# Health Consequences of Persistence in Physicians' Locations: Evidence from Decentralised Loan Repayment Programs

**Initial draft: October 2021**

Please click here for the latest version

**Anomita Ghosh**[‡] [1]


### Abstract

Do temporary labor supply programs cause physicians to move to and stay in undesirable areas? To what extent do these programs improve the health of the elderly and non-elderly population in those areas? I investigate these questions by studying state and local loan repayment programs for new eligible physicians which were rolled out over the last four decades in hundreds of counties across US states. Leveraging a new longitudinal dataset that tracks all physicians from medical school to mid-career, and exploiting both space and time variation, I find that these policies increase the number of physicians by 5% in treated counties relative to untreated counties in the state. The inflows of physicians are driven by higher paying eligible specialities. The programs continue to influence physicians' location decisions even after they end –- effects persist for at least ten years after the minimum obligation period. This is driven by employer learning effects. Furthermore, the programs modestly spur trainees to enter eligible specialities in treated states by substituting away from ineligible specialities. Treated counties also see the elderly and non-elderly increase their visits to physicians while reducing those to the emergency rooms. Using patient level data from California, I demonstrate these results are not driven by selective admission of patients to treated hospitals. My findings underline the relevance of policies that decrease financial frictions for highly skilled individuals in changing community health outcomes.


**Keywords:** labor supply, state and local government, migration flows, occupational choice, adverse health events, health expenditures

[1‡] I am grateful to David Albouy, David Molitor, Russell Weinstein, and Ben Marx for their indispensable support and guidance. I also thank Dan Bernhardt, Andrew Garin, Mateo Arbelaez, Mauricio Olivares-Gonzalez, Alex Bartik, Richard Akresh, Mark Borgschulte, Heejin Kim, Marieke Kleemans, Day Manoli, Abhiroop Mukhopadhyay, Matt Notowidigdo, Brendan Price, Danny Tannenbaum, Rebecca Thornton, Owen Zidar, and faculty at Norwegian Business School, Indian Institute of Technology Kanpur, Indian School of Business, Indian Statistical Institute Delhi, Azim Premji University, University of Purdue Agricultural Economics, University of Connecticut Agricultural Economics, Bilkent University, Orebro University School of Business, UC Irvine, and various conference participants for helpful discussion and suggestions. Any errors or omissions are my own.

**JEL Classifications:** H75, J24, J32, J61, I18

## 1 Introduction

There is a large disparity in how primary care physicians are distributed between the urban and rural areas across the United States. While the largest metro areas have over 11 physicians per 10,000 population, completely rural areas have less than 4. Looking across all geographies, the distribution of physicians varies considerably: there are 4.3 physicians per 10,000 people in places in the 10th percentile of physicians per capita, and 15.03 in places in the 90th percentile. [2] These disparities can be quite large within states. For example, there are 13.2 and 2.0 physicians per 10,000 people in the largest metro areas and completely rural areas, respectively, in Illinois.

This large geographic disparity impacts access and costs related to preventive care, inpatient admissions, emergency department (ED) visits, and other health-related events.[3] To address these geographic imbalances, public policy initiatives have pursued both demand-side incentives such as geographic variation in reimbursement rates and supply-side incentives like lumpsum payments to attract physicians to underserved areas. While the current literature has focused on demand-side incentives, there is much less evidence on the causal impact of supply-side incentives. Accordingly, in this paper, I ask whether temporary supply-side incentives are effective in moving physicians, changing health outcomes, and have lasting effects beyond the expiration of the incentives.

My analysis addresses this question through a series of state and local PCP loan repayment assistance programs. These programs were rolled out in a staggered manner over the last four decades, from 1978-2015, in hundreds of counties across forty-nine US states. They provide lumpsum funding to newly-trained eligible physicians who agree to practice for the contract period at a pre-approved site. These funds should be used to pay off the medical education loans of the eligible physicians. The advantage of this setting is that there are multiple sources of variation embedded in the design of these programs –- which allow a researcher to

[2] The 'largest metro areas' and 'completely rural area' classifications follow the Rural-Urban Continuum codes from the US Department of Agriculture.

