

Observational Studies of the Effect of Medicaid on Health: Controls Are Not Enough

Seth Freedman, *Indiana University*

Andrew Goodman-Bacon, *Opportunity and Inclusive Growth Institute, Federal Reserve Bank of Minneapolis*

Noah Hammarlund, *University of Washington*

Covariate-adjusted cross-sectional comparisons show that Medicaid patients have worse health outcomes than other patients. We evaluate the validity of this research design for estimating the causal effect of Medicaid on mortality. Even after controlling for common covariates, Medicaid patients have worse preoperative health and lower socioeconomic status than privately insured patients. Controlling for additional variables shrinks the mortality differences but still does not eliminate imbalance in other predetermined variables. These results can be explained by fairly weak assumptions about unmeasured confounders. We conclude that cross-sectional observational methods do not produce valid causal estimates of Medicaid's mortality effects.

We thank Martha Bailey, Melinda Buntin, Kitt Carpenter, Aaron Carroll, Helen Levy, Sarah Miller, Sayeh Nikpay, Harold Pollack, Dan Sacks, Kosali Simon, and Coady Wing for helpful comments and suggestions. The Robert Wood Johnson Foundation provided funding for the Healthcare Cost and Utilization Project (HCUP) data used in this paper. We also thank the Agency for Healthcare Research and Quality and its HCUP data partners (<https://www.hcup-us.ahrq.gov/db/hcupdatapartners.jsp>). All errors are our own. Contact the corresponding author, Andrew Goodman-Bacon,

[*Journal of Labor Economics*, 2021, vol. 39, no. S2]

© 2021 by The University of Chicago. All rights reserved. 0734-306X/2021/39S2-0017\$10.00

Submitted February 18, 2019; Accepted December 4, 2020

I. Introduction

Policy makers frequently rely on empirical research to understand whether Medicaid causally affects health. In turn, researchers have produced an enormous body of work estimating these effects (Levy and Meltzer 2004; Howell and Kenney 2012). Unfortunately, different studies produce dramatically different conclusions. Cross-sectional observational studies compare recipients to nonrecipients and find that Medicaid patients have worse health outcomes (e.g., LaPar et al. 2010, 2011). Instrumental variable studies compare groups whose Medicaid coverage changes for reasons outside their control and find that many, but not all, populations benefit (Currie and Gruber 1996a, 1996b; Aizer 2007; Sommers, Baicker, and Epstein 2012; Bronchetti 2014; Sommers 2014; Wherry and Meyer 2015; Goodman-Bacon 2018; Miller and Wherry 2019; Miller et al. 2019; Borgschulte and Vogler 2020). The only randomized controlled trial of Medicaid eligibility, the Oregon Health Insurance Experiment, finds significant improvements for some health outcomes and imprecise effects for others (Finkelstein et al. 2012; Baicker et al. 2013).

These results matter because Medicaid costs half a trillion dollars per year and is among the most expensive items in federal and state budgets. Advocates for reducing Medicaid funding routinely cite cross-sectional studies to argue that the program is ineffective or harmful. Former Secretary of Health and Human Services Tom Price, for example, recently claimed, “Medicaid is a program that by and large has decreased the ability for folks to gain access to care” (Office of the Press Secretary 2017). More than a third of governors who oppose the Affordable Care Act’s Medicaid expansions feel that it is a “broken program” that “harms its beneficiaries” (Sommers and Epstein 2013). If Medicaid really causes patients harm, then the case for cutting it is clear: we would save money and improve health. If not, then cuts to Medicaid involve difficult trade-offs between public spending and patient outcomes.

Empirical strategies that compare outcomes between Medicaid recipients and nonrecipients face serious challenges in separating the program’s causal effects from the fact that, by law, Medicaid serves lower-income and less healthy patients. Cross-sectional observational studies attempt to address this selection bias by controlling for measurable risk factors (Ayanian et al. 1993; Franks, Clancy, and Gold 1993; Haas and Goldman 1994; Higgins et al. 1998; Roetzheim et al. 1999; Canto et al. 2000; Kelz et al. 2004; Zacharias et al. 2005; Rosen et al. 2009; Salim et al. 2010; LaPar et al. 2012). To interpret the adjusted differences as causal effects, the control variables must completely account for factors that make patients less healthy *and* lead them to qualify for and take up Medicaid coverage.

Researchers have cast doubt on the validity of these designs for several reasons (Frakt et al. 2011). Many data sets do not include controls such as income, assets, and health history, so cross-sectional analyses must omit these variables. Studies often argue that the inclusion of prognostic risk factors (e.g., comorbidities measured at the time of a hospital admission) effectively address selection bias. If Medicaid causally affects preoperative health, though, then these are not appropriate conditioning variables (Rosenbaum 1984), so researchers with access to rich data on health still may not be able to adjust adequately for underlying differences in health between Medicaid recipients and nonrecipients. Finally, economic theory suggests that patients choose certain types of insurance because they privately know they need care and would find themselves in poor health with or without health insurance.

Since cross-sectional observational research leads to such extreme conclusions about Medicaid's causal effects, the internal validity of this design has profound policy implications. This paper assesses the validity of this research design in the context of the effect of Medicaid coverage on postsurgical mortality.

We first replicate a heavily cited cross-sectional finding (LaPar et al. 2010) that even after conditioning on a range of variables, Medicaid patients in the Nationwide Inpatient Sample (NIS) have higher postsurgical mortality than privately insured patients. We find the same result in a sample of National Health Interview Survey (NHIS) respondents who report having had a recent surgery. We show that the covariate-adjusted design fails to eliminate differences in preoperative health, demographics, and socioeconomic status not included in the original study.

Second, we demonstrate that adding more covariates shrinks the differences in mortality between Medicaid and privately insured patients but does not eliminate imbalance in other relevant covariates. Medicaid patients still have worse predetermined health than privately insured patients, even conditional on a larger set of covariates.

Third, we show that adjusted differences in mortality could result from even small amounts of unobservable selection (Altonji, Elder, and Taber 2005). For example, unobservables would not need to differ nearly as much as factors like preoperative health or socioeconomic status do to explain the entire cross-sectional mortality result. Using data from the Oregon Health Insurance Experiment, where we can calculate the amount of bias that equates cross-sectional and experimental estimates, we find very large differences in unobservables, further suggesting that such bias is a plausible driver of cross-sectional findings. Last, we focus on a specific omitted factor, detailed family income that is not recorded in discharge or health survey data, and show that imperfect measurement alone could generate enough bias to fully explain the cross-sectional result.

None of our analyses supports a causal interpretation of covariate-adjusted differences in health by insurance status, nor do they suggest that additional

controls will do so. Our analysis does not imply that insurance status has no causal effect on health but only that cross-sectional methods do not uncover it. Policy makers should not use cross-sectional evidence to evaluate Medicaid’s performance or to guide policy.

II. Data Sources, Estimation Samples, and Outcomes

To match LaPar et al. (2010), our first sample uses individual discharge records from the 2003–7 NIS, a 20% stratified random sample of hospitals in the United States. We include 858,867 patients who underwent at least one of eight major surgical operations: lung resection, esophagectomy, colectomy, pancreatectomy, gastrectomy, abdominal aortic aneurysm repair, total hip replacement, and isolated coronary artery bypass graft. The advantage of the NIS sample is its size and clinically measured outcomes, but as we discuss below it contains relatively few covariates. Our second sample uses data from the 1997–2009 NHIS, a nationally representative stratified random sample of households (Minnesota Population Center and State Health Access Data Assistance Center 2012). We include 45,309 adult respondents who report having surgery in the previous year. The NHIS sample, while significantly smaller, allows us to consider a wider range of covariates not measured in discharge records.

Our main health outcomes is an indicator that equals 1 for patients who died and 0 otherwise. In the NIS, this refers to in-hospital mortality. In the NHIS, this refers to mortality in the year of or the year immediately following the respondent’s survey year. Dates of death for NHIS respondents come from the December 31, 2011, National Death Index. Table 1 presents mortality rates in our two samples. In the NHIS, about 2% of Medicaid and 1% of privately insured patients die, compared with about 4% and 1% in the NIS. These differences most likely arise from the fact that in the NHIS we

Table 1
Mortality Rates and Sample Sizes among Surgical Patients in the Nationwide Inpatient Sample (NIS) and the National Health Interview Survey (NHIS)

Health Insurance Type	NIS: In-Hospital Mortality	NHIS: 1-Year Mortality
Private	.013 (324,633)	.008 (24,793)
Medicaid	.038 (38,458)	.023 (3,484)
Medicare	.044 (472,948)	.065 (13,348)
Uninsured	.033 (22,828)	.006 (3,684)

cannot select patients who had particular operations and patients must survive long enough after surgery to be interviewed.

III. Methods

Our analysis has three parts. First, we replicate the results in LaPar et al. (2010) and show that they are biased by observables. Second, we show that adding controls reduces the estimated effects and, in the NHIS, still does not eliminate imbalance in predetermined variables. Finally, we show that even the more fully controlled results are sensitive to mild assumptions about the nature of the unobservables (Altonji, Elder, and Taber 2005).

A. Replication and Balance Tests

We follow LaPar et al. (2010) and use a logit model to estimate adjusted odds ratios for Medicaid patients relative to privately insured patients in the NIS.¹ We focus on the estimates for Medicaid relative to private insurance (the omitted category) and report effects for other insurance types in the appendix. The NIS replication controls, X_0^{NIS} , include (linear) age, gender, elective operative status, quartiles of mean zip code income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality (AHRQ) comorbid disease. The NHIS replication controls, X_0^{NHIS} , include age, gender, family income categories, and 30 condition variables. All models use the provided survey weights.

