THE GOOD DOCTORS: COST AND MORTALITY EFFECTS OF U.S. PCPS

Jason Abaluck, Zachary Bleemer, Peter Hull, and Amanda Starc*

March 2026

Preliminary and Incomplete; Do Not Cite

Abstract

We develop measures of the health and cost impacts of primary care physicians (PCPs) serving Medicare patients in the United States. We validate the causal content of these measures using physician retirements and exits, and we use them to investigate the origins of efficient and high-quality care, the importance of PCPs for insurer network adequacy, and the potential benefits of lower-quality PCP exit. There is substantial variation in PCP quality and assignment to high-quality PCPs actually lowers expenditure slightly. Medical school quality and other training characteristics predict which physicians deliver better outcomes at lower cost. Primary care physicians are considerably more important than hospitals in accounting for variation in Medicare Advantage plan impacts. Simulations suggest that redirecting patients from bottom-quintile PCPs to average-quality providers could reduce mortality by 10-15% without increasing spending.

JEL: I11, C26, I13.

*Abaluck: jason.abaluck@yale.edu, Yale University School of Management and NBER; Bleemer: bleemer@princeton.edu, Princeton University and NBER; Hull: peter_hull@brown.edu, Brown University and NBER; Starc: amanda.starc@kellogg.northwestern.edu, The Kellogg School of Management, Northwestern University, and NBER. We thank Josh Byrne, Will Cox, and Nina Low for expert research assistance. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number P30AG012810. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

1 Introduction

An increasing body of evidence suggests that differences in the healthcare providers from which you receive care can lead to differences in outcomes for observably similar patients. Existing studies have documented substantial heterogeneity in hospital (Geweke et al., 2003; Hull, 2020; Chandra et al., 2023), physician (Badinski et al., 2023a; Ginja et al., 2025), and even insurance plan quality (Abaluck et al., 2021). We aim to extend and synthesize this work. We begin by constructing novel estimates of physician mortality and cost impacts for primary care physicians (PCPs) serving Medicare patients in the United States. We validate the causal content of these measures leveraging quasi-experimental variation in physician retirements and exits.

We then ask a series of questions. First, are there identifiable characteristics that predict which physicians deliver high-quality care at low cost? We identify observable traits that predict higher quality; most notably, training, as measured by medical school quality, predicts PCP quality. Next, we combine our physician estimates with hospital-level quality measures and link both to plan-level effects. Are some insurance plans good because of the hospitals and physicians they include in network? We find that plan and network quality depend more on the included PCPs than on hospitals. Could network market forces or regulation lead to better patient outcomes without increasing costs by rewarding plans for increasing the market-share of predictably high-quality and cost-efficient physicians? We simulate regulations that would reward plans for the quality of the PCPs they include in network.

Our analysis starts with the universe of Medicare fee-for-service claims to comprehensively estimate PCP value-added for traditional Medicare patients. The central empirical challenge is to separate whether some providers causally make patients healthy compared to other providers, or merely treat healthier patients to begin with. We construct observational measures of physician value-added, noting that variation could be due to selection or treatment or both. Our observational risk-adjusted mortality measure accounts for observable differences in patient health and statistical noise. We then show substantial variation across physicians; a physician one standard deviation better than average has mortality rates that are 13% lower than average.

We validate our measures using quasi-experimental variation in provider choice sets, following Abaluck et al. (2021). To estimate causal effects, we examine situations in which a provider-patient relationship is severed. Following Ginja et al. (2025), Molitor (2018), Fadlon and Van Parys (2020), Badinski et al. (2023b), Sabety (2023), and Kwok (2019), our research design leverages physician moves and retirements. For example, when high- or low-mortality physicians retire, their patients sort into more average-looking physicians. We ask whether mortality decreases in cohorts induced to switch from high-mortality physicians to low-mortality physicians.

Our value-added measures have predictive validity: cohorts of patients who see better value-

added PCPs are more likely to die when those providers exit, and cohorts who see worse PCPs become less likely to die following exit. That is, the IV framework shows that risk-adjusted mortality is a predictor of true PCP mortality effects. We estimate an IV forecast coefficient of 0.71. This finding does not rule out selection across PCPs. Yet, on average, the variation in risk-adjusted PCP mortality accurately predicts variation in causal mortality effects (just as hospital quality measures can reliably predict the causal impact of hospitals.) We repeat the same procedure using "Traditional" Medicare expenditures as the left-hand side outcome and validating physician expenditure effects. The regression then recovers how being assigned to a higher- or lower-quality PCP affects total utilization. Our estimates suggest that being displaced to a higher mortality PCP actually leads to higher overall utilization, though the effects are relatively modest. This suggests there is no explicit tradeoff between quality and cost at the provider level.

Next, we estimate the relationship between PCP mortality effects and physician characteristics. With regard to human capital, the Medical College Admissions Test (MCAT) and students' medical school and residency rankings strongly predict doctor quality. Female and more-experienced doctors are also higher-quality. Value-added is partially observed by patients: among doctors with public 'quality' ratings on websites like Vitals.com or WebMD.com, an additional star (on a scale from one to five) is associated with 0.2 improved mortality value-added.

Finally, we quantify the aggregate implications of our measured PCP quality variation. We first show that PCP quality is a far more important driver of plan-level outcomes than hospital quality. Using share-weighted averages of PCP value-added as a plan's "inclusive value," we find a strong relationship between PCP inclusive values and plan-level mortality. The analogous relationship with hospital inclusive values is substantially weaker, and hospital quality adds little explanatory power once PCP quality is controlled for. We then present two counterfactual simulations. Following Chetty et al. (2014b), we simulate a system-wide reallocation in which patients are redirected away from the bottom 5% of PCPs; mortality falls by approximately 5.3% of baseline. While infeasible in practice, this exercise illustrates the magnitude of gains from directing patients to better PCPs. We then consider a more feasible alternative: each insurer drops the bottom 5% of its in-network PCPs. This plan-level intervention recovers approximately 93% of the gains from the system-wide counterfactual, underscoring the potential for network design to improve patient outcomes.

Our paper makes several contributions. First, we develop general and validated measures of PCP quality and cost effects. In doing so, we incorporate both a broader literature that measures quality within healthcare (including hospitals, as discussed above, as well as plans; e.g. Abaluck et al. (2021)) and elsewhere in the economy (in particular, teachers; e.g. Kane and Staiger (2008) and Chetty et al. (2014a)). In addition to speaking to the literature looking at the impact of physician quality disclosure and pay-for-performance programs, our results have important conse-

quences for physician incentive systems and compensation (Clemens and Gottlieb, 2017; Gottlieb et al., 2023). Second, we help synthesize the sources of variations in healthcare outcomes, which speaks to a large literature examining the causal impact of place, plans, providers, and hospitals Finkelstein et al. (2016, 2021).

The paper is organized as follows. Section 2 describes the data and institutional setting. Section 3 describes the measurement and validation of PCP value added. Section 4 relates provider value added to physician characteristics. Section 5 describes policy simulations highlighting the importance of PCPs and networks. Section 6 concludes.

2 Setting and Data

Medicare is a federal health insurance program in the United States designed primarily for individuals aged 65 and older. "Traditional" Medicare (TM) covers a variety of services. Part A covers hospital stays, hospice care, and some skilled nursing facility stays. Part B covers outpatient services, preventive care, and some doctor visits. Part D is the prescription drug coverage component, and is administered by private firms. Individuals can also opt for Medicare Advantage plans offered by private insurers.

TM enrollees benefit from a wide network of providers, including most U.S. physicians and hospitals. In fact, 93% of physicians in Internal Medicine accept Medicare assignment, ensuring nearly universal access to care.[1] This broad network makes TM one of the least restrictive insurance options in the U.S. Yet research highlights the crucial role of physicians, especially PCPs, in improving patient outcomes (Badinski et al., 2023b). Medicare beneficiaries can choose to opt out of TM by enrolling in Medicare Advantage (MA), which consists of privately managed care plans.[2] One key difference between MA and TM is selective contracting. Unlike TM's nearly universal provider network, MA plans often restrict their networks. While networks may lower costs or improve quality by directing patients to efficient, high-quality providers, they can also limit access and continuity of care.

In this paper, we use extensive data on Medicare, including claims, to measure the impact of PCPs. In this section, we outline the data and key variables of interest; additional detail is available in the Appendix. We use data on the universe of Medicare beneficiaries aged 65 or older in one of 50 U.S. states or the District of Columbia from 2012 to 2019. For each beneficiary in each year, we

[1] https://www.kff.org/medicare/how-many-physicians-have-opted-out-of-the-medicare-program/

[2] These plans aim to enhance consumer choice and cost efficiency while matching TM's basic benefits. MA insurers receive a capitated, risk-adjusted payment for each enrollee, incentivizing them to manage care differently from TM. A significant number of beneficiaries have transitioned to MA in recent years. Yet our own recent work has shown that MA plans vary dramatically in quality, as measured by mortality rates. Existing publicly available star ratings are not correlated with causal plan effects and previous work has not documented mechanisms. The rapid growth of MA emphasizes the need to understand what drives plan quality, especially as networks become central to competition.

observe the identity of their selected plan (both MA and TM), their local market (county), standard beneficiary demographics (age, sex, race, and dual-eligible status), and their end-of-year mortality status. For 20% of traditional Medicare enrollees, we further observe inpatient and outpatient claims. This allows us to assign beneficiaries to PCPs.

