65 and older	-0.205***	-0.157***	-0.235***	-0.214***
	(0.030)	(0.028)	(0.026)	(0.026)
Chronic condition prevalence among Medicare population				
Omitted: % 0–1 chronic condition prevalence				
% 2–3 Chronic condition prevalence		0.271***		0.003
		(0.065)		(0.058)
% 4–5 Chronic condition prevalence		0.451***		0.216***
		(0.067)		(0.059)
% 6+ Chronic condition prevalence		0.399***		0.240***
		(0.056)		(0.051)
% Live births with low birthweight (< 2,500 grams)		1.114***		0.711***
		(0.079)		(0.074)
State fixed effects	No		Yes	
Observations	3,002	3,002	3,002	3,002
R-squared	0.464	0.564	0.702	0.721

Sources: Urban Institute August 2021 credit bureau data combined with county-level characteristics (see table A.1 for additional details).

Notes: All models are estimated using ordinary least squares (OLS). Robust standard errors are reported in parentheses. AAPI = Asian American and Pacific Islander, AIAN = American Indian/Alaska Native. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Key findings from the regression models include the following:

- Across all models, **people living in counties with a higher share of uninsured people are significantly more likely to have medical debt in collections**. This relationship is consistent with prior causal evidence that health insurance expansions are associated with reductions in medical debt (Caswell and Waidmann 2017; Hu et al. 2018). While still positive, the size of the coefficient on percent uninsured declines when we add state fixed effects to the model, as state fixed effects capture significant variation in health insurance coverage, largely because of the Medicaid expansion under the ACA.[6]
- **Counties with a higher average income have lower rates of medical debt in collections**, even when controlling for state fixed effects and health conditions. This result is not surprising because families with higher incomes are more likely to have the means to pay their medical bills on time. The relationship between income and medical debt tends to get weaker when we add health status measures to the model.
- **Adults are more likely to have medical debt in collections in counties with higher average Medicare expenditures**.[7] Medicare spending per capita is positively associated with medical debt in collections when we do not account for health condition measures of the county population. However, the relationship is not significant when we control for chronic condition prevalence and the share of live births with low birthweight in the county. In other words, **differences in the population's health conditions explain a significant share of the association between health care expenditures and medical debt.**
- In most models, we find that **counties with a greater share of Black people have higher rates of medical debt in collections**. This is consistent with past work showing that Black people are more likely to have past-due medical debt (CFPB 2022; Karpman et al. 2022).[8] This relationship is much weaker when we add measures of the population's health status. In other words, part of the Black–white population gap in medical debt at the county-level is explained by differences in the prevalence of chronic health conditions, which is consistent with evidence that Black people have higher rates of heart disease, cancer, and other chronic conditions.[9]
- In terms of the other racial and ethnic groups, **counties with high shares of Asian American and Pacific Islander (AAPI) and American Indian and Alaska Native (AIAN) people have lower rates of medical debt in collections** across all models.[10] This finding could be driven by the fact that the AIAN population is largely served by the Indian Health Service with little or no cost-sharing requirements.
- To proxy for potential differences in consumer protections and the generosity of health and social welfare policies at the local level, we also use the margin of votes for Biden over Trump in the 2018 election as a control in our regressions. **More liberal-leaning counties typically have a smaller share of adults with medical debt in collections**. Counties with a greater margin of votes for Biden over Trump in the 2018 election have a smaller share of adults with debt in collections, even when comparing counties within the same state.