[3] In 2008, there were more ED visits per 100,000 population in rural areas of the Northeast (50,115) compared to the corresponding non-rural areas (44,464). This was also the case in Southern and Western regions (Healthcare Cost and Utilisation Project [HCUP], 2008). In general, the average cost per ED visit was higher in rural areas at \$560, while in metro areas it ranged between \$500-\$540. In particular, Medicare incurred a total cost of \$2.051 billion on ED visits in rural areas (HCUP, 2017).

isolate the effects of supply side incentives on physicians' movement, their speciality choice, and the patients' health outcomes.These variations are not limited to the space and time of implementation of these programs. The other differences include, for example, the amount of funding offered by the states and localities over the contract period. Moreover, the minimum service period, and the set of eligible specialities also vary across the states.

Using within-state variation across counties and time variation, I first estimate the overall effects of the programs on the entry and exit of eligible physicians with event study and difference-in-differences designs. The availability of a rich individual longitudinal dataset –- one that spans the majority of the policy period –- allows me to causally estimate the migration flows of physicians in response to these policies. This restricted access dataset, covering the universe of physicians, tracks each of them from their medical school to mid-career and contains detailed education, training, speciality, demographic and workplace information.

Then, leveraging the state-year variation, I evaluate the composition effects by analysing whether the policies spur interest among first-year training physicians to join the eligible specialities. This analysis is facilitated by a new, restricted access, state-year level dataset which contains detailed speciality information on the near-universe of training physicians by their program-year. [4]

Finally, relying on both county-level Medicare data and confidential patient-level discharge data from California, I provide evidence on the significant benefits of the policy for both the elderly and non-elderly adult population. This includes more visits to new physicians and lower rates of adverse events like hospitalisation and emergency department visits.

There are two major empirical issues that researchers face in the causal estimation of physicians' location choice. First, physicians choose when and where to move, i.e, their location choice is non-random. Their migration flows are likely to be correlated with preferences and local economic conditions, including wages and amenities. To deal with this issue, I exploit the time-space variation offered by a set of policies which addresses both when and where they relocate. The absence of pre-existing trends in the event studies suggests that the policies seem largely uncorrelated with pre-treatment economic conditions in treated counties, thereby strengthening the causality of my estimates.

[4] Failure to account for such composition effects may yield estimates that overstate the direct treatment effects of the programs on practicing physicians' inflows to treated areas.

The second and more substantial identification issue is the possibility of these programs being correlated with demand-side policies or other unobserved shocks. This can be a confounding factor – if the other policies or unobserved shocks occur in the same county and year as the implementation of these programs, are applicable to the same group of specialities over the same duration – but are not captured by the extensive set of county level covariates and county and state-by-year fixed effects I include in the model. I first show through placebo tests that physicians in the treated counties do not experience changes in reimbursement rates, which are labor demand shocks.[5] [6] Additionally, I show a flat and insignificant entry of experienced physicians and ineligible speciality doctors in the treated counties through placebo tests. Federal physicians are exempt from the policies and they constitute a valid counterfactual group with no access to the treated communities. Combined, these falsification exercises reinforce the conjecture that other confounders are less likely to affect the causality of my estimates.

I begin my causal analysis by establishing the take-up rate of these policies among the eligible physicians. To this end, I estimate that the programs have sizable positive effects on move-in of new practicing physicians in the treated counties. After documenting flat and insignificant pre-trends, my estimates suggest that the policies increase the number of MDs and DOs by 4.9% and 7.83% respectively, in treated counties relative to untreated counties, within a treated state. The policies, despite being transitory, seem to steadily draw MDs to the treated areas. To address the threat of physicians sorting into specialities because of various reasons as described in the paper, I estimate the hiring rates of physicians within a speciality and location, and find qualitatively similar results.

To shed light on the mechanisms driving these main effects, I show that the inflows of new physicians are largely driven by higher-paying eligible specialities.In other words, the benefits of the policies were likely insufficient to make lower-income general family and internal medicine physicians move to a treated county in the short run i.e. within five years

[5]It is worth noting that program regulations limit a particular eligible physician to receive funding from only one program at a time.