To test whether the replication specification generates comparable treatment and control groups, we replace the mortality outcome with other variables not in the set of replication controls. The estimated odds ratios are balance tests that represent the relationship between observable factors and Medicaid status, conditional on a basic set of controls. In the NIS, potential confounders include indicators for patients who report Hispanic ethnicity or a race category other than white, weekend admission (a proxy for an unavoidable hospital visit; Card, Dobkin, and Maestas 2009), and “high” disease severity or mortality risk.² In the NHIS, potential confounders include indicators for patients who are nonwhite, have a high school degree or less, or have ever smoked cigarettes.

¹ In the appendix, we also present average marginal effects, which are easier to interpret and compare across data sets and specifications. In the main text, we present odds ratios to maintain comparability to LaPar et al. (2010) and to make sure that balance tests for covariates with different means are on the same scale.

² The latter two measures come from the All Patient Refined Diagnosis Related Groups, which defines four categories of severity of illness (the extent of physiologic decompensation or organ function loss) or mortality risk. We create indicator variables for patients in the highest two categories of each (“major” or “extreme”).

B. Adding Controls

The replication specifications may yield biased estimates (and fail to balance confounders) if they do not control for enough variables. To test this, we reestimate the adjusted mortality ratios using an expanded set of controls. In the NIS, we add the potential confounders listed above, plus indicators for race and ethnicity categories, single ages, and calendar years (X_1^{NIS}). In the NHIS, we also include indicators for citizenship status, college completion, employment status, poverty status, marital status, and smoking status (X_1^{NHIS}). We also estimate specifications that control for covariates (X_L^{NIS} and X_L^{NHIS}) chosen using a post-double-selection LASSO (least absolute shrinkage and selection operator) model (Belloni, Chernozhukov, and Hansen 2013; Hammarlund 2019).³

We also use the NHIS to test whether controlling for X_0^{NHIS} and X_1^{NHIS} solves the problem of imbalance. We conduct additional balance tests using the expanded NHIS specification and predetermined early-life health variables as outcomes (X_2^{NHIS}). These include indicators for height below the 25th percentile for age and sex, diabetes diagnosis before age 18, smoking initiation before age 16, current or past homelessness, and reported current mental health difficulty.⁴ If more controls really generate comparable treatment and control groups, then we expect to find no differences in these factors by adult Medicaid status after conditioning on a larger control set.

C. Selection on Unobservables: Sensitivity Analyses

More controls may mitigate omitted variables bias, but we cannot tell if they eliminate it. Bias from omitted factors may remain if covariates are measured with error, if covariates are caused by the treatment variable, or if relevant variables have not been or cannot be included. To gauge the extent of

³ We run two LASSO models, one that selects variables predictive of Medicaid insurance and one that selects variables predictive of mortality. We then enter variables selected by either model into our logit specification as control variables. This method accounts for the fact that the exclusion of variables even modestly related to mortality that are strong predictors of Medicaid can lead to substantial omitted variable bias. Potential control variables in the NIS included variables in the main specification plus indicators for each unique ICD-9 code excluding adverse events related to medical care (Hougland et al. 2008), since they may represent intermediate outcomes. Candidate NHIS variables include a rich set of variables describing health history, mental health indicators, geography, and household measures of demographics and socioeconomic status. We exclude self-reported physical health and disability days in the past year, since they too may capture intermediate outcomes between insurance at time of surgery and mortality.

⁴ We consider a respondent to have a current mental health difficulty if they reported feeling any of the following ways “all the time”: hopeless, nervous, restless, sad, worthless, or like “everything is an effort.”

unobservable selection bias, we apply two methods developed by Altonji, Elder, and Taber (2005) based on the following bivariate probit model:

$$D = 1\{X'\beta + u > 0\}, \quad (1)$$

$$Y = 1\{X'\gamma + \alpha D + \epsilon > 0\}, \quad (2)$$

$$\begin{bmatrix} u \\ \epsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right). \quad (3)$$

Equation (1) is a model for the assignment of the Medicaid treatment variable, D . Conditional on X , variation in Medicaid status comes from u , the unobservable determinants of insurance type.⁵ Respondents could have high values of u if they spend down their assets to qualify for Medicaid, if they have large medical bills that grant them medically needy status, if they lack employer-sponsored insurance, or if they anticipate costly health problems in the future. Equation (2) is a model for the mortality outcome, Y . In this framework, α captures the effect of Medicaid on mortality, but the estimated coefficient on D could be biased by differences in ϵ , unobservable determinants of mortality, between Medicaid and privately insured patients.

The causal effect of Medicaid on (the latent index for) mortality, α , is identified by differences in the treatment indicator, D , that remain after conditioning on X . Internal validity hinges on how factors that lead patients to take up Medicaid (u) relate to other determinants of mortality (ϵ). Equation (3) formalizes this by specifying a joint distribution for u and ϵ with correlation ρ . The identifying assumption in cross-sectional research designs like La Par et al. (2010) is that $\rho = 0$, but one can imagine many circumstances that affect both “tastes” for Medicaid (or low “resistance to treatment”; Heckman and

⁵ Identification in a cross-sectional observational design comes from comparing observationally equivalent respondents with high values of u (Medicaid recipients) to those with low values of u (nonrecipients). Conditional on X' , Medicaid patients are those whose unobserved taste for Medicaid (u) is high enough for them to cross the threshold value, $-X'\beta$. Therefore, Medicaid recipients have higher average u than nonrecipients by definition: $E[u|X, D = 1] - E[u|X, D = 0] = E[u|u > -X'\beta] - E[u|u < -X'\beta] > 0$. When u follows a standard normal distribution, these expectations have the familiar inverse Mills ratio:

$$E[u|u > -X'\beta] = \frac{\phi(-X'\beta)}{1 - \Phi(-X'\beta)} = \lambda^+(-X'\beta),$$

$$E[u|u < -X'\beta] = \frac{-\phi(-X'\beta)}{\Phi(-X'\beta)} = \lambda^-(-X'\beta).$$

Because $\lim_{-X'\beta \rightarrow \infty} \lambda^+(-X'\beta) = \infty$ while $\lim_{-X'\beta \rightarrow \infty} \lambda^-(-X'\beta) = 0$, their difference grows as $-X'\beta$ gets larger. Given the latent index structure, this corresponds to settings in which treatment is rarer conditional on X' .

Vytlačil 2005) and mortality risk. For example, job loss increases both Medicaid take-up (Jolly and Phelan 2017) and mortality (Sullivan and von Wachter 2009), so patients who need and use Medicaid because of an unmeasured recent job loss (u) may also have high unmeasured risk of mortality for reasons unrelated to Medicaid (ϵ). We use two strategies that combine information we can obtain from equations (1)–(3) with assumptions about the nature of ϵ to ask how likely it is that estimates of α come entirely from selection bias rather than from a true treatment effect.

1. Varying the Correlation in Unobservables

We could estimate the parameters of equation (1), form the threshold value $X'\beta$ for each respondent, and calculate the expected value of u given her Medicaid status. Medicaid recipients with relatively high earnings, for example, would likely have a negative value of $X'\beta$, so they must have some other circumstance contained in u that makes them eligible. Then if we knew ρ , we could calculate each respondent's expected value of the mortality unobservables, ϵ , and control for it in equation (2).⁶ Unfortunately, absent an instrument for Medicaid participation that is excludable from equation (2), we cannot credibly estimate ρ .

Instead, our first approach combines what we can learn about u from equation (1) with a range of assumptions about ρ that we impose on equation (3) in order to evaluate how serious unobservable selection might be in equation (2). We jointly estimate the bivariate probit model in equations (1)–(3), imposing a range of values for ρ .⁷ This removes the amount of selection bias that would be present in $\hat{\alpha}$ if our assumption about ρ was correct. We plot how the estimated mortality odds ratios change when unobservable selection is more or less serious. If Medicaid recipients have extremely different values of u compared with observationally equivalent nonrecipients, even a weak relationship between u and ϵ can lead to large bias. In this case, the relative odds would shrink to 1 at low values of ρ , and we would conclude that even a “small amount” of selection on unobservables can explain the adjusted mortality differences.

2. Comparing Unobservables to Observables under the Null of No Treatment Effects

The previous strategy is easy to interpret at extreme values of ρ , but in the middle it can be hard to gauge what amount of selection is plausible. Altonji,

⁶ With knowledge of ρ , the magnitude of unobservables in the treatment allocation equation, which we can calculate, is informative about unobservable selection in the outcome equation: $E[\epsilon|u > -X'\beta] = \rho\lambda^+(-X'\beta)$.

⁷ We use a bivariate probit instead of a logit, as in our earlier analysis, to follow Altonji, Elder, and Taber (2005) and for computational simplicity. The normal is the only distribution with a single parameter for the correlation between two random variables, which we need to vary in fig. 3. We use Stata 14's `cmp` command to implement the restricted bivariate probit models.

Elder, and Taber (2005) propose comparing $\hat{\alpha}$ to the observed amount of selection from the covariates. Their key insight is that differences in the estimated index function, $X'\gamma$, between Medicaid recipients and nonrecipients measure the amount of selection on observables. We can use this as a benchmark to ask: if $\hat{\alpha}$ only contains selection on unobservables, then how big is it relative to the measured amount of selection on the observables? If $\hat{\alpha}$ is very large, then the unobserved components of mortality would have to differ much more than observed components in order to rationalize the hypothesis of no treatment effect, casting doubt on the assumption of no treatment effect. Alternatively, if $\hat{\alpha}$ is small, then even a small amount of unobservable selection could explain it, and the no treatment effect hypothesis would seem plausible.

The standardized amount of selection on observables is the difference between Medicaid and privately insured patients in the covariate index divided by its variance: $(E[X'\gamma|D = 1] - E[X'\gamma|D = 0])/V(X'\gamma)$. The standardized amount of selection on unobservables has the same form: $(E[\epsilon|D = 1] - E[\epsilon|D = 0])/V(\epsilon)$. The key to evaluating bias is the assumption that observable and unobservable selection are equal—if recipients and nonrecipients differ by 1 standard deviation in their observables, they also differ by 1 standard deviation in their unobservables.⁸ The equal selection assumption defines what we think of as a plausible amount of selection.