Our primary outcomes of interest are mortality and healthcare utilization. We observe end-of-year mortality status for all beneficiaries. For the 20% sample with claims data, we construct a measure of healthcare utilization following the existing literature (Finkelstein et al., 2016). The claims data also allow us to observe beneficiary demographics (age, sex, race, dual-eligible status, and county of residence) and to impute physician clinical specialty.

To determine a beneficiary's PCP, we first restrict attention to "General Practice," "Family Practice," or "Internal Medicine," based on HCFA specialty codes. If the beneficiary saw no departing PCPs, then we define the time period as the calendar year. For each calendar year, we assign a beneficiary to its modal PCP in that time period, breaking any ties randomly. We then define exiting and non-exiting PCPs, based on the procedure described in Sabety (2020). If a beneficiary saw a single departing PCP, we create 12 month periods that end on the month of exit (not including that month). For example, if a beneficiary saw a PCP that exited in October, then we define the time period as October-September. We then assign the modal PCP in each of these time periods.

We develop a physician-patient assignment algorithm to better reflect PCP retirements and transitions. [3] To illustrate, consider a beneficiary whose regular PCP retired in May 2015. This beneficiary may appear to have seen two non-departing primary care physicians–one before and one after May–without any indication of retirement. With our revised method, however, the time period adjusts to reflect May-to-April intervals. This makes sure the beneficiary is linked to their old PCP until April 2015 and their new PCP beginning in May 2015. The revised assignment captures the retirement and enables precise analysis of how changes in PCP quality impact patients. As a result, we are comparing PCP transitions this year to those that have occurred in the past or are expected to occur in the future. We adjust the relevant time period to run from the month of departure (e.g., May to April).

Table I describes our beneficiary sample. Column 1 shows the full sample of PCP-time periods, while columns 2–4 restrict to PCPs who terminate during our observation window. The average beneficiary is 75 years old; 87% of beneficiaries are white, 61% are male, and 16% are dual-eligible. In the period when PCPs terminate (column 3), nearly all beneficiaries (99%) switch doctors, compared to only 17% in non-termination periods. One-year mortality is substantially higher in termination periods (2.9%) than in non-termination periods (0.8%). (We note that we do

[3] Assuming a traditional calendar year, if a beneficiary's regular PCP retires in the middle of the year, they would not be assigned to a new PCP. To resolve this, we center the assignment around the month of PCP exit. We kept observations only if they had a lagged PCP that left during this time. This way, the analysis only examined beneficiaries affected by PCP departures.

Table I: Summary Statistics

		Sample (PCPs terminated)		
	All PCPs	All PCPs	Terminated	Non-Terminated
Beneficiary Demographics				
Beneficiary Age	75.1	75.9	76.1	75.8
% White	87.1	89.3	87.5	89.4
% Male	60.9	61.8	61.8	61.8
% Dual Eligible	15.6	15.7	15.8	15.7
% Switched Doctors	14.6	23.9	97.3	17.3
% Died	4.2	1.0	2.2	0.9
Total Beneficiary-Periods	147,429,825	6,956,681	568,035	6,388,646
Doctor Demographics				
Doctor Age	50.5	51.2	52.6	50.9
% Male Doctors	67.4	70.4	67.4	71.2
PCP Value-Added	0.003	-0.002	0.003	-0.003
Total Doctor-Periods	20,509,468	415,912	53,607	362,305

Notes: This table summarizes the analysis samples in 2012-2019. Column 1 represents all PCP-time periods, regardless of if they terminated. Column 2 is the sample of PCP-time periods we are using for figure IV, i.e., they terminate during our window of observation. Column 3 represents PCPs in the time period they terminate, and column 4 are PCPs that terminate, in time periods in which they do not terminate. Column 2 is a subset of column 1, and columns 3 and 4 are a subset of column 2.

not use this variation directly.)

We also augment our main data with additional data on PCP characteristics. The second panel of Table I describes our physician sample. Exiting PCPs are slighly older than non-existing PCPs and more likely to be female.

3 PCP VAM Construction and Validation

3.1 Observational Mortality

Despite the long-standing use of physician report cards, we are unaware of any systematic, validated measures of physician value added. We develop such measures here. We proceed in three steps. First, we construct observational value-added estimates. Next, we describe our research design that allows us to test for prediction unbiasedness. Finally, we estimate a forecast regression to validate our observational value-added measure.

We begin our analysis by computing observational differences in one-year mortality rates across PCPs operating in the same area, adjusting for observable differences in plan enrollees and statistical noise. These observational mortality estimates come from ordinary least squares

(OLS) regressions, of the form

$$Y_{it} = \sum_{p} \mu_p D_{ipt} + X_{it}'\omega + \varepsilon_{it}, \tag{1}$$

where Y_{it} is an indicator for beneficiary i dying in year t and D_{ipt} indicates her assignment to a given PCP p at the start of this period. The control vector X_{it} contains observable characteristics of enrollees (age, sex, race, and dual-eligibility status) as well as a full set of county and year fixed effects. We allow the coefficient vector ω to vary flexibly by plan size. Given the fixed effects and controls, variation in the observational mortality coefficients μ_p thus reflects within-area differences in one-year mortality rates among observably similar enrollees. We account for statistical noise in the observational mortality estimates by applying a conventional empirical Bayes correction (Morris, 1983). This correction "shrinks" the estimated μ_p towards their county- and PCP size-level mean, in proportion to their expected degree of estimation error. The shrinkage is larger for PCPs seeing few patients but minimal for the PCPs with large patient panels that make up the majority of our sample.

Estimates of Equation (1) reveal substantial variation in PCP mortality rates among observably similar beneficiaries. The estimated beneficiary-weighted standard deviation of μ_j across MA plans, after correcting for estimation error, is 0.011 percentage points or 14% of the average one-year sample mortality rate of 7.9%. Figure I plots the full distribution of shrunk observational mortality rates across PCPs. The solid line shows this distribution for our baseline specification of Equation (1), with all observable controls included in X_{it}, while the dashed line shows the corresponding distribution for a simpler specification that omits the beneficiary demographic controls.

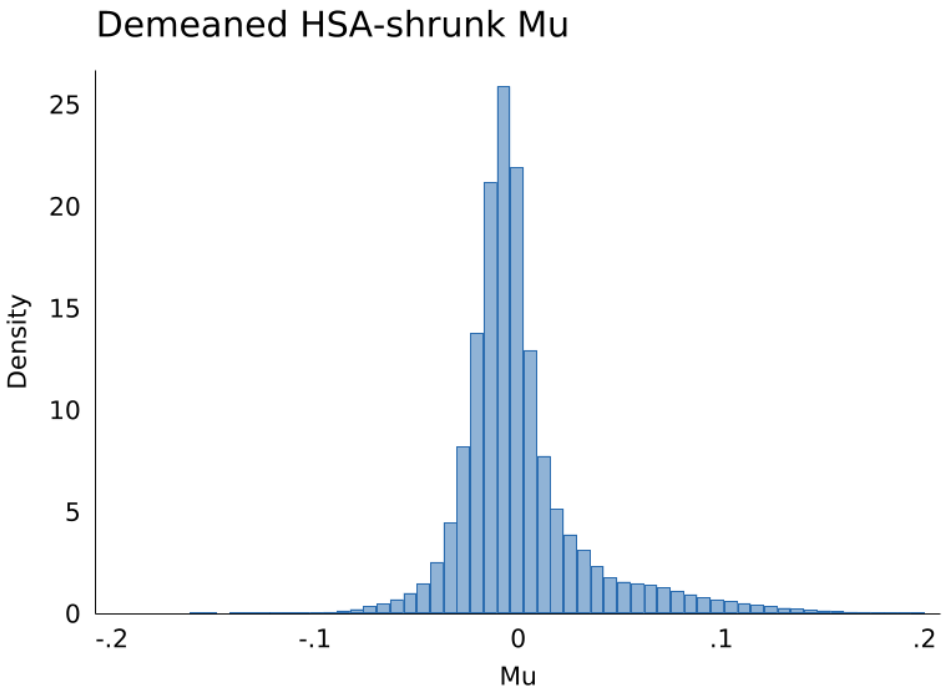


Figure I: Observational Mortality

Notes: This figure summarizes the enrollment-weighted distribution of observational mortality across PCPs, winsorized to account for extreme outliers. The solid dark line shows this distribution when observational mortality is estimated from Equation (1), with all demographic controls, while the light dashed line shows the corresponding distribution for a simpler specification that omits age, race, sex, and dual-eligible status. Average observational mortality across all PCPs is normalized to the average of the full model. Estimates are shrunk via an empirical Bayes procedure. Estimated means and standard deviations of μ_p for PCPs are for the prior distribution and shown for each estimation procedure.