[6]A physician's total compensation generally consists of salary, personal productivity, practice financial performance, bonus and other sources, where the contribution of each of these components varies by physicians' ownership status (Rama, 2018). Any productivity based component is usually considered a labor demand shock.

of adoption of the policy. Instead of moving to the remote rural treated counties, these new hires exhibit a strong preference to cluster closer to the large cities within the state. A possible explanation can be that proximity to big cities offers physicians the advantages of easier access to urban consumption amenities and commuting facilities, more opportunities for networking with other doctors, and more employment opportunities for their spouses. This finding is also corroborated by physicians' proclivity for treated counties which are richer in non-natural amenities. Importantly, using individual records, I find that physicians' preferences for proximity to big nearby cities are largely driven by the causal effects of the policy, rather than their preferences to practice closer to their medical school - majority of which are located in the big cities.

The above set of heterogeneity results suggest that the effects of these policies are unequal across the counties. There is some suggestive evidence that these policies alone may be insufficient to move physicians to the underserved communities. Instead, the incentives of these policies may act as complements to 'wage', 'neighborhood-based', and 'workplace' amenities –- for the marginal physicians induced to move by these policies.

I further investigate the characteristics of the newly hired physicians who are screened-in by the policies. First, there seems to be a persistent rise across all cohorts in the fraction of physicians who choose to practice in their state of training after implementation of the programs. This finding can be explained as follows: the financial incentives associated with these policies induce the eligible physicians to enter treated counties closer to the state's large cities, thereby increasing the attractiveness of the job, and these benefits outweigh the opportunity costs of moving out-of-state and searching for a new job or forming new professional connections.

Second, I show that while the policies encourage the enrollment of high-debt foreign physicians as well as physicians from lower ranked US medical schools, it deters relatively higher-debt physicians from unranked US medical schools.The design of the policy specifies that the new physicians will not be provided their entire loan amount if the benefits provided by the state of their practice is less than their accumulated debt, likely contributing to this deterrence effect. A natural question arises as to whether there is any adverse selection of relatively lower debt physicians due to these policies –- given the large benefit amounts offered in some states. However, I fail to find such evidence in the data. One important reason may be that physicians are unlikely to move to undesirable areas by falsifying their debt amounts — which are verifiable — just to receive the benefits of the policy.

Having characterised the inflows of new physicians and unpacking the possible channels, I examine their likelihood of staying in the treated county after exhausting their benefits. Leveraging the new individual longitudinal dataset, my estimates imply an economically small decline in the likelihood of staying in a treated county one year after the contract period of that state. I observe similar small outflows, both five and ten years after the service period. However, the net inflows of these new physicians remain significantly positive, and continue to grow over time. This suggests the stickiness in their location choices, even after the obligation period ends. I find some preliminary evidence that these retention patterns are driven by married physicians, or more precisely working couples.

While I observe that the policies lead to an increase in the physicians in treated areas relative to the untreated ones, this may simply be a reallocation of physicians within the state, with no aggregate welfare effects. However, I document a persistent rise of 2.2% in the number of first-year matched residents in treated specialities timed with the policy, which is achieved by a more than proportionate decline in the trainees in untreated specialities. Having ruled out the possibility of programs expanding to absorb the higher number of matched residents, I attribute this rise to an implicit interest of trainees in the treated specialities.[7] These results on entry into specialities suggests that it may be useful to consider how the initial supply and composition of the applicant pool of physicians change — in addition to the — usual redistribution of physicians across space and time, to conclude about the welfare effects of such programs.

After focusing on physicians' migration flows and speciality choice, I then analyse whether the programs achieve their most crucial intended effect –- in terms of the benefits to the elderly population. While theoretically it is unclear whether patients' demand for a new inexperienced physician will increase when they relocate to a treatment area, I find clear evidence of a modest but significant 2.1% increase in per capita Medicare enrollees having at least one ambulatory visit to physician in the treated counties. These effects are persistent over time, suggesting continued willingness of the elderly patients to form new physician relationships.

[7]Note that, the increase in aggregate trainees is distributed among general obstetrics/gynaecology, general pediatrics and eligible family/internal medicine specialities, while general family and internal medicine trainees, comprising the largest category, have seen declining interest over the post-policy years.

Furthermore, the policy causes a 4.98% reduction in per capita emergency room (ER) visits and a 10% decline in preventable hospital stays among Medicare beneficiaries. Consistent with declining ER visits and hospital admissions, my estimates also point to an economically meaningful reduction in hospital reimbursements per Medicare enrollee in the treated counties after the policies.