- We also find that **counties with a higher share of the population ages 65 and older have lower rates of medical debt in collections**. This finding is consistent with previous causal evidence that turning 65 and accessing Medicare significantly reduces the probability of large amounts of medical debt in collections (Caswell and Goddeeris 2020).
- We also investigate the relationship between chronic condition prevalence and medical debt. **Counties with a high share of the population with chronic conditions are more likely to have high shares of medical debt**. A 10 percentage-point increase in the share of the Medicare population with six or more chronic conditions is associated with a 2.4 percentage-point increase in medical debt in collections (model 4) in the county.
- As noted in box 1 and figure A.1, we find that **chronic condition prevalence among a county's Medicare-age population serves as good proxy for the chronic condition prevalence of younger adults in the same county**. However, to further validate that our results are not based on only the Medicare population's health conditions, we also estimate models with the disability rate for the population ages 18 to 64 included as a health condition explanatory variable (table A.2). We find that the disability rate among 18- to 64-year-old adults is positively associated with medical debt in collections in all models, with the effect being stronger when not controlling for the prevalence health conditions in the Medicare population.
- We also investigate the relationship between low birthweight rates and medical debt. Birthweight is the primary measure of a baby's health in analyses of infant health and welfare in economics. Low birthweight is typically associated with specific maternal behaviors, such as cigarette smoking during pregnancy, and is closely related to preterm birth, which can present additional complications for mothers and newborns. Infants with low birthweight may need additional high-intensity care in costly settings such as neonatal intensive care units and may need greater health care resources in early childhood. We find that **medical debt in collections is higher in counties with high shares of low birthweight births**. A 10 percentage-point increase in the share of low birthweight is associated with a 7.1 percentage-point increase in the share of medical debt in collections (model 4).

The Relative Importance of Each Measure of Medical Debt

The regression model described in the previous section provides evidence on the associations between a county's socioeconomic and health status and medical debt. We are also interested in each characteristic's power to predict medical debt in collections in a county.

For this purpose, we use random forest, a machine-learning algorithm, to predict medical debt in collections (Breiman 2001). The goal is to rank each county's socioeconomic characteristics in terms of their predictive power for medical debt in collections (box 2 presents a detailed description of this approach). Random forest is one of the best-performing machine learning algorithms and, unlike our regression analysis, has the advantage of not imposing any functional form assumptions on the relationship between the county's characteristics and medical debt.

BOX 2

Random Forest Model Implementation

A tree-based model involves recursively partitioning the dataset into groups (trees) based on a certain criterion until a predetermined stopping condition is met.[a] The random forest model is an ensemble-learning algorithm where the algorithm averages predictions over many individual trees. These individual trees are built on bootstrap samples rather than the original sample. In implementing the algorithm, we use 3,002 observations (counties) and 14 explanatory variables listed in the model (2) of table 2. All of our explanatory variables are continuous variables.

To start the model-training process, we arrange the data points in a randomly sorted order. We then equally split the data into training and test data. Next, we tune the number of algorithm iterations (the number of subtrees). We use both an out-of-bag error (tested against training data subsets that are not included in subtree construction) and a validation error (tested against the testing data) to determine the best possible model. Using this rule, we set 120 iterations and obtain a final out-of-bag error of 4.27. This is somewhat lower than the root mean square error calculated against the test data, 5.23. We also set the number of predictors considered at each split to be equal to the square root of the number of independent variables and set the minimum number of observations per leaf to be equal to one.

Variable importance is calculated by adding up the improvement in the objective function given in the splitting criterion over all internal nodes of a tree and across all trees in the forest, separately for each predictor variable. The variable importance score is then normalized by dividing all scores over the maximum score: the importance of the most important variable is always 100 percent.