The bias in $\hat{\alpha}$ that we would observe under the null that $\alpha = 0$ and the equal selection assumption approximately equals

$$\begin{aligned} \text{bias}(\hat{\alpha}) &= \frac{V(D)}{V(\tilde{D})} \frac{E[\epsilon|D = 1] - E[\epsilon|D = 0]}{V(\epsilon)} \\ &= \frac{V(D)}{V(\tilde{D})} \frac{E[X'\gamma|D = 1] - E[X'\gamma|D = 0]}{V(X'\gamma)}. \end{aligned} \quad (4)$$

The first line is the omitted variables bias formula for $\hat{\alpha}$ in a linear version of equation (2). Omitted variables bias formulas for nonlinear models are qualitatively similar (see Rosenbaum and Rubin 1983; Lin, Psaty, and Kronmal 1998). The term $V(\tilde{D})$ is the variance of the residual from a regression of D on X , and the ratio $V(D)/V(\tilde{D})$ comes from the fact that the difference in ϵ is with respect to D , but the regression coefficient uses variation in \tilde{D} .⁹ The

⁸ Equal selection is at least as plausible an assumption as the one necessary to give $\hat{\alpha}$ a causal interpretation: $E[\epsilon|D = 1] - E[\epsilon|D = 0] = 0$. Unmeasured or unmeasurable variables as well as common data limitations may make it reasonable to expect similar variation in the observable and unobservable determinants of mortality. Hospital discharge data, e.g., include categories of zip code income, which means that all variation in personal or household income within zip code categories is in the error term.

⁹ This approach is based on a linear model, in which case the bias in $\hat{\alpha}$ is

$$\frac{\text{cov}(\tilde{D}, \epsilon)}{V(\tilde{D})} = \frac{\text{cov}(D - X'\beta, \epsilon)}{V(\tilde{D})} = \frac{\text{cov}(D, \epsilon)}{V(\tilde{D})} - \frac{\text{cov}(X'\beta, \epsilon)}{V(\tilde{D})},$$

second line follows directly from the equal selection assumption and shows how to calculate expected bias. We calculate γ under the null of no Medicaid effect from a mortality probit model that excludes the Medicaid dummy. We then calculate $E[X'\gamma|D = 1] - E[X'\gamma|D = 0]$ as the mean difference in the probit index function between individuals with Medicaid and those with private insurance and $V(X'\gamma)$ as the variance of the index function.

Equation (4) gives the plausible amount of unobserved selection under the null of no treatment effect. Plausibility comes from the equal selection assumption. The ratio of $\hat{\alpha}$ to the second line of equation (4) shows how much unobservables would have to vary relative to observables to explain our estimate of $\hat{\alpha}$. If the ratio is large, then the null that $\alpha = 0$ can be true only if there is much more selection on unobservables than observables. If the ratio is small, then the estimate of $\hat{\alpha}$ is consistent with the null under weaker assumptions about unobservables. It would take only a small amount of unobservable selection to rationalize the finding. We present estimates of $\hat{\alpha}$, estimates of observable selection from equation (4), and their ratios.

IV. Results

Figure 1 presents our replication results controlling for X_0^{NIS} or X_0^{NHIS} . In the NIS, Medicaid patients are about twice as likely to die in the hospital as private-pay patients (odds ratio, 1.98; 95% confidence interval [CI], 1.86–2.11), which is nearly identical to LaPar's result (odds ratio, 1.96; 95% CI, 1.84–2.10; table 6). Despite differences in the sample and covariates, we reproduce the NIS mortality result in the NHIS almost exactly (odds ratio, 2.06; 95% CI, 1.44–2.95). Table 1 shows that mortality rates are relatively low, even after surgery, so odds ratios are close to risk ratios.

Despite the range of control variables in this specification, figure 2 shows that Medicaid patients are not comparable to privately insured patients. Medicaid patients in the NIS have significantly elevated preoperative risk. They are more likely to be admitted on the weekend (odds ratio, 1.21; 95% CI, 1.16–1.25), to have high predicted mortality risk (odds ratio, 1.55; 95% CI,

where the second term equals zero because ϵ is orthogonal to the column space of X . Multiplying and dividing by $V(D)$ allows us to write the bias as a function of the ordinary least squares (OLS) coefficient from a regression of ϵ on D , which is just the difference in the expectation of ϵ between Medicaid and privately insured patients:

$$\frac{V(D)}{V(\tilde{D})} \frac{\overbrace{\text{cov}(D, \epsilon)}^{\text{OLS coefficient of } \epsilon \text{ on } D}}{V(D)} = \frac{V(D)}{V(\tilde{D})} (E[\epsilon|D = 1] - E[\epsilon|D = 0]).$$

This yields eq. (4).

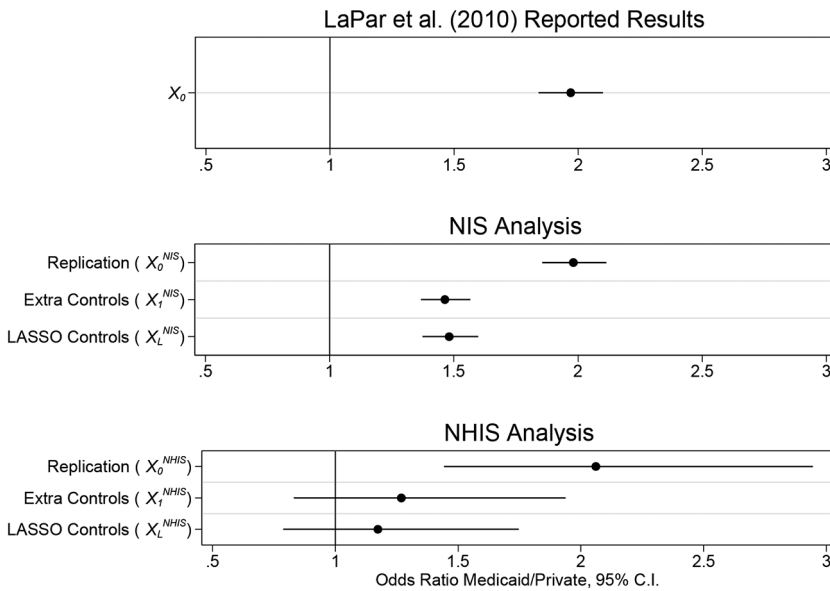


FIG. 1.—Adjusted odds ratios for mortality: Medicaid versus private insurance. This figure reports adjusted odds ratios for in-hospital mortality (Nationwide Inpatient Sample [NIS]) and mortality in survey year of or following surgery (National Health Interview Survey [NHIS]). The reference group is private insurance as primary payer status, and odds ratios for Medicare and uninsured patients are reported in tables A1 and A2. NIS models use the NIS discharge weight. Covariates from LaPar et al. (2010) include patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. Extra controls include age dummies, nonwhite, major loss of function, and major mortality risk. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. NHIS models include adults (age 18 or older) from the 1997–2009 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Covariates mimicking LaPar et al. (2010) are age, sex, condition indicators, and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor’s degree holder, and dummies for smoking frequency. CI = confidence interval.

1.51–1.60), and to have high predicted severity (odds ratio, 1.44; 95% CI, 1.40–1.48). They are also more likely to be nonwhite (odds ratio, 2.01; 95% CI, 1.96–2.05), a documented risk factor for adverse in-hospital outcomes (Fiscella et al. 2000). Medicaid patients in the NHIS also differ significantly in demographics, socioeconomic status, and health behaviors. They are more

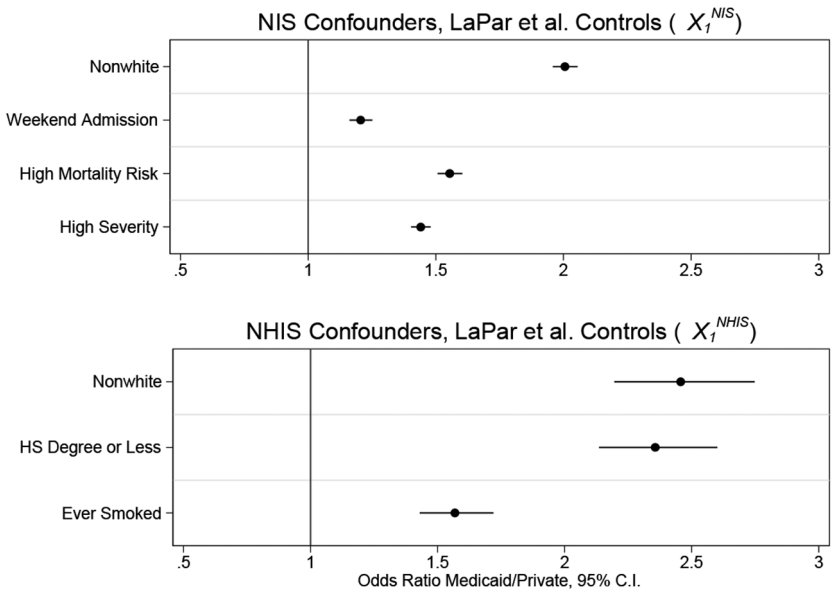


FIG. 2.—Adjusted odds ratios for confounders: Medicaid versus private insurance. This figure reports odds ratios for potential confounders, X_1^{NIS} and X_1^{NHIS} . The reference group is private insurance as primary payer status, and odds ratios for Medicare and uninsured patients are reported in tables A1 and A2. Nationwide Inpatient Sample (NIS) models use the NIS discharge weight. Covariates from LaPar et al. (2010) include patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. Extra controls include age dummies, nonwhite, major loss of function, and major mortality risk. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. National Health Interview Survey (NHIS) models include adults (age 18 or older) from the 1997–2009 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Covariates mimicking LaPar et al. (2010) are age, sex, condition indicators, and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor's degree holder, and dummies for smoking frequency. CI = confidence interval; HS = high school.

likely than private patients to be nonwhite (odds ratio, 2.46; 95% CI, 2.20–2.75). They are also more likely to have a high school diploma or less (odds ratio, 2.36; 95% CI, 2.13–2.60) and to have ever smoked (odds ratio, 1.57; 95% CI, 1.43–1.72).