I supplement the above county-level analysis using richer patient- level data from California, whereby I provide comparable evidence of decline in new admissions of Medicare patients in treated hospitals. The reason I perform this additional analysis is to address the fact –- the average treatment effects within counties may be confounded by inherent differences across hospitals, and patients of varying severity in illness selecting into them.

This paper makes four main contributions to the literature. To my knowledge, it is the first to use both time and granular space variation provided by a set of temporary, income-based labor supply policies, to isolate their causal effects on the dynamics of physicians' movement over their careers in the US. [8] I believe that the program I study is a more direct way of making physicians move relative to labor demand or health insurance based policy changes that are examined in the current literature (Costa et al, 2019; Falcettoni, 2020; Huh, 2019; Khoury et al, 2021 ; Kulka, 2019).[9] [10] Equipped with a long panel of rich individual data, a group of policies having multiple sources of variation, and precise identification of treated counties, I am able to find relatively large take-up rates and effectiveness of these policies, as compared to majority of the above studies.

The paper which is most closely related to mine is Falcettoni (2020). The author uses labor demand shocks, in particular Medicare fee for service reimbursement rates, to identify causal effects on Medicare doctors' location decisions across the US. Furthermore, the time period

[8] Carrillo and Feres (2019) –- using only spatial variation –- analyse the effects of More Physicians Program in a developing country, i.e. Brazil, on physician supply and utilisation of medical care. They do not consider the dynamics of physicians' mobility decisions.

[9] The negative urban wage premium for doctors implies that they strongly prefer to live in large cities (Lee, 2010). This suggests that supply-side incentive is a more direct approach to relocate doctors.

[10] Papers that study physicians' location choices in different countries, largely focusing on labor demand incentives, include Bolduc et al (1996) in Canada. Holmes (2005) explores physicians' static location decisions in the US, with the location choice set being endogenously determined –- leading to downward biased estimates. Importantly, these papers do not have enough variations to disentangle the effects of a supply side policy, nor can they examine the dynamics of physicians' movement over their careers.

of their analysis is 2012-2016, which does not cover the pre-policy period for most states.[11] There are several ways in which this paper differs from theirs. *First*, I use a labor supply shift as the identifying source of variation- specifically a combination of within-state variation across counties and year variation provided by a policy experiment. *Second*, I examine an important goal of the policy that aims to lower disparities in health access, outcomes and expenditures in treated areas, using both county level data and rich patient level data to address selection issues. *Third*, the rich individual longitudinal dataset I employ allows me to study the retention choices of the universe of physicians in treated counties upto ten years after the obligation period –- which varies by state.

This paper complements the growing *experimental* literature that examines whether financial incentives can attract frontline providers to disadvantaged regions in developing countries ( Kruk et al, 2010; Rao et al, 2012; Dal Bo et al, 2013; Ashraf et al, 2020; Deserranno, 2019) by providing causal evidence on the overall treatment effect and composition bias arising from a set of decentralised policies in a developed country. Identifying the stages of composition changes over a physician's career –- due to government intervention –- can help policymakers in designing policies that meet their intended objectives. Additionally, my setting offers an opportunity to study the long-term effects in physicians' location after the program ends, and the corresponding health benefits, distinguishing it from the experimental literature.

*Second*, this paper also speaks to the literature suggesting that trainees' choice of specialisation responds to the expected income of a speciality (Hay, 1991; Hurley, 1991; Nicholson, 2002) by incorporating the effect of public policy on trainees' propensity to join an eligible speciality. This potential contribution of public policy has been largely unstudied in the literature.[12]

*Third*, this paper also clarifies a previously unexplored aspect of how the exit of primary care providers affects the health of Medicare patients (Agha, Frandsen and Rebitzer, 2017;

[11]A possible empirical challenge of these Medicare reimbursement shocks is that they may violate the exclusion restriction. Some studies have identified Medicare reimbursement rates as important predictors of hospital closures.

[12]A notable exception is Wasserman(2019), who uses a natural experiment to show how changes in non-monetary incentives like time requirements in certain specialities affect trainees' likelihood of joining those specialities. Given the differences in our setting, my paper offers additional insights on both the speciality and location choices of physicians, beyond their training stage.