3.2 Research Design

We use an instrumental variables (IV) framework, leveraging PCP exits and leaves, to measure the validity of observational mortality differences in predicting differences in causal mortality effects. While not identifying mortality effects for individual physicians, this approach is sufficient to estimate the expected mortality impact of reallocating beneficiaries across observably different PCPs. We first our research design and then discuss econometric implementation.

Several studies in the literature leverage physician moves and exits, either in a differences-in-differences or "movers design" framework. For example, Sabety (2020), Fadlon and Van Parys (2020), and Kwok (2019) compare patient outcomes between patients with PCPs who exit and PCPs who stay. Identification relies on comparing outcomes within a PCP's group of patients to the matched control group, relative to an omitted time group. Molitor (2018) and Badinski et al.

(2023b) both leverage similar variation within a "movers" framework. Specifically, they study utilization changes for Medicare patients and cardiologists who move across HRRs. They ask whether utilization better reflects the "origin" (indicating large patient or physician effects) or the "destination" (indicating large physician or place effects).

Our research design combines elements of both of these approaches, leveraging the fact that some patients must switch PCPs. If outcome variation is driven by patient characteristics (e.g., health status, preferences), switching PCPs should not significantly change utilization. Conversely, if PCP practice styles (e.g., spending habits) drive outcomes, patients switching PCPs should have mortality rates that match that of their new PCP's other patients. Of course, there might be a direct effect of PCP exits on outcomes, reflecting adjustment costs. As a result, we will compare patients that are forced to switch away from *high-observational-mortality* PCPs to patients that are forced to switch away from *low-observational-mortality* PCPs.

To build intuition for our quasi-experimental approach, consider a set of patients seeing a PCP with a high observational mortality rate μ_p. Since PCP choice is highly inertial, most of these patients will remain with their PCP throughout the following year. Suppose, however, that a PCP exits the market for a plausibly idiosyncratic reason (such as retirement). This termination would force the PCP's patients to choose an alternative physician. Under standard regression-to-the-mean, they will tend to switch to a new PCP that is more typical in terms of μ_p. If the observational mortality rates were causal, then all else equal we would expect the mortality of this patient cohort to fall commensurate to the decline in μ_p. Identical logic holds for patients seeing exogenously exiting PCPs with low observational mortality rates: subsequent PCPs are likely to be more typical in terms of μ_p, relative to patients seeing non-exiting low observational mortality PCPs. If observational mortality variation reflects causal effects, then mortality should rise. Combining these two exit quasi-experiments can reveal the predictive content of our observational mortality rate estimates while allowing for direct termination effects that are common to the high- and low-mortality PCP exits.

3.3 Estimation and Results

Our first goal is to measure the relationship between observational mortality μ_p and true mortality effects β_p. Formally, we seek to estimate the MA forecast coefficient λ, defined by the projection of causal mortality effects β_p on observational mortality μ_p. Normalizing the means of both parameters to zero, this projection can be written

$$\beta_p = \lambda \mu_p + \eta_p, \tag{2}$$

where η_p is mean-zero and uncorrelated with μ_p by definition. This regression is infeasible in the sense that the dependent variable β_p is neither observed nor estimated, despite measurement of the independent variable μ_p. The forecast coefficient nevertheless captures the predictive validity of the observational mortality measures. For example, μ_p is an on-average unbiased predictor of causal mortality effects when $\lambda = 1$, while observational mortality has little association with true causal effects when λ is small.[4] We emphasize that Equation (2) reflects an equilibrium statistical relationship, given by existing patterns of selection, and that λ is not a structural parameter.

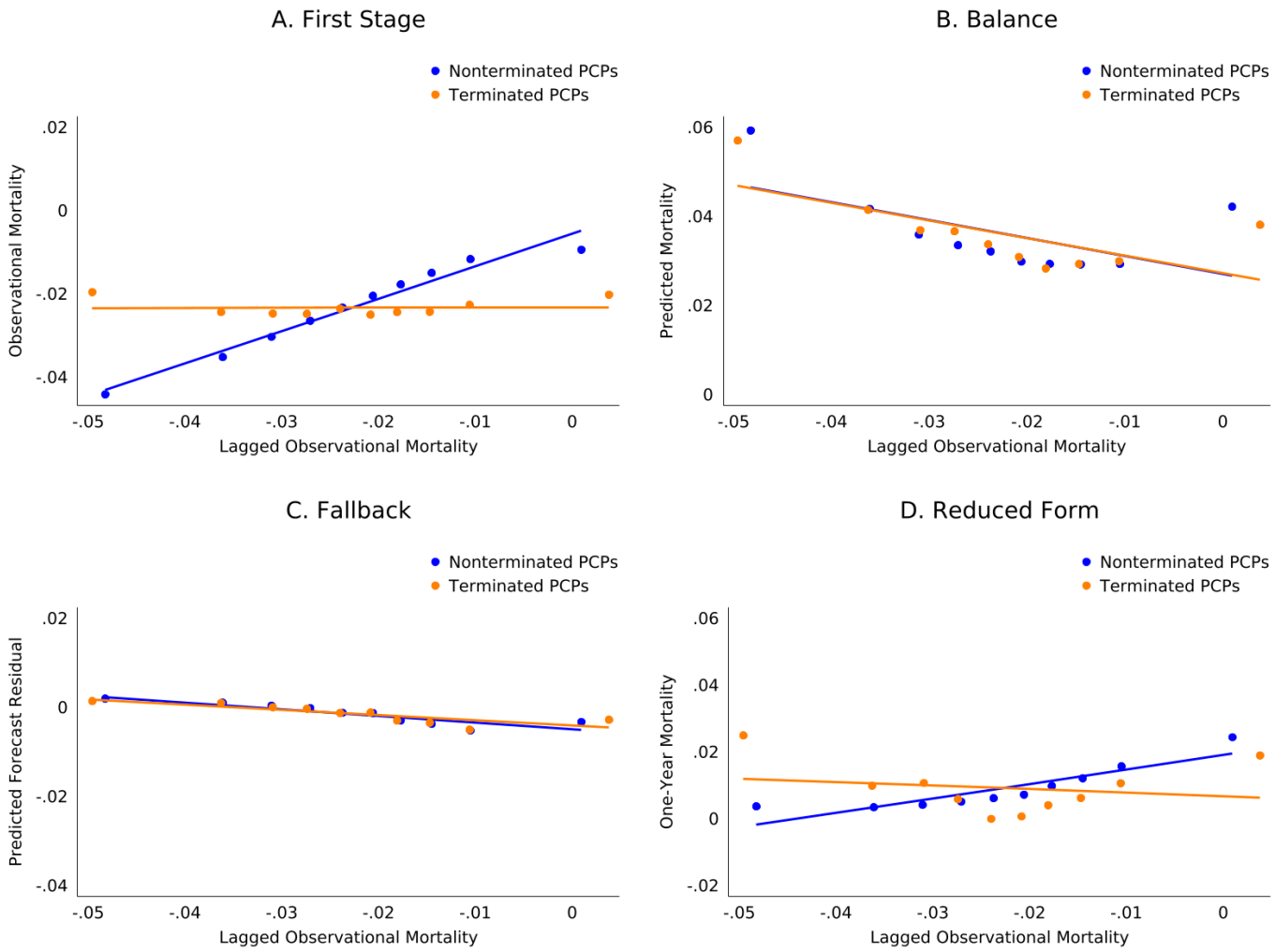
We begin by describing the variation in the data. Panel A of Figure II illustrates the first-stage relationship. We plot contemporaneous PCP μ_p by deciles of lagged observational mortality. We use this specification to plot the estimated contemporaneous PCP observational mortality for enrollees who, in the previous year, had PCPs in different deciles of observational mortality that did and did not exit. While observational mortality of the lagged PCP predicts current plan observational mortality among the non-exiting group, the relationship is essentially flat for exiting PCPs. The flattening reflects the combination of inertia and regression-to-the-mean. We emphasize that only the difference in the slopes used for IV identification.

Our approach requires three assumptions: first-stage relevance, balance, and a fallback condition (Abaluck et al., 2021). We build support for the balance condition by testing whether the instrument predicts observable differences in beneficiary health. We replace the observational mortality outcome with a prediction of one-year beneficiary mortality. In contrast to the clear first-stage effect, there is no differential trend in predicted mortality for exiting versus non-exiting PCPs. We next build support for the fallback condition. We test whether our instrument predicts an observable proxy for the forecast residual η_i. We construct the proxy by first regressing observational mortality on a set of observable PCP characteristics (here, tenure). We then take the residual from projecting the fitted values from this regression (as an observable proxy of β_p) on μ_p. This residual yields an observable proxy for η_p. Panel C illustrates our test. As before, we see no systematic relationship between exits and the predicted PCP unobservable at any decile of lagged observational mortality.