When we add the extra controls described above (X_1^{NIS} and X_1^{NHIS}), figure 1 shows that the estimated mortality differences fall by more than half. We cannot simultaneously control for both sets of additional variables because

they exist in separate data sets. Doing so may further shrink or eliminate the adjusted mortality differences. In the post-double-selection LASSO model, we find very little change in the NIS odds ratio when we more flexibly control for health conditions coded in International Classification of Diseases, Ninth Revision (ICD-9), codes but a decrease in the NHIS odds ratio where we control for many more socioeconomic status variables.

One way to test the internal validity of our expanded model is to test for balance in even more covariates (X_2^{NHIS}). Figure 3 shows that Medicaid patients in the NHIS are more likely to be below the 25th percentile of height for their age and sex group (odds ratio, 1.18; 95% CI, 1.04–1.33), to have ever been homeless (odds ratio, 2.22; 95% CI, 1.82–2.71), or to have a mental health difficulty (odds ratio, 1.37; 95% CI, 1.17–1.62). Among diabetics, Medicaid patients are more likely to have been diagnosed as minors (odds ratio, 1.47; 95% CI, 0.81–2.69). Among ever-smokers, Medicaid patients are more likely to have started before age 16 (odds ratio, 1.20; 95% CI, 1.04–1.38). Adding extra controls for socioeconomic status, education, race, and smoking status does not eliminate uncontrolled differences by insurance status.

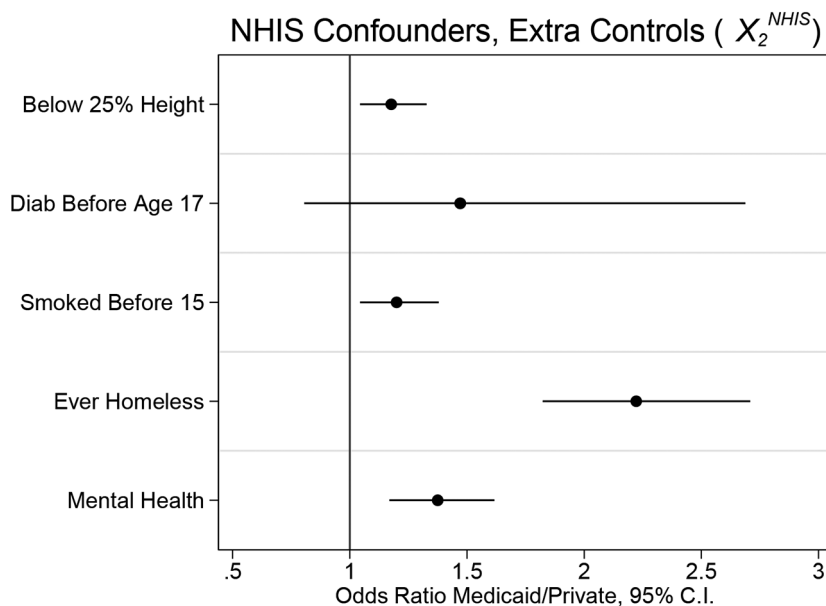


FIG. 3.—Adjusted odds ratios for early-life health confounders: Medicaid versus private insurance. This figure reports odds ratios for predetermined health variables, X_2^{NHIS} . The reference group is private insurance as primary payer status. Controls include the baseline National Health Interview Survey (NHIS) variables (X_0^{NHIS}) and the additional controls used in figure 1 (X_1^{NHIS}). CI = confidence interval; diab = diabetes.

Figure 4 plots average odds ratios calculated from bivariate probit models that vary the correlation between unobserved determinants of mortality and unobserved determinants of Medicaid coverage. The top two points assume that $\rho = 0$ and compare estimates with and without our extra controls

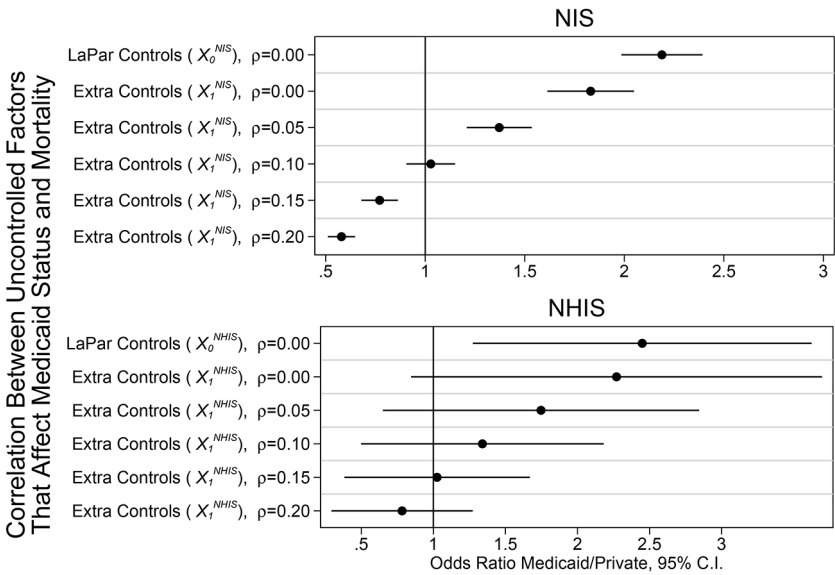


FIG. 4.—Sensitivity of estimates to the magnitude of selection bias. This figure reports adjusted average odds ratios for in-hospital mortality (Nationwide Inpatient Sample [NIS]) and mortality in survey year of or following surgery (National Health Interview Survey [NHIS]) from bivariate probit models for Medicaid insurance and mortality. The first two specifications set ρ , the correlation between unobserved determinants of mortality and unobserved determinants of Medicaid coverage, to zero and are equivalent to a single probit model for mortality. The other specifications represent implied odds ratios under the assumption that ρ equals various values. Samples include Medicaid and privately insured patients. NIS models use the NIS discharge weight. Covariates from LaPar et al. (2010) include patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. Extra controls include age dummies, nonwhite, major loss of function, and major mortality risk. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. NHIS models include adults (age 18 or older) from the 1997–2009 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Covariates mimicking LaPar et al. (2010) are age, sex, condition indicators, and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor’s degree holder, and dummies for smoking frequency. CI = confidence interval.

to be comparable to the logit results from figure 1.¹⁰ When we allow for the possibility of unobserved selection through correlated error terms, the estimated average odds ratios fall by about half (relative to 1) when $\rho = 0.05$ and fall to about 1, implying no mortality differential, when $\rho = 0.10$. At higher values of ρ , we estimate that Medicaid patients have lower mortality than privately insured patients.

Table 2 formalizes the comparison of unobservable and observable selection by comparing the amount of unobserved selection necessary to account for the adjusted mortality differences to the amount of observed selection implied by our statistical adjustments. Row 1 shows the probit coefficient, which equals the amount of bias from unobservables under the null hypothesis of no treatment effect ($\alpha = 0$). Row 2 shows the bias that would result under the equal selection assumption. The ratio of these two numbers (row 3) shows how unobservable selection would have to compare to observable selection in order for bias to explain the entire estimate. In both data sets, this ratio is much less than 1, indicating that the adjusted mortality differences can be explained by unobservables that vary less than the included covariates. The equal selection assumption more than explains the adjusted mortality differences. Standardized gaps in unobserved mortality determinants between Medicaid and private patients only need to be 16%–28% as large as the standardized gaps in observed mortality determinants to fully account for the estimated effect of Medicaid. The estimates appear quite sensitive to even small unobservable differences.

V. Discussion

The results above replicate a common and influential finding that Medicaid patients have higher mortality rates than privately insured patients even after adjusting for observable risk factors (LaPar et al. 2010, 2011). Our analysis shows that adjustments do not adequately control for the selection of patients with lower socioeconomic status and worse health into Medicaid.

A. Imbalance

We find clear evidence that Medicaid patients in the NIS enter the hospital in worse health than privately insured patients. They have 44% higher odds of having a high-severity diagnosis, 55% higher odds of having high predicted mortality risk, and 21% higher odds of having been admitted on a weekend. These results show that the adjusted mortality ratios in LaPar et al. (2010), which are repeatedly cited in policy debates (Suderman 2010; Gottlieb 2011;

¹⁰ These specifications lead to slightly higher average odds ratios because of the different functional form and restricted sample that includes only Medicaid and private patients, but the pattern is the same: Medicaid patients are much more likely to die following surgery than private patients, and this effect is smaller but still present with extra control variables.

Table 2
Amount of Selection on Unobservables Relative to Selection on Observables Required to Attribute the Entire Medicaid Effect to Selection Bias

Quantity	Description	NIS	NHIS
Single-equation probit coefficient on a Medicaid dummy	Amount of unobserved selection bias under null of no treatment effect	.185	.271
$\frac{V(D)}{V(\tilde{D})} \frac{(E[X'\gamma D = 1] - E[X'\gamma D = 0])}{V(X'\gamma)}$	Amount of unobserved selection bias under null of no treatment effect and the assumption that unobserved selection is of the same magnitude as observed selection	.671	1.757
Ratio	How much would unobservables have to shift relative to observables to explain the entire estimated effect?	.275	.154

NOTE.—This table uses the same expanded specifications applied to the sample of Medicaid and privately insured patients as in the second row of fig. 3. NHIS = National Health Interview Survey; NIS = Nationwide Inpatient Sample.

Roy 2011; Cruz 2013; American Action Forum 2014; Antos et al. 2015; Indiana 2016), represent the combination of Medicaid’s causal effects and differences in preoperative risk. When we control for these factors, we find adjusted mortality differences that are about half as large as in the published results.