David and Kim, 2018; Fadlon et al, 2020; Sabety 2021) i.e. how the entry of new physicians due to the policies affects patients' allocation of care between physician settings and ERs/inpatient hospitals. Contrary to these papers, my results indicate that the policies cause patients to meaningfully substitute their source of care from ERs to physician settings.

*Finally*, I contribute to our understanding of the nascent but growing literature on doctor migration that isolates doctor-specific factors from environment-specific factors (for example,Molitor(2018)), by providing a set of policy experiments that can plausibly explain why physicians' practice environments might change at the start of their careers.

The rest of the paper is organised as follows. In Section 2, I briefly describe the institutional details the programs. In Section 3, I describe the data sources and sample construction. In Section 4, I discuss the identification strategy. In Section 5, I present the main results, and also show evidence of mechanisms explaining these results. In Section 6, I discuss the robustness of my results and rule out alternative interpretations. Finally, I conclude in section 7 with a policy-induced estimate of marginal value of public funds(MVPF).

# 2 Background

## 2.1 Institutional details of the policy

The state and local PCP loan repayment assistance programs provide a set of incentives and guidelines –- varying by state and locality –- to hire physicians in the underserved areas.These programs may be considered as providing information to first-time job seekers about places lacking physicians, where job searches may be more successful. Given that states and local communities know what is best for themselves and therefore design policies accordingly,there are multiple sources of variation embedded in the design of these programs which a researcher can exploit in their analysis. These policies offer a certain amount of funding to newly trained eligible physicians who agree to practice full time for the duration of the minimum service period in a *pre-approved site*.[13]The funding should be used to pay for medical education student loans incurred by these eligible physicians. Importantly, these funds cannot be considered as sign-in bonuses or relocation allowances –- whose expenditures

[13]Additionally, some states like Oregon, Colorado and New Hampshire also provide lower funding to eligible physicians who practice part time in a treated area.

are discretionary –- and not tied to a specific purpose. In general, the loans that are not eligible for repayment under this program comprise those that were consolidated with any other type of debt or another person's debt, Parent PLUS loans and loans from a friend or family member. Funding is a 1:1 match, varying by state and is jointly provided by the state government and local community or hiring organisation.[14] The funding is generally disbursed according to the formula: $min\{Benefit\ amount, Loan\ amount\}$ where benefit amount is the state specific funding offered by the policy. This formula implies that those physicians with large loan amounts exceeding the state mandated benefit amount will not be provided the full amount of their loan. Similarly, if the loan amount is smaller than the benefit amount, the participants are not paid the full amount of the benefit offered by the program. Note that, this design may prevent those physicians with significantly high debts relative to the state specific funding, from enrolling in that state's program.

Figure 1 presents the county level distribution of the amount of funding per physician for the entire duration of the minimum service period. [15] While most counties have a funding of $100,000 or less, eligible counties in Mississippi, Nebraska, New York, South Dakota and Texas have generous benefits of more than $120,000. With most states offering equal amount of funding each year of the minimum service period; states like Arizona, Oklahoma and Texas provide different amounts over the duration of the obligation period. Despite the uncertainty in the amount of budget allocated by state and local governments for these policies –- which affects the number of hires in a year –- there are no substantial fluctuations in the state prescribed benefits per eligible physician over the years.[16]

The minimum service period also differs by state, ranging between 1 and 5 years (Figure 2). All new physicians are required to stay for the fixed term as prescribed by the state, and this term length has remained unchanged throughout my sample period. In other words, there is no variation in contract periods for physicians within a state –- unlike the randomized term length assigned to incoming state bureaucrats in Arkansas, Florida, Illinois and Texas. In

[14]For example, in California, the funding comes from licensure fees from the Medical Board of California and Osteopathic Board of California, Managed Care fines and penalties, and other donations.

[15]I have used the terms 'funding' and 'benefits' as synonymous in the paper. Both indicate the state level benefits, as specified in the law.