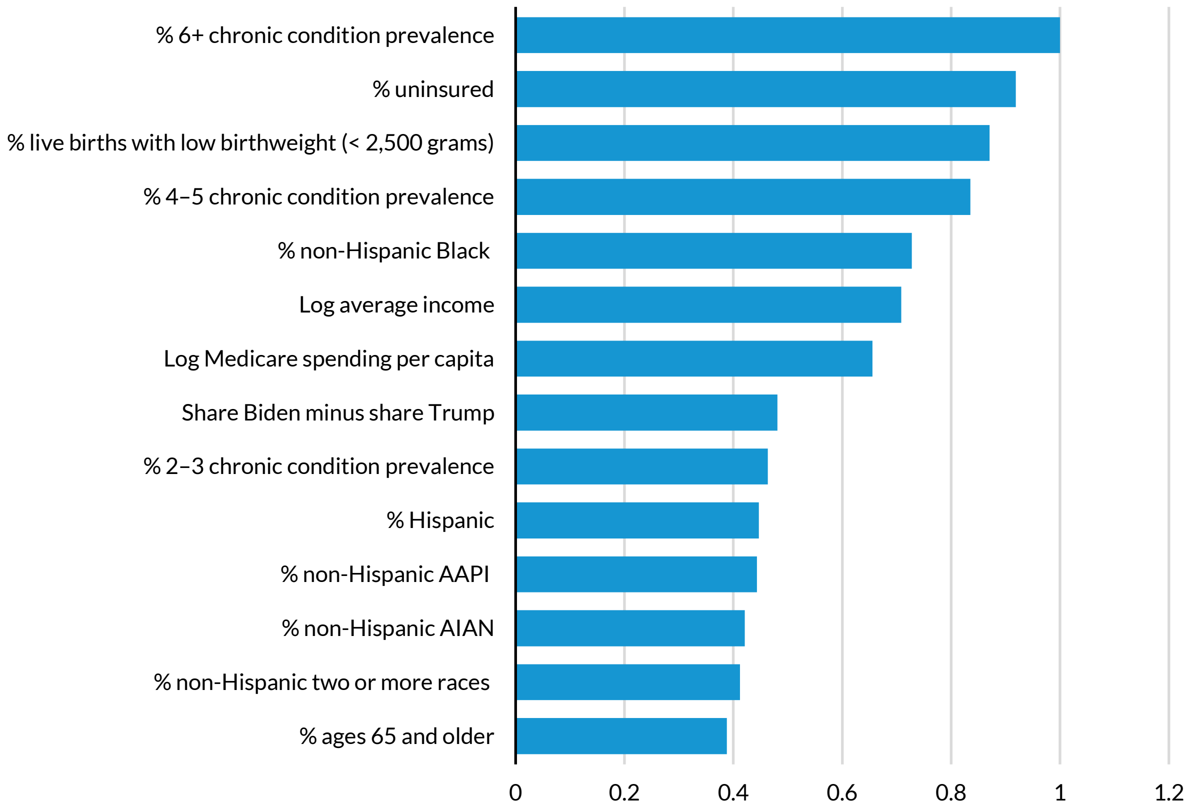
[a] Matthias Schonlau and Rosie Yuyan Zou, "The Random Forest Algorithm for Statistical Learning," *The Stata Journal: Promoting Communications on Statistics and Stata* 20 (2020), no. 1: 3–29, https://doi.org/10.1177/1536867X20909688.

From this analysis, we find the following:

- **The single most important predictor of a county's share of medical debt in collections is the percentage of the Medicare population with six or more chronic conditions (figure 3)**. Among all characteristics included in our model, having a high share of the population with multiple chronic conditions is the strongest predictor for forecasting whether residents are unable to pay their medical bills. The importance of this predictor is normalized to be 100 percent.
- **The second most important predictor is the share of population without health insurance**, (about 92 percent importance relative to the most important predictor), **followed by the low birthweight rate** (about 87 percent importance relative to the most important predictor).
- **Other county socioeconomic characteristics, such as the share of the Black population** (73 percent relative to the most important predictor) **and average income** (71 percent) **are also important predictors** but with a slightly smaller predictive power in the model.
- These results highlight the **importance of accounting for both the population's health insurance coverage and prevalence of chronic health conditions** when investigating the determinants of medical debt.

FIGURE 3

The Relative Importance of Predictors for Percentage with Medical Debt in Collections



Sources: Urban Institute August 2021 credit bureau data combined with county-level characteristics (see table A.1 for additional details).