If one believes that health at admission is itself affected by (past) Medicaid coverage, though, then these measures of health at admission are outcomes instead of covariates (Rosenbaum 2002). They represent causal mechanisms for Medicaid’s mortality effects and should not be controlled. (Note that this claim applies equally well to the included comorbidities, which are treated as crucial control variables in the observational literature.) In fact, this is a common interpretation: “most doctors refuse to see Medicaid patients . . . these patients can’t get access to routine checkups and preventive care . . . that, in turn, leads to poorer health outcomes” (Roy 2011).

Even under the extreme assumption that preoperative risk is not a source of bias, a causal interpretation of the mortality ratios still requires that Medicaid and privately insured patients have similar predetermined characteristics, such as demographics or earlier-life health. Our results show that they do not. In our expanded model, Medicaid patients are more likely than privately insured patients to be nonwhite, to have a high school degree or less, and to have ever smoked. When we control for these factors in the NHIS, the estimated mortality ratios fall by about half.

One might conclude that our expanded models solve the underlying biases in the replication results and that the remaining mortality differences really are causal. Our analysis refutes this. For one, even in our expanded model, Medicaid patients in the NHIS still appear to be less healthy than privately insured patients in ways that cannot have been caused by Medicaid and cannot be controlled for in administrative data like the NIS. They are shorter (a common proxy for earlier life health and economic conditions;

Steckel 1995), more likely to have been diagnosed with diabetes as minors, more likely to have initiated smoking before age 16, and more likely to have ever been homeless. Adult Medicaid coverage cannot have caused any of these differences, yet they matter for health and directly falsify a cross-sectional design, even one based on a richer set of control variables than are available in most observational data sets.

B. Selection on Unobservables

Our sensitivity analyses show that even small amounts of unobservable selection can explain adjusted mortality differences by insurance status. For example, we find that the correlation between unobserved determinants of Medicaid participation and mortality need only be 0.1 to account for their positive covariate-adjusted association.¹¹ One challenge with this result is how to judge whether it is actually a small amount of unobserved selection.

One strategy to gauge the size of this correlation in unobservables is to calculate the correlation between the index of observables in the Medicaid equation, $X'\beta$ in equation (1), and the index of observables in the mortality equation, $X'\gamma$ in equation (2).¹² This effectively asks how the observed portion of Medicaid participation relates to the observed part of mortality. In the NHIS sample, the two are strongly correlated, with an estimated ρ of 0.33. We also obtain high correlations between the indices of condition variables only ($\rho = 0.25$), between indices of demographic and socioeconomic

¹¹ For comparison, when Altonji, Elder, and Taber (2005) apply this method to estimates of the effect of Catholic high school on test scores, their preferred estimates fall by about half, remain statistically significant for $\rho = 0.3$, and remain positive for values of ρ as high as 0.5.

¹² Following Rosenbaum and Rubin (1983), we also consider how large ρ could be if u and ϵ contained a single omitted dummy variable, W , plus independent standard normal error. Note that this violates the model in eqq. (1)–(3), since u and ϵ cannot be jointly normally distributed if they include a dummy. This example simply illustrates how ρ relates to specific omitted variables. Denote the coefficient on W in the Medicaid equation by π (this reflects imbalance), the coefficient on W in the mortality equation by ϕ , and the probability that $W = 1$ by p . These assumptions imply that

$$\rho = \frac{\pi\phi p(1-p)}{\sqrt{\pi^2 p(1-p) + 1} \cdot \sqrt{\phi^2 p(1-p) + 1}}.$$

This expression reaches 0.1 only under fairly extreme values. Even if we assume that $p = 0.5$ (which maximizes the variance of W) and that both π and ϕ equal 0.5 (large values given the scale of the probit index functions), ρ would equal only 0.058. The only two variables in the NHIS mortality equation that have coefficients above 0.5 are the dummies for having cancer or muscular dystrophy. Twenty-one variables in the Medicaid equation have probit coefficients that exceed 0.5 in absolute value (mostly income and employment variables). Therefore, the covariate-adjusted mortality differences cannot be explained by a single omitted dummy variable.

variables only ($\rho = 0.31$), and even between indices of education and poverty variables only ($\rho = 0.65$). Therefore, as we find in table 2, the cross-sectional finding could easily be explained by sets of unobservables that are much more weakly correlated than our observables.

We also benchmark plausible values of ρ using a sample of 23,295 baseline survey respondents from the Oregon Health Insurance Experiment. Using the lottery experiment, we estimate a local average treatment effect of Medicaid on mortality of -0.005 (0.003). We then repeat the exercise in figure 4 to find value of ρ that equates a cross-sectional estimate with the experimental effect.¹³ We do this separately for the treatment and control groups because winning the lottery changes selection into Medicaid by making it easier to obtain: 13.3% of lottery losers ever receive Medicaid, compared with 44% of winners. About 10% of non-Medicare recipients in our observational samples are on Medicaid, so selection patterns probably resemble Oregon Health Insurance Experiment lottery losers more than lottery winners. As expected, we find stronger negative selection (i.e., a more positive Medicaid/mortality relationship) in the control group than in the treatment group.

In both groups, however, we find that the ρ necessary to equate the experimental estimate and the cross-sectional estimate exceeds 0.1. For lottery winners, the correlation between Medicaid and mortality unobservables that make the cross-sectional estimates equal the experimental effect is about 0.17. For lottery losers it is about 0.27. Therefore, the one context where it is possible to benchmark unobservable selection bias to an experimental effect shows selection into Medicaid that it is noticeably larger than required to fully eliminate the positive cross-sectional relationship between Medicaid and mortality.

Our second sensitivity analysis showed that unobservables needed to vary only by 16%–28% as much as the observables to explain the cross-sectional relationship between Medicaid and mortality. Here we provide some specific quantifiable examples of unobserved variables that suggest that differences in unobservables of this magnitude are likely.

Income is one obvious variable that determines Medicaid eligibility, matters for health, and is completely or partly omitted from many data sets. The NIS, for example, records only the quartile of mean per capita income of each patient's zip code. Two patients from the same income quartile may be from areas with very different income levels or may be the poorest and richest residents of a given zip code. Two adjacent zip codes on the south side of Chicago (60621 and 60620) both fell in the bottom income quartile in 2007, but average income in one was 1.5 times as large as in the other (Internal Revenue

¹³ Control variables included dummies for race/ethnicity, age (over or under age 55), sex, education, family income bins (grouped for responses of \$15,000 or greater), and baseline self-reported health status.

Service 2017). Such coarse economic controls mean that comparisons by Medicaid status are also largely comparisons by poverty status.

Income is a clear example of an unobserved confounder whose bias we can measure using auxiliary data. Assume that the mortality regression includes mean income for some grouping g , \bar{x}_g , but excludes deviations of family income relative to those means: $\tilde{x}_i \equiv x_i - \bar{x}_{g(i)}$. Abstracting from other variables and denoting the coefficient on income in the mortality equation by ϕ , the equal selection assumption (4) is

$$\frac{E[\bar{x}_g \phi | D = 1] - E[\bar{x}_g \phi | D = 0]}{V(\bar{x}_g \phi)} = \frac{E[\tilde{x}_i \phi | D = 1] - E[\tilde{x}_i \phi | D = 0]}{V(\tilde{x}_i \phi)}. \quad (5)$$

We can directly calculate the ratio of the unobserved and observed income only using data on income and Medicaid status (and do not even need to know ϕ):

$$\frac{E[\tilde{x}_i | D = 1] - E[\tilde{x}_i | D = 0]}{E[\bar{x}_g | D = 1] - E[\bar{x}_g | D = 0]} \frac{V(\bar{x}_g)}{V(\tilde{x}_i)}. \quad (6)$$

This quantity is comparable to the results in table 2 that unobservables need to vary only by 16%–28% as much as observables to explain the cross-sectional results. This exercise uses income alone to judge whether that amount of unobserved differences is plausible.

We calculate the relative differences in omitted versus included income components using auxiliary data from the 2009 American Community Survey (ACS) and various definitions of the included income groups.¹⁴ We find that the ratio is almost 4 when we use state means; 0.3 when we use means by state, race, sex, and education categories; 0.24 when we use national family income quartiles; and 0.09 when we use national family income deciles. All of these groupings are more detailed than what is available in the NIS, for example, and so we take this as evidence that omitted components of income alone likely vary enough to explain the cross-sectional findings.¹⁵

¹⁴ The ACS does not provide information on respondents' zip codes, so we cannot exactly match the income variable from the NIS. We can, however, measure income by many more than four categories, which increases the variance in the included measure, \bar{x}_g , at the expense of variation in the excluded measure, \tilde{x}_i .

¹⁵ Prognostic factors are similarly mismeasured. The NIS includes indicators for comorbidities identified by Elixhauser et al. (1998), and commentators have described these as especially important controls (Roy 2011). But those comorbidities were chosen explicitly to "eliminate the main reason for hospitalization" and are not meant to capture differences in severity (Elixhauser et al. 1998). Adjusting for these secondary comorbidities alone means that strong differences in preoperative risk remain between Medicaid and privately insured patients.

C. Interpreting Bias

In addition to our empirical results on the potential role for unobservable selection, the structure of Medicaid policy and economic theory further suggest that this bias is likely quite serious. Selection likely comes from health information known only to patients. Those who know they are in worse health are more likely to take up insurance because the financial protections are worth more to them (Akerlof 1970; Rothschild and Stiglitz 1976). This adverse selection leads to bias from unobservables because the health of observationally similar patients who make different insurance choices would differ under any circumstances.