[16]The funds allocated for these programs vary by state. For instance,in New York, the total funding earmarked has been as high as 9 million for an application cycle.

general, moving away from the treated area before the minimum contract period involves payment of penalties - and is therefore rarely observed.[17] In some states, recipients can apply for one or two years of renewed funding after completion of their minimum service period.[18]

Other important stipulations of the program include restrictions on medical school and citizenship. While Alabama, Indiana and Maine have in-state medical school restrictions to be eligible for the policy, other states do not specify such restrictions. Moreover, this policy is not only targeted to US citizens although states like Iowa, New Mexico, Utah, Vermont and Hawaii specifically state US citizenship as criteria for receiving the benefits. Equally important is the set of eligible specialities among physicians –- which differs across the states and is described in section 3.2.

Finally, I compile information on the rollout dates of the program either from the state health department websites or from the state-specific original legislation documents (Figure 2).[19]

After outlining the major elements of the program design, it becomes imperative to mention how they are communicated to the medical students and trainees –- which in turn –- determine their willingness to participate in these policies. This information is available on the websites of the medical institutions and state health/education departments, as well as guidance is provided by the financial aid offices and program directors of these institutions. Hence, it is likely that the students are aware of these policies–- though the information can be provided to them in a more systematic way, through nudges. Their ultimate take-up decisions, then, depend on the ease of the application and approval processes, and the fairness

[17]These penalties can sometimes be as large as 200% of the principal amount of the loan. It can also be the sum of : amount paid to the participant for any un-served period, the number of months not served multiplied by $7500 and the interest on the first two components. It can also cause cancellation of their license to practice medicine. Sometimes, if a health care facility applied for the award for its incoming physician and the physician leaves the job before the contract period, the facility is responsible for the refund and penalty.

[18]These states include Louisiana, New Mexico, Ohio, West Virginia, Alaska, Delaware, Illinois, Georgia, Maryland, Michigan, Oklahoma, Virginia.

[19]https://www.ruralhealthinfo.org/funding/states provides a list of funding opportunities offered in rural areas of each state. From this list, I choose those specific programs which are applicable to primary care physicians. California is an exception where both physicians and surgeons are eligible for the benefits. With the above restriction in mind, the year of rollout of these programs I have considered in my paper, may differ from the year the program was first introduced in that state –- for a few states.

of the provisions in the appointment contract–- which vary by state and employer.[20]

In general, most states have pre-defined criteria for screening the applicants. This restricts the scope of discretionary hiring practices by recruiters within a state, as compared to the private sector. Despite these criteria, of course, recruiters may be mistaken or exhibit bias in their hiring decisions –- which may influence the gaps in the hiring of physicians across space and time. Or, they may also have private signals about the applicants' quality that are used in the hiring process. These include- generally hard to observe soft skills - like public service motivation. Additionally, organised monitoring of the local recruiters is absent in many states –- which may allow for weak execution of the set criteria.

## 2.2 Speciality choice

The choice of speciality determines the career path and income of practicing doctors. This decision is critical –- since it entails significant human capital inputs throughout the medical school, a demanding internship period, and hefty switching costs.[21]

Medical students choose their preferred speciality when they apply for residency programs in the beginning of fourth year of their medical school. After that, the residency programs select which applicants to interview. When the interviews are complete, most of the first year residency slots are assigned on a single day through the National Resident Matching Program (NRMP).

Literature suggests that speciality choice responds to changes in monetary payoffs of specialities (Nicholson (2002), Gagné and Léger (2005)). [22]However, current research has not addressed in a causal framework whether public policy that temporarily increases income of

[20]Conversations with some physicians suggest that the contract provisions for the eligible new hires are not always clear and transparent –- which may limit the large scale take-up of these programs. The Department of Education has taken a positive step in this direction by recently announcing efforts to improve the application and approval processes of the programs. See https://www.ed.gov/news/press-releases/fact-sheet-public-service-loan-forgiveness-pslf-program-overhaul.

[21]Because of the high switching costs, doctors are unable to train in a different specialty once they have completed their training in one. For example, currently a practicing general family physician's average annual salary of $261,000 is more than 4x a first year training physician's salary of $\sim$ $ 58,000 (Freida database and Doximity report).