Notes: We use a machine learning random forest algorithm to predict the share of adults with medical debt in collections. Variable importance is calculated by adding up the improvement in the objective function given in the splitting criterion over all internal nodes of a tree and across all trees in the forest, separately for each predictor variable. In the implementation of random forest, the variable importance score is normalized by dividing all scores over the maximum score; the importance of the most importance variable is always 100 percent. AAPI = American Asian and Pacific Islander, AIAN = American Indian and Alaska Native.

Conclusion

This brief highlights the geographic variation in medical debts across counties and characteristics of counties with the highest shares of people with medical debt in collections. We also estimate the association between population demographic, social, economic, and health characteristics and medical debt. We consistently find that factors related to insurance coverage, household income, chronic illness prevalence, and other measures of population health status are closely tied to the geographic variation in medical debt. We find that the county's prevalence of chronic conditions is the strongest predictor of a county's medical debt in collections.

There are a few possible explanations for the strong relationship between population health and the share of adults who are unable to pay their medical bills. On one hand, there is higher demand for medical care in counties where the population is in poor health. Adults with chronic conditions need to seek care more often and therefore will have more medical bills to pay, even when covered by health insurance. On the other hand, people with medical debt are more likely to forgo needed health care because of its cost.

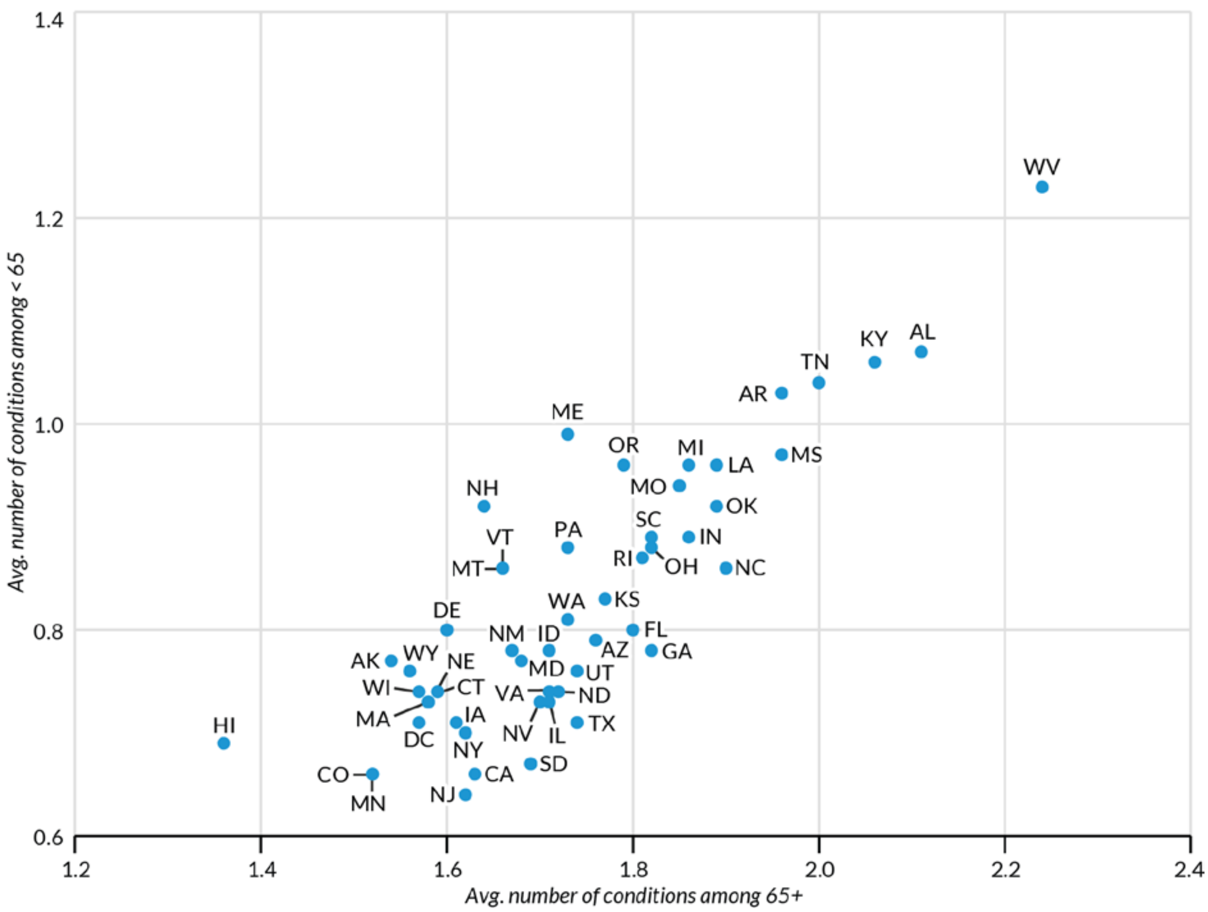
While this brief presents descriptive results, past literature offers some causal evidence on policies that reduce medical debt. Studies of the ACA and pre-ACA coverage expansions provide strong evidence that expanding Medicaid reduces medical debt in collections and other measures of financial distress by extending coverage to the uninsured (Caswell and Waidmann 2017; Finkelstein et al. 2012; Gross and Notowidigdo 2011; Hu et al. 2018). The No Surprises Act, which took effect in January 2022, also offers researchers an opportunity to better understand the effects of protections for consumers with health insurance coverage against out-of-network medical bills.