Neither imperfect controls nor inherently unobservable confounders can be addressed by adding more covariates. For this reason, experimental (Finkelstein et al. 2012; Baicker et al. 2013) and quasi-experimental (Currie and Gruber 1996a, 1996b; Aizer 2007; Sommers, Baicker, and Epstein 2012; Bronchetti 2014; Sommers 2014; Wherry and Meyer 2015; Goodman-Bacon 2018; Miller and Wherry 2019) studies rely on policy changes that generate treatment and control groups independently of patient choices (Stukel et al. 2007). The internal validity of these studies is up for debate as well (Kaestner 2012), but standard practice is to describe the identifying variation clearly and justify the identifying assumptions using balance tests, falsification tests, and robustness checks. The pattern of treatment effects that emerge from quasi-experimental work differ by era, patient type, and study design but often find that public insurance provides valuable health benefits. There are important cases in which it does not, but no credibly designed study concludes that public insurance harms health.

On the other hand, the ability to predict clinical outcomes in a diverse group of patients clearly matters for providers, and cross-sectional observational studies can answer this question. If the goal is to provide guidance about which patients may be at risk of adverse health outcomes on the basis of easily observed characteristics, none of the biases we identify above are a problem. It might not matter to a clinician whether Medicaid patients do poorly after surgery because of Medicaid's presumed deficiencies or because they are a vulnerable population. Helping providers identify the riskiest patients on the basis of salient characteristics can provide value independent of a causal question. Nevertheless, these associations, whatever their clinical utility, should not be used as a basis for policy prescriptions that depend crucially on a program's causal effects.

VI. Conclusion

Our aim is not to argue for one kind of Medicaid policy over another. Instead, the goal of this study is to emphasize why cross-sectional observation research cannot contribute to debates over Medicaid's causal effects on health. That said, the descriptive finding that public insurance recipients have worse

health outcomes than privately insured patients provides crucial context for how these programs operate. These correlations reveal complex socioeconomic patterns that interact with Medicaid’s statutory purpose of serving low-income patients. They are an important challenge for Medicaid, not a consequence of it.

Appendix

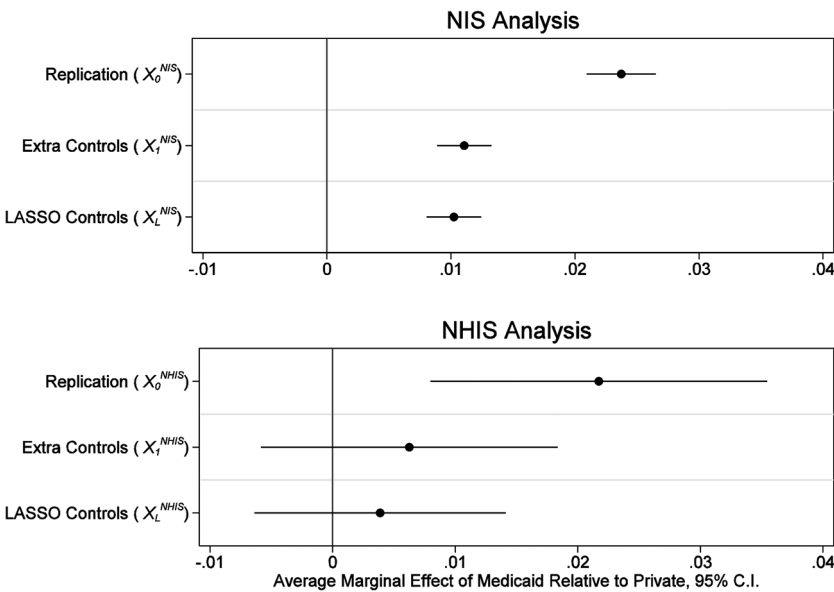


FIG. A1.—Average marginal effects for mortality: Medicaid versus private insurance. This figure reports adjusted average marginal effects for in-hospital mortality (Nationwide Inpatient Sample [NIS]) and mortality in survey year of or following surgery (National Health Interview Survey [NHIS]). The reference group is private insurance as primary payer status. NIS models use the NIS discharge weight. Covariates from LaPar et al. (2010) include patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. Extra controls include age dummies, nonwhite, major loss of function, and major mortality risk. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. NHIS models include adults (age 18 or older) from the 1997–2009 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Covariates mimicking LaPar et al. (2010) are age, sex, condition indicators, and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor’s degree holder, and dummies for smoking frequency. CI = confidence interval.

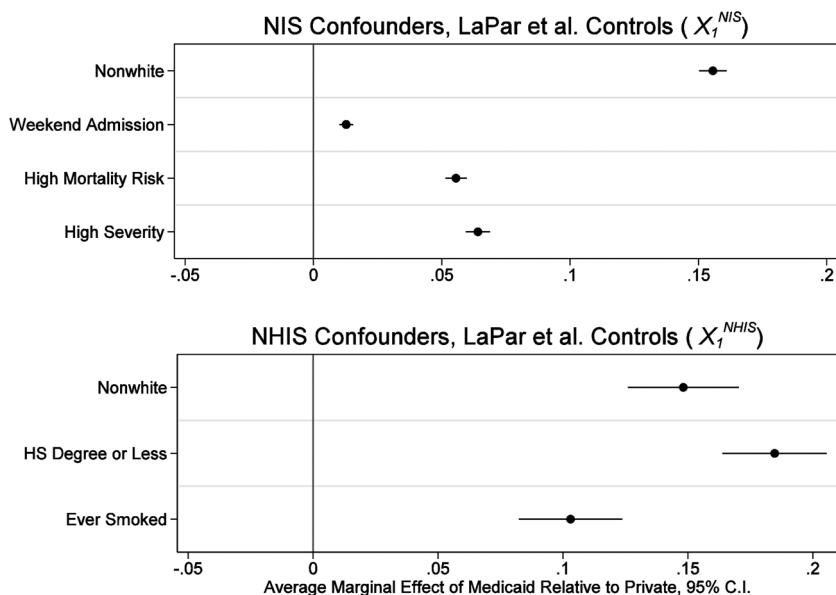


FIG. A2.—Average marginal effects for confounders: Medicaid versus private insurance. This figure reports average marginal effects for potential confounders, X_1^{NIS} and X_1^{NHIS} . The reference group is private insurance as primary payer status. Nationwide Inpatient Sample (NIS) models use the NIS discharge weight. Covariates from LaPar et al. (2010) include patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. Extra controls include age dummies, nonwhite, major loss of function, and major mortality risk. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. National Health Interview Survey (NHIS) models include adults (age 18 or older) from the 1997–2009 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Covariates mimicking LaPar et al. (2010) are age, sex, condition indicators, and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor's degree holder, and dummies for smoking frequency. CI = confidence interval; HS = high school.

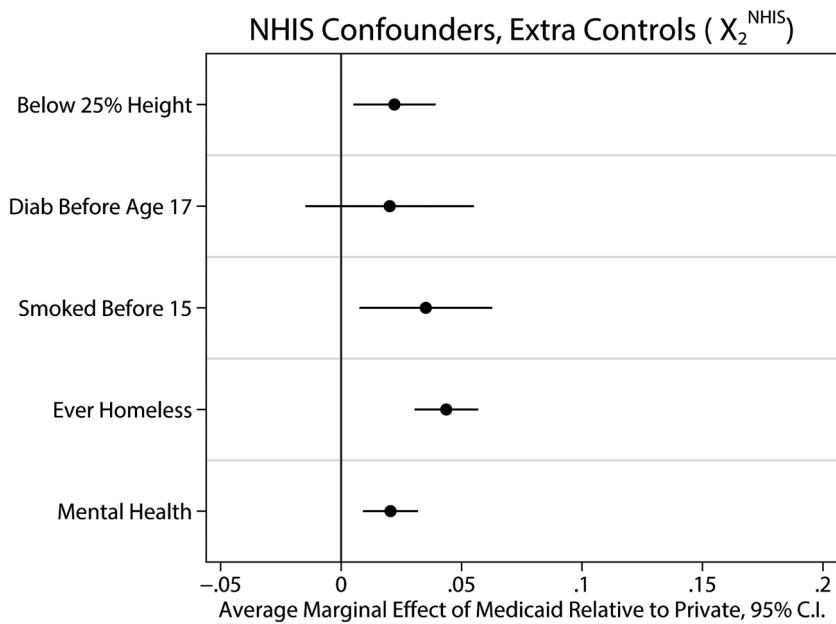


FIG. A3.—Average marginal effects for early-life health confounders: Medicaid versus private insurance. This figure reports average marginal effects for predetermined health variables, X_2^{NHIS} . The reference group is private insurance as primary payer status. Controls include the baseline National Health Interview Survey (NHIS) variables (X_0^{NHIS}) and the additional controls used in figure 1 (X_1^{NHIS}). CI = confidence interval; diab = diabetes.