In contrast to Panels B and C, our reduced form (Panel D) echoes the first stage. Lagged PCP observational mortality $\mu_{p,t-1}$ predicts contemporaneous *individual mortality* Y_it for patients seeing the same PCP. However, the relationship is much weaker for patients forced to see new PCPs.

[4] This definition of the forecast coefficient aligns $1-\lambda$ with the notion of "forecast bias" in the education value-added literature (Kane and Staiger (2008), Chetty et al. (2014a)).

Figure II: Graphical Tests of Assumptions and the Reduced Form



Notes: Panel A of this table is based on estimation of Equation (3) and presents the OLS coefficient in a first-stage regression of observational mortality on the instrument. Panel B replaces observational mortality as the dependent variable with a prediction of one-year mortality based on beneficiary demographics. Panel C uses as the dependent variable a prediction of the forecast residual based on physician characteristics (tenure, three digit zip code, medical school quality, and degree type). Panel D presents the forecast coefficient.

To formalize this logic, we next present these IV estimates. We begin by estimating an OLS first-stage regression of

$$\mu_{it} = \pi_Z Z_{it} + X_{it}'\pi_X + v_{it}, \tag{3}$$

where again μ_{it} denotes the physician observational mortality for beneficiary i at time t and $Z_{it} = \mu_{i,t-1} \times T_{i,t-1}$ is the interaction of observational mortality of the lagged plan and an indicator for

lagged plan termination.

While it is not feasible to estimate Equation (2) directly, we can relate it to observed enrollee mortality by estimating

$$Y_{it} = \lambda_p \mu_{it} + X_{it}'\gamma + \varepsilon_{it} + \eta_{it}, \tag{4}$$

where $\mu_{it} = \sum_p \mu_p D_{ipt}$ denotes the observational mortality of beneficiary i given her PCP D_{ipt} and $\eta_{it} = \sum_p \eta_p D_{ipt}$ is the corresponding forecast residual of her selected plan.

As predicted, the first stage is negative in all specifications, reflecting regression-to-the-mean. The magnitudes are stable across specifications, and each estimate is highly significant. Panel B presents balance tests using predicted mortality. Across specifications, the coefficients are consistently small (though often marginally significant). The direction of the coefficients varies, indicating that any balance failures are not systematic.[5] Fallback tests in Panel C show coefficients extremely close to zero, with confidence intervals including zero in all specifications. There is little evidence for residual correlation between the instrument and the forecast error term η.

Finally, Panel D reports the relationship between the instrument and the forecast coefficient. Across specifications, the forecast coefficient notably increases with additional controls, going from 0.63 in the PCP demographic-only model to 0.71 with richer (five-digit ZIP and chronic conditions) controls. Recall that a coefficient of one would indicate predictive unbiasedness. Our forecast coefficient in the final, preferred specification, is less than one (and one lies slightly outside the confidence interval). Nonetheless, we find that our observational measure has substantial predictive power.

3.4 Utilization

We next ask if PCP quality has a causal impact on spending. That is, are observationally better doctors systematically higher or lower cost? The relationship between provider quality and cost is ambiguous: Doyle Jr et al. (2010) argue that higher quality doctors use fewer resources but Doyle Jr et al. (2015) show that higher (Medicare) spending hospitals have lower mortality.

To estimate the causal relationship between observational mortality and total spending, we first construct a spending variable. We then replace Y_{it} in equation 4 with the spending variable S_{it}. We note that the first stage—along with associated balance and fallback tests—remain unchanged. We then estimate:

[5] For the specification in column (3), the coefficient sign flips, but as noted, this model jointly controls for covariates used to generate predicted mortality. Controlling for construction variables likely undermines the interpretability of the balance test in column (3).

Table II: Tests of Assumptions

	(1)	(2)	(3)
Dep. Var.: Observational Mortality		A. First Stage	
Instrument	−0.480	−0.510	−0.510
	(0.014)	(0.015)	(0.021)
F Statistic	367.5	260.5	246.5
Dep. Var.: Predicted Mortality		B. Balance	
Instrument	−0.011	−0.023	0.029
	(0.011)	(0.016)	(0.009)
Dep. Var.: Predicted Forecast Residual		C. Fallback	
Instrument	0.015	0.006	0.015
	(0.002)	(0.002)	(0.003)
Dep. Var.: Forecast Coefficient		D. λ_p	
Instrument	0.561	0.597	0.710
	(0.049)	(0.059)	(0.113)
Dep. Var.: Log Expenditure		E.	
Instrument	3.652	3.672	3.522
	(0.789)	(.906)	(1.295)
Specification	PCP demo.	+3 digit zip	+5 digit zip & CC
Demographic Controls	No	Yes	Yes
N Beneficiary-Years	6,956,681	6,956,681	6,508,756
N Beneficiary-Years (Carrier)	1,098,388	1,098,388	1,098,388

Notes: In column (1) the instrument is the interaction of observational mortality of the lagged plan and a lagged plan termination indicator. In column (2) the instrument is the interaction of an indicator for above-median observational mortality of the lagged plan and a lagged plan termination indicator. In all specifications, we control for the observational mortality of the lagged plan and termination main effects, county-by-year fixed effects, year- and county-specific termination effects, and interactions of lagged plan characteristics (as described in the text). Standard errors are clustered by county and reported in parentheses.

$$S_{it} = \lambda_p, s\mu_{it} + X_{it}'\gamma_s + \varepsilon_{it,s} + \eta_{it,s}. \tag{5}$$

The results are in the final panel of Table II. Moving to a PCP with 1 percentage point higher mortality increases spending by approximately 3.5 to 3.7 percent, depending on the specification. The coefficient is remarkably stable across specifications, ranging from 3.522 in the richest specification with 5-digit ZIP codes and chronic conditions to 3.672 in the specification with 3-digit ZIP codes.

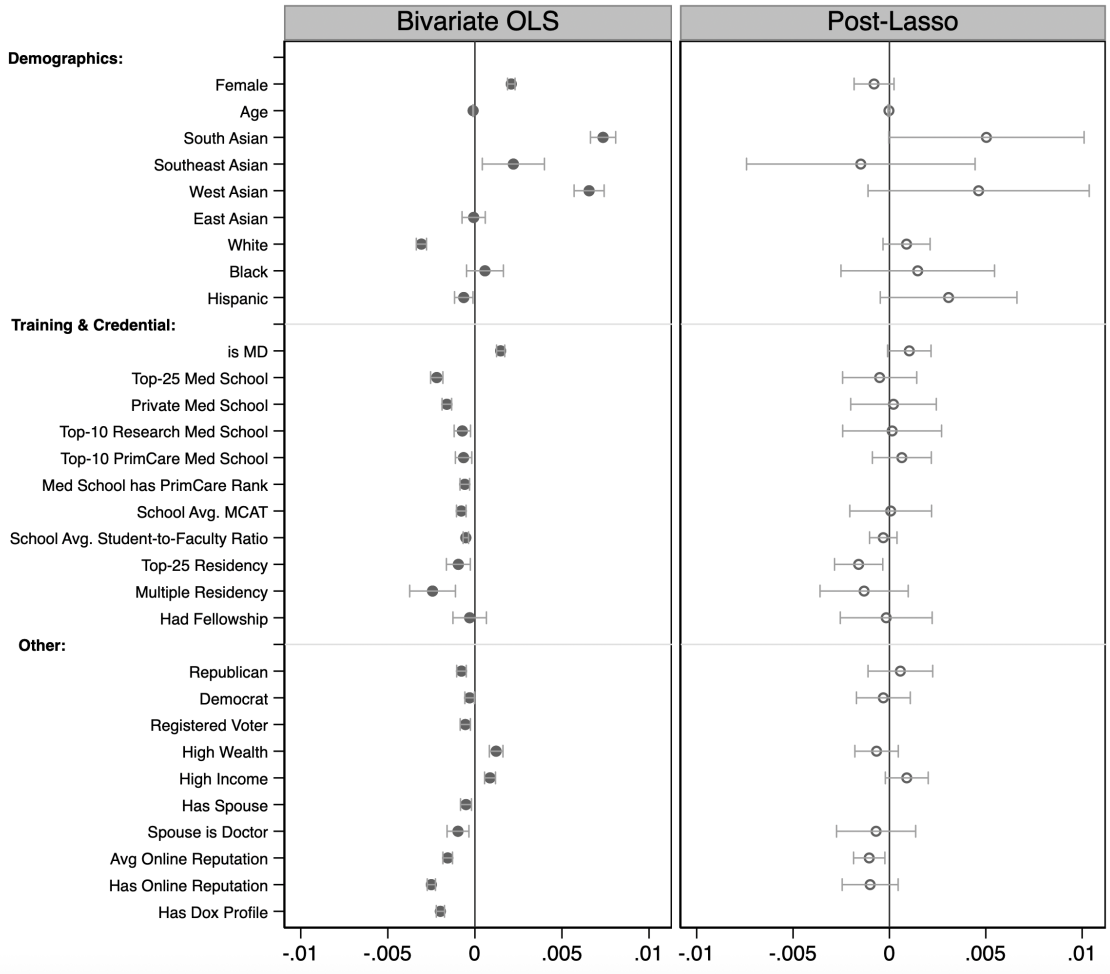
These results indicate there is no explicit tradeoff between quality and cost at the physician level. Higher-mortality PCPs are associated with higher overall utilization, though the magnitude of the effect is relatively modest. This finding suggests that directing patients to lower-mortality physicians could simultaneously improve health outcomes and reduce Medicare expenditures.