Finally, the three major credit-reporting agencies recently decided to remove most medical debt in collections from credit reports starting in July 2022. Though newer credit-scoring models place less weight on medical debt in collections, removing most medical collections tradelines from credit reports means they will no longer affect a person's credit score (CFPB 2022). Moving forward, this decision could mitigate the observed relationships between medical debt and county-level characteristics. Credit scores affect people's ability to borrow, influencing their ability to buy a home or car or pay for emergency expenses. Credit can be vital to families who need to smooth their expenses until the next paycheck or pay for an emergency expenditure like a medical visit.

Appendix

FIGURE A.1

Estimated Relationship between the Number of Chronic Conditions among the Elderly and Nonelderly Adult Populations at the State Level



Source: 2018 survey data from the Behavioral Risk Factors Surveillance System.
Note: Each dot represents one state observation. The estimate correlation coefficient is 0.85.

TABLE A.1

Variable Definitions

Metric	Definition	Time	Source
Outcome			
Share with medical debt in collections	Share of people with a credit bureau record who have medical debt in collections	Aug-21	Credit bureau data
Socioeconomic measures			
Share without health insurance coverage	Share of people who do not have health insurance coverage	2015–19 Average	ACS
Average household income	Average household income in 2021 dollars	2015–19 Average	ACS
Share non-Hispanic white	Share of non-Hispanic white population	2020	Census
Share non-Hispanic Black	Share of non-Hispanic Black population	2020	Census
Share non-Hispanic AAPI	Share of non-Hispanic AAPI population	2020	Census
Share non-Hispanic AIAN	Share of non-Hispanic AIAN population	2020	Census
Share non-Hispanic two or more races	Share of non-Hispanic two or more races population	2020	Census
Share Hispanic	Share of Hispanic population	2020	Census
Population	County population totals	2020	Census
Dem margin	Share votes Biden minus share votes Trump	Nov-20	MIT Election Lab
Health measures			
Chronic condition prevalence among Medicare beneficiaries	21 chronic conditions are grouped into four categories (0–1, 2–3, 4–5 and 6 or more). Prevalence estimates are expressed as a percentage of beneficiaries in fee-for-service population	2018	CMS Multiple Chronic Conditions
Prevalence 0–1	Prevalence of individuals with 0–1 MCC's		
Prevalence 2–3	Prevalence of individuals with 2–3 MCC's		
Prevalence 4–5	Prevalence of individuals with 4–5 MCC's		
Prevalence 6+	Prevalence of individuals with 6+ MCC's		
Low birthweight	Percentage of live births with low birthweight (< 2,500 grams).	2013–19 Average	County Health Rankings/NCHS Natality Files
Any disability (ages 18 to 64)	Share of nonelderly adults with a reported disability	2015–19 Average	ACS
Medicare spending per capita	Weighted mean of total Medicare spending per capita; weighted by prevalence; standardized by geography	2018	CMS Multiple Chronic Conditions