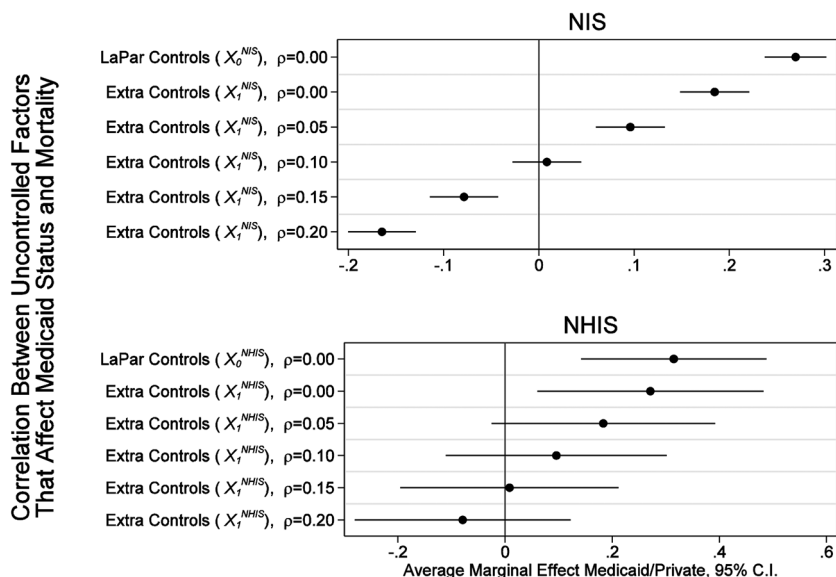


FIG. A4.—Sensitivity of estimates to the magnitude of selection bias, average marginal effects. This figure reports average marginal effects for in-hospital mortality (Nationwide Inpatient Sample [NIS]) and mortality in survey year of or following surgery (National Health Interview Survey [NHIS]) from bivariate probit models for Medicaid insurance and mortality. The first two specifications set ρ , the correlation between unobserved determinants of mortality and unobserved determinants of Medicaid coverage, to zero and are equivalent to a single probit model for mortality. The other specifications represent implied marginal effects under the assumption that ρ equals various values. Samples include Medicaid and privately insured patients. NIS models use the NIS discharge weight. Covariates from LaPar et al. (2010) include patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. Extra controls include age dummies, nonwhite, major loss of function, and major mortality risk. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. NHIS models include adults (age 18 or older) from the 1997–2009 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Covariates mimicking LaPar et al. (2010) are age, sex, condition indicators, and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor's degree holder, and dummies for smoking frequency. CI = confidence interval.

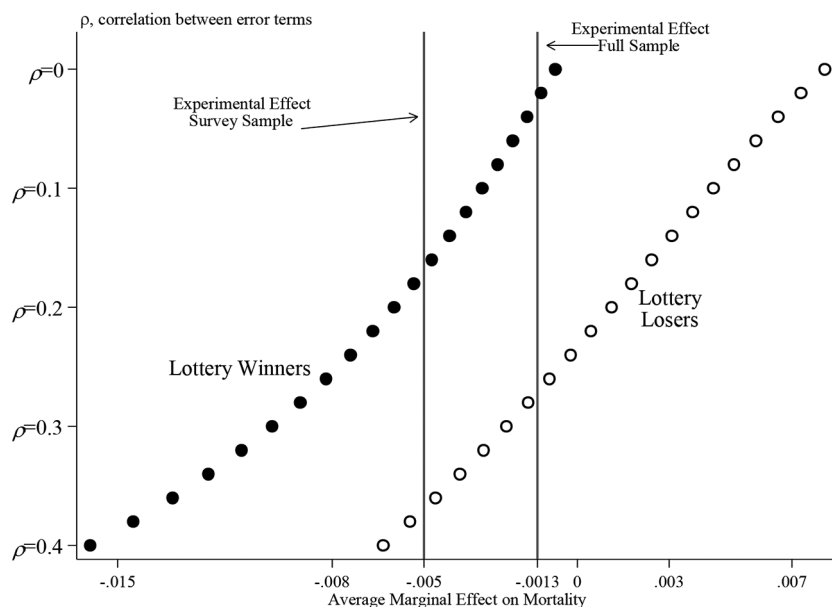


FIG. A5.—Sensitivity of cross-sectional estimates in the Oregon Health Insurance Experiment (OHIE) to the magnitude of selection bias. This figure includes 23,295 baseline survey respondents in the OHIE. We split the sample into lottery winners and losers. The X -axis measures the average marginal effect of ever receiving Medicaid in the OHIE follow-up period on dying in the same period. Estimates come from bivariate probit models that vary the correlation in unobservables between the Medicaid selection equation and the mortality outcome equation, ρ , on the Y -axis (ordered in the same way as fig. 4). Average marginal effects of Medicaid on mortality for lottery winners (44% ever on Medicaid) are in closed circles, and those for lottery losers (13% ever on Medicaid) are in open circles. The two vertical lines indicate the experimental effect in the survey sample (-0.005 ; $SE = 0.003$) and in the full sample reported in Finkelstein et al. (2012). Controls include dummies for white, Hispanic, black, over age 55, female, all values of baseline self-reported health status, education, and family income (all responses of \$15,000 or greater are grouped).

Table A1
Nationwide Inpatient Sample (NIS) Outcomes, Adjusted Odds Ratios by Payer Type

Variable	Death (LaPar et al. Controls) (1)	Death (Additional Controls) (2)	Nonwhite (3)	Weekend Admission (4)	Major Mortality Risk (5)	Major Loss of Function (6)
Medicare	1.55*** (1.49–1.62)	1.27*** (1.21–1.34)	.93*** (.92–.94)	1.08*** (1.06–1.11)	1.43*** (1.40–1.46)	1.35*** (1.33–1.37)
Medicaid	1.98*** (1.86–2.11)	1.46*** (1.37–1.57)	2.01*** (1.96–2.05)	1.21*** (1.16–1.25)	1.55*** (1.51–1.60)	1.44*** (1.40–1.48)
Uninsured	1.74*** (1.61–1.89)	1.55*** (1.42–1.69)	1.48*** (1.43–1.52)	1.37*** (1.32–1.43)	1.35*** (1.30–1.41)	1.26*** (1.22–1.30)
Observations	858,867	858,867	860,756	860,756	860,756	860,756

NOTE.—This table reports odds ratios for adjusted in-hospital mortality and other outcomes. All models are weighted by the NIS discharge weight. The reference group is private insurance as primary payer status. Outcomes are adjusted for patient age, gender, elective operative status, mean income, hospital geographic region, teaching hospital status, type of surgical operation, primary payer status, and categories for Agency for Healthcare Research and Quality comorbid disease. The sample includes all patients in the NIS data set for the years 2003–7 undergoing at least one of eight major surgical operations. Shown in parentheses are 95% confidence intervals.

*** $p < .01$.

Table A2
National Health Interview Survey (NHIS) Outcomes, Adjusted Odds Ratios by Payer Type

Variable	Death (LaPar et al. Controls) (1)	Death (Additional Controls) (2)	Nonwhite (3)	Less than High School Degree (4)	Does Not Smoke (5)	Short (6)	Diabetes before Age 18 (7)	Smoked before Age 15 (8)	Homeless (9)	Mental Health (10)
Medicare	1.82*** (1.44–2.30)	1.48*** (1.10–1.99)	.98 (.90–1.08)	1.44*** (1.34–1.54)	.83*** (.77–.89)	1.07 (.94–1.22)	2.30*** (1.36–3.87)	1.03 (.90–1.18)	1.39*** (1.07–1.80)	1.22*** (1.02–1.46)
Medicaid	2.06*** (1.44–2.95)	1.27 (.83–1.94)	2.46*** (2.20–2.75)	2.36*** (2.13–2.60)	1.57*** (1.43–1.72)	1.18*** (1.04–1.33)	1.47 (.81–2.69)	1.20*** (1.04–1.38)	2.22*** (1.82–2.71)	1.37*** (1.17–1.62)
Uninsured	.86 (.54–1.38)	.64* (.40–1.03)	2.00*** (1.81–2.21)	1.84*** (1.69–2.00)	1.49*** (1.37–1.63)	1.12** (1.01–1.25)	1.33 (.68–2.61)	1.13** (1.00–1.28)	2.20*** (1.86–2.60)	1.41*** (1.21–1.65)
Observations	45,309	42,187	45,691	45,693	45,693	45,498	5,223	23,006	32,839	45,489

NOTE.—Covariates from LaPar et al. (2010) are age, sex, condition indicators (see below), and categories of family income (as opposed to zip code income). Additional covariates are dummies for age, detailed race, Hispanic ethnicity, citizen status, marital status, high school graduate, bachelor's degree holder, and dummies for smoking frequency. (Weight regression also includes height as a covariate.) The sample includes sample adults (age 18 or older) from the 1997–2014 NHIS who were eligible for the mortality follow-up and report having had any surgery in the past 12 months. Shown in parentheses are 95% confidence intervals, clustered on survey strata.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

References

- Aizer, Anna. 2007. Public health insurance, program take-up, and child health. *Review of Economics and Statistics* 89, no. 3:400–415.
- Akerlof, George A. 1970. The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84, no. 3:488–500.
- Altonji, Joseph, Todd Elder, and Christopher Taber. 2005. Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy* 113, no. 1:151–84.
- American Action Forum. 2014. Expanding Medicaid—harming those who need it most. In *Insight*, February 19.
- Antos, Joseph, James C. Capretta, Lanhee J. Chen, Scott Gottlieb, Yuval Levin, Thomas P. Miller, Ramesh Ponnuru, Avik Roy, Gail R. Wilensky, and David Wilson. 2015. Improving health and health care: An agenda for reform. American Enterprise Institute, Washington, DC.
- Ayanian, John Z., Betsy A. Kohler, Toshi Abe, and Arnold M. Epstein. 1993. The relation between health insurance coverage and clinical outcomes among women with breast cancer. *New England Journal of Medicine* 329, no. 5:326–31.
- Baicker, Katherine, Sarah L. Taubman, Heidi L. Allen, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Eric C. Schneider, Bill J. Wright, Alan M. Zaslavsky, and Amy N. Finkelstein. 2013. The Oregon experiment — effects of Medicaid on clinical outcomes. *New England Journal of Medicine* 368, no. 18:1713–22.
- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2013. Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies* 81, no. 2:608–50.
- Borgschulte, Mark, and Jacob Vogler. 2020. Did the ACA Medicaid expansion save lives? *Journal of Health Economics* 72:102333.
- Bronchetti, Erin Todd. 2014. Public insurance expansions and the health of immigrant and native children. *Journal of Public Economics* 120:205–19.
- Canto, John G., William J. Rogers, William J. French, Joel M. Gore, Nisha C. Chandra, and Hal V. Barron. 2000. Payer status and the utilization of hospital resources in acute myocardial infarction. *Archives of Internal Medicine* 160, no. 6:817–23.
- Card, David, Carlos Dobkin, and Nicole Maestas. 2009. Does Medicare save lives? *Quarterly Journal of Economics* 124, no. 2:597–636.
- Cruz, Ted. 2013. Congressional Record. <https://www.govinfo.gov/content/pkg/CREC-2013-03-13/pdf/CREC-2013-03-13-pt1-PgS1719.pdf>.
- Currie, Janet, and Jonathan Gruber. 1996a. Health insurance eligibility, utilization of medical care, and child health. *Quarterly Journal of Economics* 111, no. 2:431–66.
- . 1996b. Saving babies: The efficacy and cost of recent changes in the Medicaid eligibility of pregnant women. *Journal of Political Economy* 104, no. 6:1263–96.