4 Correlates of Physician Quality

What predicts physician value-added? In this section, we consider a number of physician and physician practice characteristics that may be correlated with patient outcomes. We begin with physician tenure (as measured at the beginning of our sample). We then consider medical school quality (knowing specialty selection can be different by medical school). *We then use pre-medical school measures of potential aptitude, such as medical school admission scores.* Finally, we consider physician practice characteristics. Specifically, we examine organizational factors such as practice size and ownership.

Figure **??** presents initial results. Each row represents a different PCP characteristics. The left panel presents coefficients from separate bivariate OLS regressions. The points in the right panel are coefficients from post-Lasso multivariate OLS. Table III reports the bivariate OLS coefficients.[6] Across a range of measures, medical school quality predicts physician value-added.

[6] Additional information is presented in Table V.



Notes: This figure summarizes the relationship between physician characteristics and PCP VAM measures. Estimates are shrunk via an empirical Bayes procedure. Estimated means and standard deviations of μ_p for PCPs are for the prior distribution and shown for each estimation procedure.

We first consider demographics. Female and white doctors tend to have lower (better) mortality value-added, while South and West Asian doctors tend to have somewhat higher value-added. Mortality value-added declines in experience, with older doctors seeing less patient mortality.

Medical training is a strong predictor of mortality value-added. Doctors with lower mortality tend to have higher MCAT admission scores and to have attended highly-ranked US medical schools, whereas doctors who attended Caribbean medical schools generally have poorer mortality value-added. The popular US News and World Report medical rankings include separate rankings for both high-quality research and primary care training; while both are separately predictive of better value-added, in a bivariate regression with both covariates the weight loads fully onto the primary care ranking. Doctors who completed highly-ranked residencies and who completed fellowships (as reported on their Doximity profiles) also tend to have better value-added. Strikingly, though, doctors with osteopathic medical training tend to outperform those with more traditional

allopathic training, despite the greater selectivity and prestige of the latter schools.

Finally, we match primary care doctors to the one-to-five-star ratings available from two online sources, Vitals.com and WebMD.com. The 65.6% percent of doctors with at least 10 such cumulative ratings tend to be above-average, and ratings are strongly predictive of low mortality value-added: doctors with one additional star rating have better value-added by 0.17 percentage points.

5 Implications of Quality Variation

In this section, we quantify the aggregate impacts of variation in physician quality. We begin by establishing that PCP quality—not hospital quality—is the primary driver of plan-level outcomes. We then consider two counterfactual simulations. First, we examine an infeasible but informative system-wide exercise: how much would redirecting Medicare beneficiaries away from the lowest-quality PCPs reduce mortality? Second, we turn to a more feasible alternative: plans can capture nearly all of these gains simply by excluding the bottom 5% of their in-network PCPs.

5.1 PCP Networks Predict Plan Quality

In previous work, we argued that there were large differences in causal mortality effects across plans. We explored three mechanisms by which plans may influence beneficiary health: cost-sharing, utilization controls, and provider networks. Results on benefit generosity were suggestive. However, there is still significant variation in plan quality that was left unexplained. Provider networks are a natural candidate mechanism.

In this subsection, we ask three additional questions. First, is there meaningful variation in average PCP quality across private insurance plans? Second, does it matter (for plan quality)? Finally, can insurers effectively screen PCPs and steer enrollees to high quality clinicians?

To quantify network quality at the plan level, we aggregate up to the plan-level by creating an average PCP quality by plan. We refer to this object as the PCP inclusive value, and define it as $v_j = \sum_p \mu_p s_{pj}$, where s_{pj} again represents PCP shares among a plan's enrollees. We follow Feyman et al. (2024) to construct the shares. Their paper constructs measures of network breadth by utilizing prescription drug event (PDE) data. The authors use PDE data from beneficiaries enrolled in stand-alone prescription drug plans from 2011 to 2019. Providers are counted as "in-network" if beneficiaries in a county received prescriptions from them. Their method avoids the inaccuracies of provider addresses and directly captures provider-patient interactions.

A plan with a PCP inclusive value one standard deviation lower than average sends its beneficiaries to PCPs with a lower mortality rate. As a result, it is plausible that network qual-

Table III: PCP Characteristics and Observational Mortality

Provider Characteristic	Coefficient	(SE)
Demographics		
Female	-0.00209	(0.00011)
Age	-0.00010	(0.00001)
South Asian	0.00736	(0.00037)
Southeast Asian	0.00221	(0.00091)
West Asian	0.00656	(0.00044)
East Asian	-0.00007	(0.00034)
White	-0.00307	(0.00015)
Black	0.00058	(0.00054)
Hispanic	-0.00064	(0.00027)
Training & Credentials		
MD (vs. DO)	0.00148	(0.00012)
Top-25 med school	-0.00219	(0.00018)
Private med school	-0.00161	(0.00014)
Top-10 research med school	-0.00072	(0.00024)
Top 10 primary care med school	-0.00065	(0.00024)
Med school has primary care rank	-0.00058	(0.00014)
Med school Avg. MCAT	-0.00078	(0.00014)
Med school Avg. student-to-faculty ratio	-0.00052	(0.00008)
Top-25 residency	-0.00095	(0.00035)
Multiple residency	-0.00243	(0.00067)
Had fellowship	-0.00030	(0.00049)
Other Characteristics		
Republican	-0.00077	(0.00014)
Democrat	-0.00030	(0.00014)
Registered voter	-0.00055	(0.00015)
High wealth	0.00122	(0.00020)
High income	0.00087	(0.00016)
Has spouse	-0.00051	(0.00016)
Spouse is doctor	-0.00097	(0.00032)
Avg. online reputation	-0.00156	(0.00014)
Has online review	-0.00250	(0.00012)
Has dox profile	-0.00198	(0.00012)

Notes: This table reports OLS estimates of the regression of observational mortality on physician characteristics. The average non-missing value across plans replaces missing physician characteristics. Standard errors are clustered by county and reported in parentheses.

ity—measured in this way—predicts plan quality in the aggregate. We then first plot the relationship between a plan's inclusive value v_j (on the x-axis) and a plan's observational mortality μ_j (on the y-axis). To the extent that plans that attract "good" PCPs (as measured by PCP VAM)

are themselves "good" (as measured by plan observational mortality), there will be a strong relationship between the two. If PCPs were responsible for all of the variation in outcomes, we would expect something close to a 1-for-1 relationship. Figure III reports the result. Obviously, in practice, we see a slope that is far less than one; nonetheless, there is a strong relationship between the two measures.

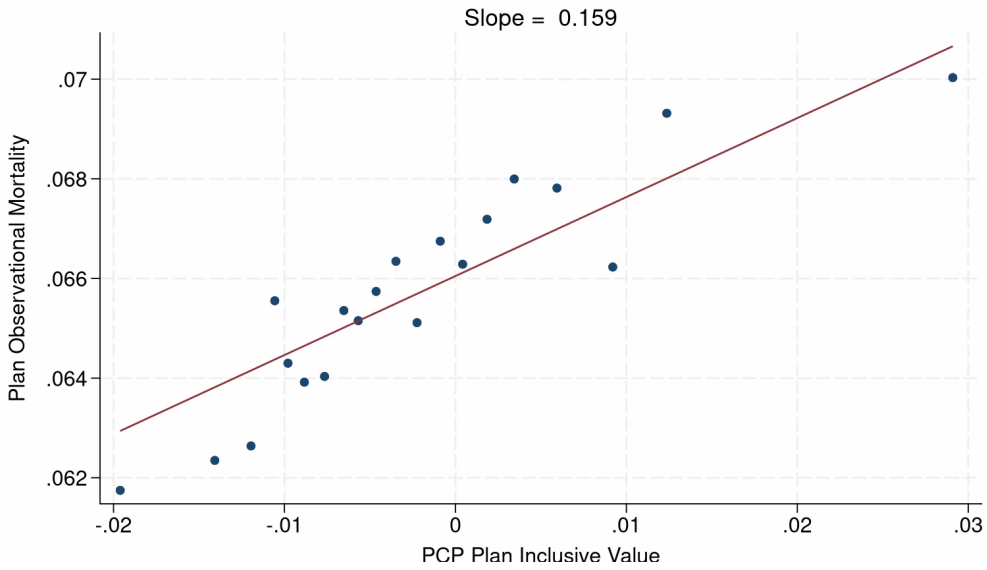
Next, we show that this relationship persists conditional on covariates. Furthermore, PCP quality has more predictive power than hospital quality.[7]

OLS estimates of θ describe the relationship between our hospital inclusive value measure and plan-level observational mortality μ_j. The first column of Table IV shows PCP inclusive values and plan value added are positively correlated. Column 4 further shows that worse PCP inclusive values are associated with higher mortality. Column 2 shows that hospital inclusive value is also correlated with plan value-added, though the relationship is somewhat noisier. Column 3 reveals a compelling pattern: PCP and hospital-inclusive values are correlated. When we regress plan value-added on both, only the relationship with PCP inclusive value remains statistically significant. Hospital-inclusive value is actually "wrong-signed" in the final regression (and also imprecisely measured). Taken together, our results indicate that hospital inclusive values are only weakly correlated with plan-level observational mortality, μ_j. We believe two broad facts explain these patterns in the data. First, in MA, hospital networks are much broader than PCP networks. Second, only a small fraction of enrollees are admitted to the hospital; a much higher fraction see PCPs.