Note: AAPI = Asian American and Pacific Islander, AIAN = American Indian and Alaska Native.

TABLE A.2

Estimated Relationship between Consumer Medical Debt in Collections and County Characteristics, Including Disability for Nonelderly Adults

	(1)	(2)	(3)	(4)
% Uninsured	0.497***	0.524***	0.125***	0.166***
	(0.031)	(0.029)	(0.033)	(0.032)
Log Medicare spending per capita	10.918***	-0.008	4.841***	0.116
	(0.893)	(1.114)	(0.882)	(1.037)
Log average income	-5.783***	-4.724***	-7.559***	-5.922***
	(0.812)	(0.763)	(0.696)	(0.699)
Race and ethnicity				
Omitted: % white non-Hispanic				
% Hispanic	0.028***	0.037***	-0.006	-0.021*
	(0.011)	(0.010)	(0.012)	(0.012)
% non-Hispanic Black	0.110***	0.006	0.085***	0.039***
	(0.012)	(0.014)	(0.012)	(0.013)
% non-Hispanic AIAN	-0.237***	-0.168***	-0.049**	-0.045**
	(0.021)	(0.020)	(0.019)	(0.019)
% non-Hispanic AAPI	-0.204***	-0.265***	-0.175***	-0.199***
	(0.047)	(0.043)	(0.045)	(0.043)
% non-Hispanic two or more races	0.432***	0.491***	0.002	0.004
	(0.078)	(0.072)	(0.084)	(0.082)
Share Biden minus share Trump	-0.005	-0.010**	-0.022***	-0.022***
	(0.005)	(0.005)	(0.005)	(0.005)
% Ages 65 and older	-0.239***	-0.180***	-0.259***	-0.231***
	(0.029)	(0.028)	(0.026)	(0.026)
% Any disability (ages 18 to 64)	0.439***	0.202***	0.194***	0.119***
	(0.034)	(0.034)	(0.030)	(0.030)
Chronic condition prevalence among Medicare Population				
Omitted: % 0–1 chronic condition prevalence				
% 2–3 Chronic condition prevalence		0.266***		0.004
		(0.065)		(0.058)
% 4–5 Chronic condition prevalence		0.463***		0.224***
		(0.066)		(0.059)
% 6+ Chronic condition prevalence		0.330***		0.208***
		(0.057)		(0.051)
% Live births with low birthweight (< 2,500 grams)		1.013***		0.674***
		(0.080)		(0.075)
State fixed effects	No		Yes	
Observations	3,002	3,002	3,002	3,002
R-squared	0.493	0.569	0.707	0.723

Sources: Urban Institute August 2021 credit bureau data combined with county-level characteristics (see table A.1 for additional details).

Notes: All models are estimated using ordinary least squares (OLS). Robust standard errors are reported in parentheses. AAPI = Asian American and Pacific Islander, AIAN = American Indian and Alaska Native. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes

[1] The Consumer Financial Protection Bureau (CFPB) also specified that "companies that try to collect on medical bills that are prohibited by the No Surprises Act, or who furnish information to credit bureaus about such invalid debts, may face significant legal liability under the Fair Debt Collection Practices Act and the Fair Credit Reporting Act (CFPB 2022). See also Maanasa Kona, "State Balance-Billing Protections," Commonwealth Fund, February 5, 2021, https://www.commonwealthfund.org/publications/maps-and-interactives/2021/feb/state-balance-billing-protections.