- Elixhauser, Anne, Claudia Steiner, D. Robert Harris, and Rosanna M. Coffey. 1998. Comorbidity measures for use with administrative data. *Medical Care* 36, no. 1:8–27.
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. The Oregon Health Insurance Experiment: Evidence from the first year. *Quarterly Journal of Economics* 127, no. 3:1057–106.
- Fiscella, Kevin, Peter Franks, Marthe R. Gold, and Carolyn M. Clancy. 2000. Inequality in quality: Addressing socioeconomic, racial, and ethnic disparities in health care. *Journal of the American Medical Association* 283, no. 19:2579–84.
- Frakt, Austin, Aaron E. Carroll, Harrold A. Pollack, and Uwe Reinhardt. 2011. Our flawed but beneficial Medicaid program. *New England Journal of Medicine* 364, no. 16:e31.
- Franks, Peter, Carolyn M. Clancy, and Marthe R. Gold. 1993. Health insurance and mortality: Evidence from a national cohort. *Journal of the American Medical Association* 270, no. 6:737–41.
- Goodman-Bacon, Andrew. 2018. Public insurance and mortality: Evidence from Medicaid implementation. *Journal of Political Economy* 126, no. 1: 216–62.
- Gottlieb, Scott. 2011. Medicaid is worse than no coverage at all. *Commentary*, March 10.
- Haas, Jennifer S., and Lee Goldman. 1994. Acutely injured patients with trauma in Massachusetts: Differences in care and mortality, by insurance status. *American Journal of Public Health* 84, no. 10:1605–8.
- Hammarlund, Noah. 2019. Racial treatment disparities after machine learning surgical-appropriateness adjustment. Unpublished manuscript.
- Heckman, James J., and Edward Vytlacil. 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73, no. 3:669–738.
- Higgins, Robert S. D., Gaetano Paone, Steven Borzak, Gordon Jacobsen, Edward Peterson, and Norman A. Silverman. 1998. Effect of payer status on outcomes of coronary artery bypass surgery in blacks. *Circulation* 98, suppl. 19:II46–II49; discussion, II49–II50.
- Houglund, Paul, Jonathan Nebeker, Steve Pickard, Mark Van Tuinen, Carol Masheter, Susan Elder, Scott Williams, and Wu Xu, eds. 2008. *Advances in patient safety: New directions and alternative approaches*, vol. 1. Rockville, MD: Agency for Healthcare Research and Quality.
- Howell, Embry M., and Genevieve M. Kenney. 2012. The impact of the Medicaid/CHIP expansions on children: A synthesis of the evidence. *Medical Care Research and Review* 69, no. 4:372–96.
- Indiana, State of. 2016. Healthy Indiana Plan demonstration. Project no. 11-W-00296/5. Annual report, reporting period February 1, 2015, to January 31, 2016. <https://www.medicaid.gov/Medicaid-CHIP-Program-Information>

- /By-Topics/Waivers/1115/downloads/in/Healthy-Indiana-Plan-2/in-healthy-indiana-plan-support-20-annl-rpt-feb-jan-2016-04292016.pdf.
- Internal Revenue Service. 2017. Individual income tax zip code data (SOI). Washington, DC: Internal Revenue Service.
- Jolly, Nicholas A., and Brian J. Phelan. 2017. The long-run effects of job displacement on sources of health insurance coverage. *Journal of Labor Research* 38, no. 2:187–205.
- Kaestner, Robert. 2012. Mortality and access to care after Medicaid expansions. *New England Journal of Medicine* 367, no. 25:2453–54.
- Kelz, Rachel Rapaport, Phyllis A. Gimotty, Daniel Polsky, Sandra Norman, Douglas Fraker, and Angela DeMichele. 2004. Morbidity and mortality of colorectal carcinoma surgery differs by insurance status. *Cancer* 101, no. 10:2187–94.
- LaPar, Damien J., Castigliano M. Bhamidipati, Carlos M. Mery, George J. Stukenborg, David R. Jones, Bruce D. Schirmer, Irving L. Kron, and Gorav Ailawadi. 2010. Primary payer status affects mortality for major surgical operations. *Annals of Surgery* 252, no. 3:544–50; discussion, 550–51.
- LaPar, Damien J., Castigliano M. Bhamidipati, Dustin M. Walters, George J. Stukenborg, Christine L. Lau, Irving L. Kron, and Gorav Ailawadi. 2011. Primary payer status affects outcomes for cardiac valve operations. *Journal of the American College of Surgeons* 212, no. 5:759–67.
- LaPar, Damien J., George J. Stukenborg, Richard A. Guyer, Matthew L. Stone, Castigliano M. Bhamidipati, Christine L. Lau, Irving L. Kron, and Gorav Ailawadi. 2012. Primary payer status is associated with mortality and resource utilization for coronary artery bypass grafting. *Circulation* 126:S132–S139.
- Levy, Helen, and David Meltzer. 2004. What do we really know about whether health insurance affects health? In *Health policy and the uninsured*, ed. Catherine G. McLaughlin. Washington, DC: Urban Institute Press.
- Lin, D. Y., B. M. Psaty, and R. A. Kronmal. 1998. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* 54, no. 3:948–63.
- Miller, Sarah, Sean Altekruze, Norman Johnson, and Laura R. Wherry. 2019. Medicaid and mortality: New evidence from linked survey and administrative data. NBER Working Paper no. 26081, National Bureau of Economic Research, Cambridge, MA.
- Miller, Sarah, and Laura R. Wherry. 2019. The long-term effects of early life Medicaid coverage. *Journal of Human Resources* 54, no. 3:785–824.
- Minnesota Population Center and State Health Access Data Assistance Center. 2012. Integrated Health Interview Series: Version 5.0. Minneapolis: University of Minnesota.
- Office of the Press Secretary. 2017. Press briefing by Press Secretary Sean Spicer. White House, March 7. <https://trumpwhitehouse.archives.gov/briefings-statements/press-briefing-press-secretary-sean-spicer-030717/>.

- Roetzheim, Richard G., Naazeen Pal, Colleen Tennant, Lydia Voti, John Z. Ayanian, Annette Schwabe, and Jeffery P. Krischer. 1999. Effects of health insurance and race on early detection of cancer. *Journal of the National Cancer Institute* 91, no. 16:1409–15.
- Rosen, Heather, Fady Saleh, Stuart R. Lipsitz, John G. Meara, and Selwyn O. Rogers. 2009. Lack of insurance negatively affects trauma mortality in US children. *Journal of Pediatric Surgery* 44, no. 10:1952–57.
- Rosenbaum, Paul R. 1984. The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society Series A* 147, no. 5:656–66.
- . 2002. *Observational studies*. 2nd ed. Springer Series in Statistics. New York: Springer.
- Rosenbaum, P. R., and D. B. Rubin. 1983. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society Series B* 45, no. 2:212–18.
- Rothschild, Michael, and Joseph Stiglitz. 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics* 90, no. 4:629–49.
- Roy, Avik. 2011. Why Medicaid is a humanitarian catastrophe. *Forbes*, March 2. <http://www.forbes.com/sites/theapothecary/2011/03/02/why-medicaid-is-a-humanitarian-catastrophe/>.
- Salim, Ali, Marcus Ottochian, Joseph DuBose, Kenji Inaba, Pedro Teixeira, Linda S. Chan, and Daniel R. Margulies. 2010. Does insurance status matter at a public, level I trauma center? *Journal of Trauma: Injury, Infection, and Critical Care* 68, no. 1:211–16.
- Sommers, Benjamin D. 2014. Impact of Medicaid expansions on mortality. Working paper, Harvard School of Public Health.
- Sommers, Benjamin D., Katherine Baicker, and Arnold M. Epstein. 2012. Mortality and access to care among adults after state Medicaid expansions. *New England Journal of Medicine* 367, no. 11:1025–34.
- Sommers, Benjamin D., and Arnold M. Epstein. 2013. U.S. governors and the Medicaid expansion—no quick resolution in sight. *New England Journal of Medicine* 368, no. 6:496–99.
- Steckel, Richard H. 1995. Stature and the standard of living. *Journal of Economic Literature* 33, no. 4:1903–40.
- Stukel, Thérèse A., Elliot S. Fisher, David E. Wennberg, David A. Alter, Daneil J. Gottlieb, and Marian J. Vermeulen. 2007. Analysis of observational studies in the presence of treatment selection bias: Effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *Journal of the American Medical Association* 297, no. 3:278–85.
- Suderman, Peter. 2010. Paying more for less. *Reason*, August 6. <http://reason.com/archives/2010/08/06/paying-more-for-less>.
- Sullivan, Daniel, and Till von Wachter. 2009. Job displacement and mortality: An analysis using administrative data. *Quarterly Journal of Economics* 124, no. 3:1265–306.

- Wherry, Laura R., and Bruce D. Meyer. 2015. Saving teens: Using a policy discontinuity to estimate the effects of Medicaid eligibility. *Journal of Human Resources* 51, no. 3:556–88.
- Zacharias, Anoar, Thomas A. Schwann, Christopher J. Riordan, Samuel J. Durham, Aamir Shah, and Robert H. Habib. 2005. Operative and late coronary artery bypass grafting outcomes in matched African-American versus Caucasian patients: Evidence of a late survival-Medicaid association. *Journal of the American College of Cardiology* 46, no. 8:1526–35.