Our further results indicate that physicians are more important drivers of plan quality than hospitals. This finding is consistent with Tilipman (2022), who finds that consumers' preferences in employer-sponsored health insurance are much more strongly driven by access to broad physician networks than by hospital networks. Access to a broad physician network accounts for about 80% of the total utility derived from network breadth. Our finding is also consistent with Badinski et al. (2023b), who find that physician practice styles can explain at least a third of geographic variation in utilization.

[7] Our procedure involves identifying each beneficiary's primary hospital based on the number of visits each year using the MEDPAR data. We again note that MEDPAR covers over 80% of hospitals across all payers, making it a reliable and comprehensive source. We then merge primary hospitals with annual plan information from the CMS denominator files at the beneficiary level. We calculate hospital shares by dividing the number of beneficiary-years assigned to a specific hospital by the total beneficiary-years in that plan. The CDS hospital quality measures from 1992 to 2014 are averaged for each cohort and hospital, excluding data before 2006. We calculate each plan's "inclusive value" by multiplying the providers' mean measure by its share within the plan and summing these values; that is, it is a patient flow-weighted average of provider value-added. We focus here on AMI mortality; robustness checks are available in the appendix.

Figure III: Observational Mortality and Inclusive Values



Notes: This figure summarizes the relationship between enrollment-weighted distribution of provider inclusive value across plans and plan-level observational mortality. Estimates are shrunk via an empirical Bayes procedure. Estimated means and standard deviations of μ_p for PCPs are for the prior distribution and shown for each estimation procedure.

Table IV: Networks are Correlated with Plan Value Added and Individual Mortality

	Unshrunk μ_j			One Year Mortality		
PCP	0.1979		0.1924	0.3553		0.3864
Inclusive Value	(0.0171)		(0.0392)	(0.0269)		(0.0507)
Hospital		0.1173	0.0499		0.1143	-0.0400
Inclusive Value		(0.0217)	(0.0303)		(0.0524)	(0.0624)
N	16,907,013	17,125,368	16,685,303	16,907,013	17,125,368	16,685,303

Notes. This table reports OLS coefficient estimates for the PCP and hospital inclusive values. The dependent variable is unshrunk plan value-added in the first three columns and realized one-year mortality in the next three columns. In columns (1) and (4), the independent variable of interest is PCP inclusive value. In columns (2) and (5), the independent variable of interest is hospital inclusive value. In columns (3) and (6), both inclusive value measures are included as controls. In all specifications, we control for the applicable lagged inclusive value(s) and termination main effects, county-by-year fixed effects, year- and county-specific termination effects, interactions of lagged plan characteristics (as described in the text), and beneficiary demographics. Standard errors are clustered by county and reported in parentheses.

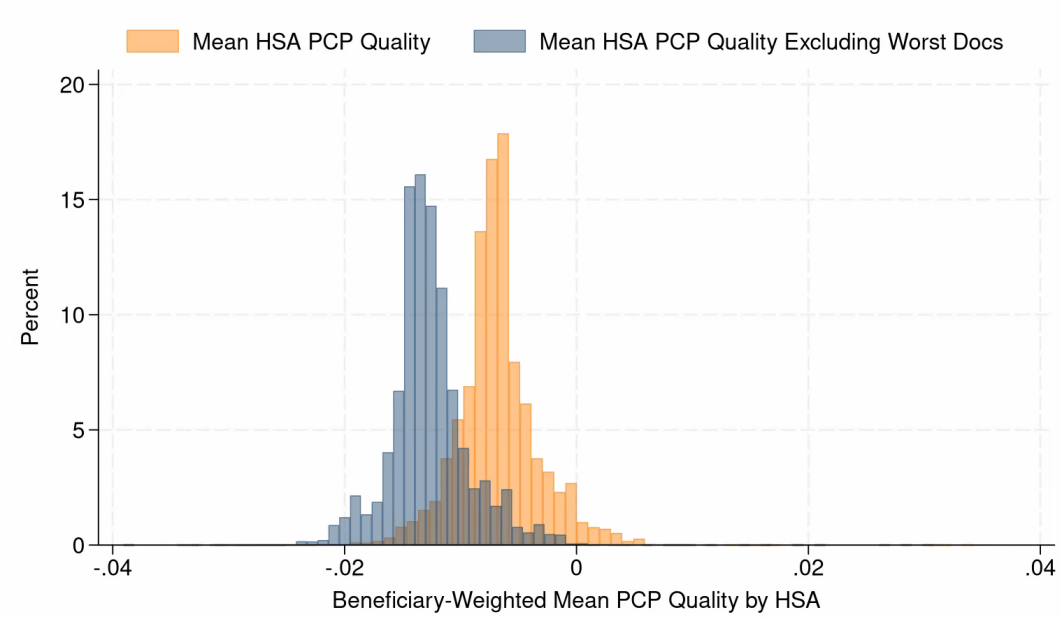
5.2 Aggregate Mortality Effects

Following Hanushek (2011) and Chetty et al. (2014b), we begin by simulating redirecting patients from the lowest 5% of PCPs to average-quality ones. To do so, we make several assumptions. First, we assume that PCP VAM is known without uncertainty. Second, we assume the impact of a one-standard deviation improvement in PCP VAM on mortality is constant over time. Third, we assume there are no capacity limits or general equilibrium effects. Such effects could potentially attenuate gains if all beneficiaries moved to higher-quality PCPs. Finally, we assume that beneficiaries sort to new PCPs in proportion to their market shares, implicitly imposing an IIA structure.

Mechanically, we begin by constructing a patient-weighted average value-added measure across PCPs. Specifically, let $\bar{x} = \sum_p S_p \mu_p$, where S_p denotes the share of patients who see PCP p, and μ_p is PCP p's (possibly normalized) VAM. Next, define a set of counterfactual patient shares S_p° that assigns zero weight to PCPs in the bottom 5 percent of the VAM distribution and rescales the remaining shares so that they continue to sum to one: patients of non–bottom-5-percent PCPs are reweighted proportionally by a factor of $1/0.95$. Using these counterfactual shares, construct the corresponding patient-weighted average VAM, $\bar{x}^\circ$. Finally, compute $\Delta = \bar{x}^\circ - \bar{x}$, which captures in percentage points how much the average PCP VAM would rise if all patients currently assigned to the bottom 5 percent of PCPs were redistributed proportionally across the remaining PCPs.

Panel A of Figure IV shows the distribution of changes in average PCP value-added under the system-wide counterfactual in which the bottom 5% of PCPs are removed from the market and their patients redistributed proportionally to remaining providers. This counterfactual leads the average beneficiary to see a PCP with a value-added measure 0.59 percentage points lower than in the baseline. Given our forecast coefficient of 0.71, this implies an expected reduction in one-year mortality of approximately [0.59 × 0.71 = 0.42] percentage points, or roughly [0.0042/0.079 = 5.3%] of baseline mortality. Of course, this simulation is unrealistic for a number of reasons. Among other issues, it may be difficult to implement for political economy reasons. (Policymakers are, in fact, loathe to sever patient-doctor relationships; see Congressional Budget Office (2010).) Yet there are other demand-side mechanisms that can be used to steer patients to low-cost or high-quality providers, including insurance plans.