[2] Following the terminology in the data analyzed, this brief uses the term "Hispanic" to describe people of Latin American descent. The authors acknowledge this may not be the preferred identifier, and we remain committed to employing inclusive language whenever possible.

[3] Breno Braga, Alexander Carther, Kassandra Martinchek, Signe-Mary McKernan, and Caleb Quakenbush, "Debt in America: An Interactive Map," Urban Institute, updated March 31, 2021, https://apps.urban.org/features/debt-interactive-map/?type=medical&variable=perc_debt_med.

[4] Lawrence County, Ohio, is the only non-Southern state county among the top 100 counties with the most medical debt in collections in the country.

[5] All models are estimated using unweighted ordinary least squares. Robust standard errors are reported in parentheses.

[6] "Health Insurance Coverage of the Total Population," Kaiser Family Foundation, 2019, https://www.kff.org/other/state-indicator/total-population/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

[7] Skinner and Fisher (2010) conclude that most of the variation in Medicare spending is attributable geographic differences in medical practice styles. However, Sheiner (2014) contends that geographic variation in Medicare spending may merely reflect underlying socioeconomic and demographic characteristics that drive the demand for medical care.

[8] Signe-Mary McKernan, Steven Brown, and Genevieve M. Kenney, "Past-Due Medical Debt a Problem, Especially for Black Americans," *Urban Wire* (blog), March 27, 2017, https://www.urban.org/urban-wire/past-due-medical-debt-problem-especially-black-americans.

[9] Latoya Hill, Samantha Artiga, and Sweta Haldar, "Key Facts on Health and Health Care by Race and Ethnicity," Kaiser Family Foundation, January 26, 2022, https://www.kff.org/racial-equity-and-health-policy/report/key-facts-on-health-and-health-care-by-race-and-ethnicity/.

[10] Different from our findings, Martinchek and colleagues show that majority–AIAN communities have higher shares of medical debt in collections than majority-white communities: Kassandra Martinchek, Alex Carther, Breno Braga, Caleb Quakenbush, and Signe-Mary McKernan, "Credit Health during the COVID-19 Pandemic," Urban Institute, updated March 8, 2022, https://apps.urban.org/features/credit-health-during-pandemic/#:~:text=Credit%20scores%20improved%2C%20and%20the,of%20693%20in%20February%202020. One potential explanation is that the statistics do not account for differences in income across communities.

References

Adams, Alyce S., Raymond Kluender, Neale Mahoney, Jinglin Wang, Francis Wong, and Wesley Yin. 2021. "The Impact of Financial Assistance Programs on Health Care Utilization." Working paper 29227. Cambridge, MA: National Bureau of Economic Research.

Braga, Breno, Signe-Mary McKernan, and Andrew Karas. 2017. "Is Financial Knowledge Associated with Past-Due Medical Debt?" Washington, DC: Urban Institute.

Breiman, Leo. 2001. "Random Forests." Berkeley: University of California Statistics Department.

Caswell, Kyle J., and John H. Goddeeris. 2020. "Does Medicare Reduce Medical Debt?" *American Journal of Health Economics* 6 (1): 72–103.

Caswell, Kyle J., and Timothy A. Waidmann. 2017. "The Affordable Care Act Medicaid Expansions and Personal Finance." *Medica Care Research and Review* 76 (5): 538–71.

CFPB (Consumer Financial Protection Bureau). 2022. *Medical Debt Burden in the United States*. Washington, DC: CFPB.

Cooper, Zack, James Han, and Neale Mahoney. 2021. "Hospital Lawsuits Over Unpaid Bills Increased by 37 Percent in Wisconsin from 2001 to 2018." *Health Affairs* 40 (12): 1830–35.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J. Notowidigdo. 2018. "Myth and Measurement: The Case of Medical Bankruptcies." *New England Journal of Medicine* 378 (12): 1076–78.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127 (3): 1057–106.