Figure IV: Counterfactual Changes in Average PCP Quality (cont. on next page)



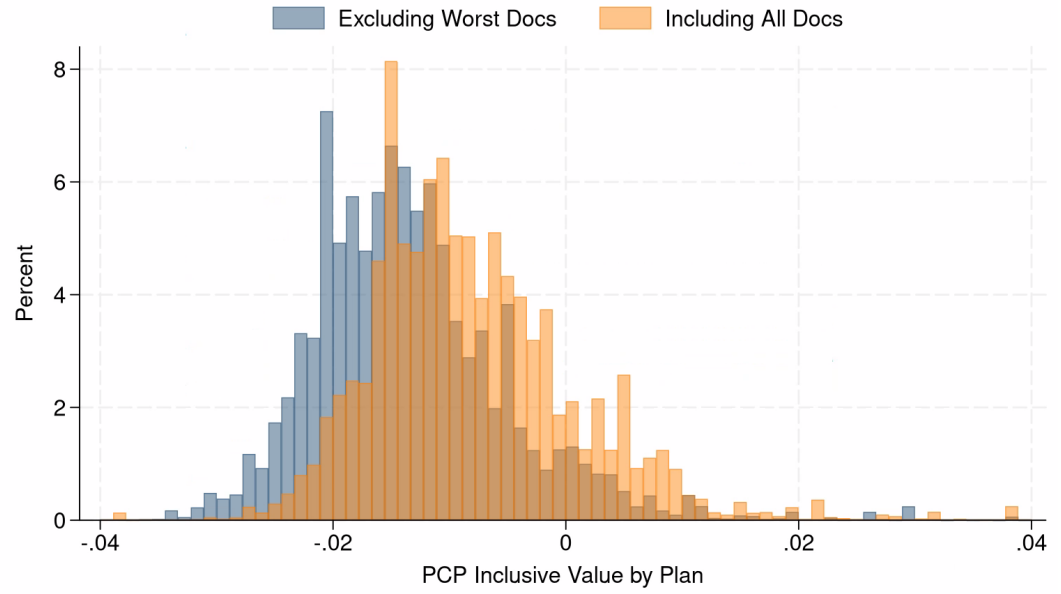
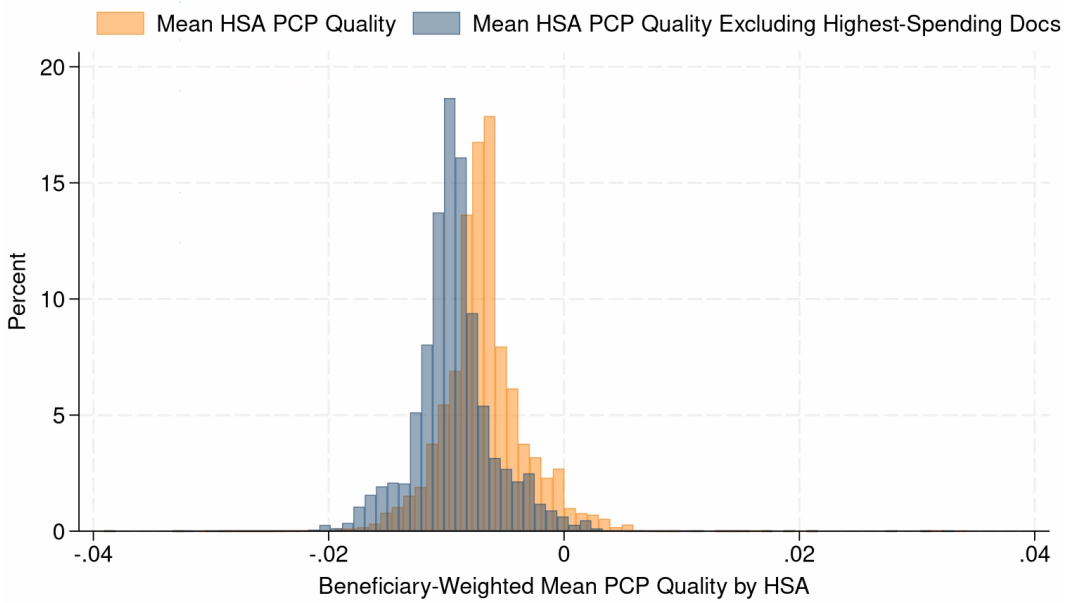


Figure IV: Counterfactual Changes in Average PCP Quality (continued)



*N*otes: Panel A plots the change in HSA-level average PCP value-added under a counterfactual in which the bottom 5% of PCPs (by observational mortality) in each plan are removed from the market and their patients are redistributed proportionally to the remaining PCPs in their HSA. Panel B plots the change in average PCP value-added under a counterfactual in which each insurance plan excludes the bottom 5% of its in-network PCPs and redistributes those patients proportionally among its remaining in-network providers.

5.3 Plan-Level Mortality Effects

This result raises an interesting question. Can insurers effectively steer their enrollees to quality providers? If so, will such steering affect plan quality in a meaningful way? We begin by replicating the previous simulation at the plan level. We relate our estimates to plan quality and demonstrate that plans can lead to considerable improvements in enrollee health by directing them to higher-quality providers. It underscores the importance of provider networks and consumer awareness in improving outcomes.[8]

Panel B of Figure IV presents an alternative, more feasible counterfactual in which each insurance plan excludes the bottom 5% of its in-network PCPs and redistributes those patients proportionally among its remaining in-network providers. Remarkably, this plan-level intervention achieves nearly identical gains: the average beneficiary would see a PCP with a value-added mea-

[8]Our approach connects to policy levers employed by both states and the federal government. Commonly known as "network adequacy" rules, regulations on networks can take many forms. Geographic access standards—which require plans to ensure members can reach providers within specified distance or travel times—are common. Requiring minimum provider-to-member ratios (1:2000 in MA) is also common. Plans must also typically ensure that a minimum percentage of PCPs are accepting new patients.

sure 0.55 percentage points lower than baseline. Put differently, plans can capture approximately 93% of the gains from the infeasible system-wide counterfactual simply by dropping low-quality doctors from their networks. Furthermore, as described above, there is no clear economic tradeoff: dropping these low-quality PCPs would also lead to lower overall utilization and costs.

Of course, insurers need not screen PCPs based on quality. True quality could be unobservable to insurers. Alternatively, focusing on quality might conflict with their goal of optimizing profits. (Indeed, WTP is basically unchanged in the previous counterfactual, indicating that such a change would not increase revenue significantly.) A natural approach – consistent with insurer incentives – would be to screen PCPs based on their cost to the insurer. We also run a counterfactual simulation in which each insurance plan excludes the top 5% of high-expenditure PCPs from its network. Again, patients are redistributed proportionally among the remaining in-network providers. Panel C of Figure IV shows that, under this counterfactual, the average beneficiary would see a PCP with a value-added measure 0.0026 percentage points lower than baseline. That is, screening on cost can lead to improvements in quality, though the magnitude is modest.

We also note that plans may not know the full distribution of μ_p. We next ask whether – based on the results in Section 4 – insurers can effectively screen on observable characteristics. We consider the following scenario: an insurer chooses a set of characteristics, such as medical school credentials and quality, on which to "screen." While Section 4 suggests that moving an individual to a PCP from a higher-ranked medical school may not causally change their mortality, we note that *plans* that direct patients to lower μ_p PCPs have better outcomes *on average*. To the extent that plans can predict which PCPs are higher quality on average, they can improve aggregate outcomes.

6 Conclusion

The paper presents develops and validates new, comprehensive measures of primary care physician quality in Medicare and uses them to investigate the determinants of physician and insurer quality. There is notable variation in mortality rates among PCPs, even for similar patients. Patients who switch from high-mortality to lower-mortality PCPs are subsequently less likely to die (and vice-versa for switches in the opposite direction). Higher-quality provider networks are also associated with lower patient mortality. A portion of the observed health plan quality comes from the value that included PCPs add. Simulations indicate that moving patients from the bottom 5% of PCPs to average-quality ones could reduce mortality significantly, even only reallocating patients within networks.

Our results can inform consumers, private insurers, and policymakers. They suggest that guiding patients to higher-value PCPs could save lives. The results also emphasize the importance of provider networks in determining the quality of insurance plans, which has significant implica-

tions for regulations surrounding "narrow networks," price transparency, and incentives aimed at promoting high-quality primary care.

Our work also opens several paths for future research. One area to use the value-added measures we constructed to analyze quantitatively the relative importance of factors like increased spending, care coordination, or practice patterns as determinants of physician quality. Research should also examine whether physician quality improves with competition, public reporting, or pay-for-performance incentives. Finally, a promising area involves simulating equilibrium effects: how would markets and patient outcomes change if networks improved, if patients could better identify high quality physicians, or if copays redirected patients across physicians?

References

Abaluck, J., M. C. Bravo, P. Hull, and A. Starc (2021). Mortality Effects and Choice Across Private Health Insurance Plans. *The Quarterly Journal of Economics*.

Badinski, I., A. Finkelstein, M. Gentzkow, and P. Hull (2023a). Geographic variation in health care utilization: The role of physicians. Technical Report Working Paper XXXX, National Bureau of Economic Research.

Badinski, I., A. Finkelstein, M. Gentzkow, and P. Hull (2023b). Geographic variation in healthcare utilization: The role of physicians.

Chandra, A., M. Dalton, and D. O. Staiger (2023). Are hospital quality indicators causal? Technical Report Working Paper YYYY, National Bureau of Economic Research.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review 104*(9), 2593–2632.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014b). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review 104*(9), 2633–2679.

Clemens, J. and J. D. Gottlieb (2017). In the shadow of a giant: Medicare's influence on private physician payments. *Journal of Political Economy 125*(1), 1–39.

Congressional Budget Office (2010, 3). Cost estimate for the patient protection and affordable care act (h.r. 3590) and the health care and education reconciliation act of 2010 (h.r. 4872). Cbo cost estimate, Congressional Budget Office, Washington, DC. Accessed December 18, 2025.

Doyle Jr, J. J., S. M. Ewer, and T. H. Wagner (2010). Returns to physician human capital: Evidence from patients randomized to physician teams. *Journal of health economics 29*(6), 866–882.

Doyle Jr, J. J., J. A. Graves, J. Gruber, and S. A. Kleiner (2015). Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy 123*(1), 170–214.

Fadlon, I. and J. Van Parys (2020). Primary care physician practice styles and patient care: Evidence from physician exits in medicare. *Journal of health economics 71*, 102304.

Finkelstein, A., M. Gentzkow, and H. Williams (2016). Sources of geographic variation in health care: Evidence from patient migration. *Quarterly Journal of Economics 131*(4), 1681–1726.

Finkelstein, A., M. Gentzkow, and H. L. Williams (2021). Place-based drivers of mortality: Evidence from migration. *American Economic Review*. Forthcoming.

Geweke, J., G. Gowrisankaran, and R. J. Town (2003). Bayesian inference for hospital quality in a selection model. *Econometrica 71*(4), 1215–1238.

Ginja, R., J. Riise, B. Willage, and A. Willén (2025). Does your doctor matter?

Gottlieb, J. D., M. Polyakova, K. Rinz, H. Shiplett, and V. Udalova (2023). Who values human capitalists' human capital? the earnings and labor supply of us physicians.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review 30*(3), 466–479.

Hull, P. (2020). Estimating hospital quality with quasi-experimental data. *Working Paper*.

Kane, T. J. and D. O. Staiger (2008). Estimating teacher impacts on student achievement: An experimental evaluation. NBER Working Paper No. 14607.

Kwok, J. H. (2019). How do primary care physicians influence healthcare. *Evidence on practice styles and switching costs from medicare. URL: https://static1. squarespace. com/static/5bd6632951f4d49caf7eecd5 5*, 1564679228053.

Molitor, D. (2018). The evolution of physician practice styles: evidence from cardiologist migration. *American Economic Journal: Economic Policy 10*(1), 326–356.

Morris, C. N. (1983). Parametric empirical bayes inference: theory and applications. *Journal of the American statistical Association 78*(381), 47–55.

Sabety, A. (2020). The value of relationship-specific capital in health care. Working Paper.

Sabety, A. (2023). The value of relationships in healthcare. *Journal of Public Economics 225*, 104927.

Tilipman, N. (2022). Employer incentives and distortions in health insurance design: Implications for welfare and costs. *American Economic Review 112*(3), 998–1037.

A Data Appendix

B Additional Results

C Comparison with Hospital Quality

We also use existing, validated hospital quality measures. We begin with the observational value-added (VA) measures in Chandra et al. (2023) (henceforth "CDS"). They estimate short-run mortality outcomes for Medicare patients referred to hospitals by ambulance for emergency conditions. CDS validate their measures of hospital quality (at the year level) using a research design based on hospital closures. Our main measure is 30-day survival, incorporating risk-adjusters such as demographics and comorbidities. The estimation procedure involves averaging condition-specific RAM quality indices across patient conditions.

The CDS approach proceeds in a series of four steps. First, they estimate an initial hierarchical model of mortality as a function of hospital fixed effects and individual risk-adjusters. Construct the residual. Second, regress the residuals on the logarithm of hospital volume. Construct patient-level residuals and calculate the average residual for each hospital-year combination. Third, estimate the variance of the hospital quality residuals and the correlation of these residuals over time. Finally, form the best linear predictions of hospital quality using only data from other years.

D Notes

Table V: PCP Characteristics and Observational Mortality

	OLS		IV		
Provider Characteristic	Coefficient	(SE)	Coefficient	(SE)	F-stat
Demographics					
Female	-0.00209	(0.00011)	0.07441	(0.01323)	173.147
Age	-0.00010	(0.00001)	-0.00024	(0.00006)	18000
South Asian	0.00736	(0.00037)	-0.01946	(0.03249)	26.170
Southeast Asian	0.00221	(0.00091)	0.01597	(0.05533)	16.376
West Asian	0.00656	(0.00044)	-0.00254	(0.05063)	13.314
East Asian	-0.00007	(0.00034)	-0.03043	(0.03978)	13.710
White	-0.00307	(0.00015)	0.00199	(0.01058)	214.148
Black	0.00058	(0.00054)	0.07688	(0.04341)	25.377
Hispanic	-0.00064	(0.00027)	0.03650	(0.02886)	27.306
Training & Credentials					
MD (vs. DO)	0.00148	(0.00012)	-0.00192	(0.00623)	645.886
Top-25 med school	-0.00219	(0.00018)	-0.00242	(0.01221)	226.720
Private med school	-0.00161	(0.00014)	-0.00221	(0.00854)	352.097
Top-10 research med school	-0.00072	(0.00024)	0.00646	(0.02806)	42.977
Top 10 primary care med school	-0.00065	(0.00024)	-0.03471	(0.02762)	60.548
Med school has primary care rank	-0.00058	(0.00014)	0.01826	(0.01043)	214.482
Med school Avg. MCAT	-0.00078	(0.00014)	0.02984	(0.01822)	24.559
Med school Avg. student-to-faculty ratio	-0.00052	(0.00008)	0.01298	(0.00850)	27.222
Top-25 residency	-0.00095	(0.00035)	0.03758	(0.09723)	2.236
Multiple residency	-0.00243	(0.00067)	0.02124	(0.05053)	18.739
Had fellowship	-0.00030	(0.00049)	0.01405	(0.01556)	187.167
Other Characteristics					
Republican	-0.00077	(0.00014)	-0.01374	(0.00910)	287.048
Democrat	-0.00030	(0.00014)	0.01515	(0.00942)	271.898
Registered voter	-0.00055	(0.00015)	-0.01663	(0.00815)	381.196
High wealth	0.00122	(0.00020)	0.00252	(0.01516)	107.133
High income	0.00087	(0.00016)	0.00059	(0.01390)	125.500
Has spouse	-0.00051	(0.00016)	-0.03093	(0.00906)	353.282
Spouse is doctor	-0.00097	(0.00032)	0.02338	(0.05230)	6.458
Avg. online reputation	-0.00156	(0.00014)	-0.01206	(0.01528)	56.910
Has online review	-0.00250	(0.00012)	-0.00981	(0.00857)	517.807
Has dox profile	-0.00198	(0.00012)	-0.01433	(0.00725)	637.336
UCSF Applicant Characteristics					
MCAT score	-0.11	(0.06)			
MCAT score, conditional	-0.09	(0.06)			
MCAT score, within school	-0.14	(0.07)			
College GPA	-1.88	(0.80)			
College GPA, conditional	-1.46	(0.83)			
College GPA, within school	-1.80	(1.17)			
College STEM GPA	-1.42	(0.63)			
UCSF screen score	-0.16	(0.07)			
UCSF screen score, within school	-0.23	(0.12)			
UC Undergraduates Characteristics					
GPA	-0.63	(0.29)			
Organic Chemistry grade	0.29	(0.23)			
Biology GPA	0.02	(0.25)			
Chemistry GPA	0.42	(0.20)			
Physics GPA	0.06	(0.22)			
Math GPA	0.12	(0.21)			
Humanities Double Major	-0.78	(0.52)			
UC Medical Students Characteristics					
Medical GPA	0.18	(0.53)			
Medical class rank	-0.008	(0.007)			
Medical class rank, UCSF	-0.03	(0.01)			

*Notes:*This table reports OLS and IV estimates of the regression of observational mortality and physician mortality effects, respectively, on physician characteristics. The dependent variable is observational mortality in Panel A and one-year mortality in Panel B. The IV specifications instrument by the interaction of lagged physician characteristics and terminations, controlling for main effects. The average non-missing value across plans replaces missing physician characteristics. Standard errors are clustered by county and reported in parentheses.

Notation	Definition
Y_{it}	1(bene i dies in year t)
μ_p	PCP p VAM measure
D_{ipt}	1(bene i sees PCP p in year t)
X_{it}	bene i year t covars
ε_{it}	bene i year t health shock
β_p	causal effect of PCP p
η_p	forecast error term
Z_{it}	$\mu_{i,t-1} \times T_{i,t-1}$ instrument, where $\mu_{i,t-1} = \sum_p \mu_p D_{it-1}$
W_{pt}	wages of PCP p in year t
τ_{pt}	(marginal) tax rate of PCP p in year t
τ_{pt}^{sim}	simulated version of τ