Gross, Tal, and Matthew J. Notowidigdo. 2011. "Health Insurance and the Consumer Bankruptcy Decision: Evidence from Expansions of Medicaid." *Journal of Public Economics* 95 (7–8): 767–78.

Hamel, Liz, Mira Norton, Karen Pollitz, Larry Levitt, Gary Claxton, and Mollyann Brodie. 2016. *The Burden of Medical Debt: Results from the Kaiser Family Foundation/New York Times Medical Bills Survey*. Oakland, CA: Kaiser Family Foundation.

Hu, Luojia, Robert Kaestner, Bhashkar Mazumder, Sarah Miller, and Ashley Wong. 2018. "The Effect of the Affordable Care Act Medicaid Expansions on Financial Wellbeing." *Journal of Public Economics* 163:99–112.

Karpman, Michael, and Kyle J. Caswell. 2017. "Past-Due Medical Debt among Nonelderly Adults, 2012–15." Washington, DC: Urban Institute.

Karpman, Michael, Kassandra Martinchek, and Breno Braga. 2022. "Medical Debt Fell during the Pandemic. How Can the Decline Be Sustained?" Washington, DC: Urban Institute.

Kluender, Raymond, Neale Mahoney, Francis Wong, and Wesley Yin. 2021. "Medical Debt in the US, 2009–2020." *JAMA* 326 (3): 250–56.

Mendes de Leon, Carlos F., and Jennifer J. Griggs. 2021. "Medical Debt as a Social Determinant of Health." *JAMA* 326 (3): 228–29.

Rabin, David L., Anuradha Jetty, Stephen Petterson, and Allison Froehlich. 2020. "Under the ACA Higher Deductibles and Medical Debt Cause Those Most Vulnerable to Defer Needed Care." *Journal of Health Care for the Poor and Underserved* 31 (1): 424–40.

Sheiner, Louise. 2014. "Why the Geographic Variation in Health Care Spending Cannot Tell Us Much about the Efficiency or Quality of Our Health Care System." *Brookings Papers in Economic Activity* (fall 2014):1–72.

Skinner, Jonathan, and Elliott S. Fisher. 2010. "Reflections on Geographic Variations in US Health Care." Lebanon, NH: Dartmouth Institute for Health Policy and Clinical Practice.

About the Authors

Fredric Blavin is a principal research associate in the Health Policy Center at the Urban Institute. His areas of expertise include a wide range of topics, including hospital finances, telehealth and health information technology, health care reform, private health insurance markets, provider supply, health care spending, child and maternity health, and Medicaid policy. Blavin received his PhD in managerial science and applied economics from the University of Pennsylvania.

Breno Braga is a labor economist and principal research associate in the Center on Labor, Human Services, and Population at the Urban Institute. His research has covered topics such as the role of local conditions in asset accumulation and the local factors associated with debt in collections. His articles have been published in academic journals including the *Journal of Labor Economics*. Braga received his MA in economics from the Pontifical Catholic University of Rio de Janeiro and his PhD in economics from the University of Michigan.

Anuj Gangopadhyaya is a senior research associate in the Health Policy Center. His research focuses on the impact of safety net programs on health and well-being, family income, and education achievement outcomes for children in families with low incomes. He has focused on the impact of Medicaid eligibility expansion on children's education achievement, maternal and child health effects of the earned income tax credit program, and the impact of the Affordable Care Act Medicaid expansion on adult labor supply and fertility rates of women of reproductive age. Gangopadhyaya received his PhD in economics from the University of Illinois at Chicago.

Acknowledgments

This brief was funded by Arnold Ventures. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission. We also would like to thank Michael Karpman, Signe-Mary Mckernan, Aaron Williams, and Stephen Zuckerman for their helpful comments.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.