[22]Other factors affecting the choice of specialities include physician's interest, ability, lifestyle considerations, prestige and expected remuneration, availability of residency slots and perceived job availability (USDHHS, 2008).

certain practicing physicians affects trainees' decisions to join those specialities. Theoretically, the effect of increase in practicing doctors' income on trainees' speciality choice can be ambiguous in direction. This is because choice of a preferred speciality depends not only on expected lifetime earnings conditional on entering that speciality (which increases for treated specialities because of the policy) but also on probability of entering that speciality and expected lifetime earnings associated with other specialities. Amongst other reasons, this uncertainty of entering a speciality is because of the existence of residency caps established by the federal government.

To appropriately estimate the Average Treatment Effect on the Treated (ATT) coefficients of the policies I analyse, it is essential to address the issue of selection into specialities. The second motivation for researching trainees' specialisation in my setting, is to determine the intertemporal general equilibrium implications of these programs. In other words, the welfare impacts would be significantly different, if the policies resulted not just in a simple reallocation of physicians between treated and untreated areas, but also in the entry of newly minted physicians into the eligible specialties.

# 3 Data

To estimate physician responses to the policy, I explore two main longitudinal datasets- one at the individual level to study decisions made by practicing physicians and one at the state level to study decisions made by training physicians. To estimate health effects of the population due to the policy, I utilize three administrative longitudinal datasets-one at the county level for health outcomes of Medicare population, one at the county level for mortality outcomes and the final one at individual level for hospital entry, utilization and charge outcomes of patients in California. Below, I describe the details of each dataset used in this paper.

## 3.1 Data sources

### 3.1.1 Primary care physician (physician) data

To examine how primary care physician entry and exit responds to temporary labor supply policies, I first start with count data of MDs (Doctor of Allopathic Medicine) and DOs(Doctor of Osteopathic Medicine) at a county-year level. I obtain this data from Area Health Resource

File and American Medical Association Physician Masterfile for the years 1995-2017.[23]The count data provides aggregated totals of active physicians by federal/non-federal status, speciality, major professional activity, county and year. However, this dataset is not able to distinguish between existing physicians and new physicians who enter the labor market. Nor can it track the migration decisions of the individual physicians over the span of their labor market careers. Finally, this data does not enable me to compare the outcomes of physicians within specialities, and therefore cannot address the fact that physicians may select into certain specialities based on observable and unobservable characteristics.

In view of the above limitations, my analysis also crucially relies on a new longitudinal dataset of individual physicians containing their demographic, education and employment location details. I build this dataset from multiple sources in the following manner. First, I extract the education and training details of all primary care eligible physicians (900,000+ of them) from Doximity, LinkedIn and workplace based websites. This data includes medical school attended, year of graduation from the medical school, training institution, start and end dates of training, fellowship institution, start and end dates of fellowship. From this information, I can determine the year they start practicing. I select the subset of primary physicians who choose their first job between 1996-2017, as they are likely to be affected by the roll-out of the policy. For each of these physicians, I obtain deidentified data from American Medical Association which allows me to track each physician's migration choice,starting from their medical school till the end of 2017. [24]In particular, this rich micro-data allows me to observe the following relevant variables for my analysis-workplace location to the level of zip code, primary and secondary speciality, type of practice, employment arrangement, year of state license, demographics like age, sex, birth date and birth place. I link the deidentified data from AMA with the education data I scraped from the websites and reidentify the whole dataset with the names of the physicians. I use a sufficient set of matching attributes –- which vary by physicians –- to merge the two datasets. This process ensures that all records are matched and there are no duplicate matches. One reason why I need to link these two

[23]The AHRF has MD data for the years 1995-2008, 2010-2017. The DO data, on the other hand, has limited availability for the years 1995,2003, 2004, 2007 and 2010-2017. I procure the data for the missing years from American Medical Association.

[24]Note that, the deidentified data from AMA does not provide the physician's name and exact street address of office.

datasets is that some medical training information is lacking in the AMA data I have, which is a crucial piece of information in my analysis. [25] To this dataset, I add the state of medical school of each practicing physician from online sources.[26] I source the information on rank of US medical school and average US medical school debt of each practicing physician from US News and World Report primary care 2020. For unranked medical schools and foreign medical schools, I obtain the average medical school debt from their respective institution websites. [27] I further add the state of training institutions of each practicing physician from National Residency Matching Program (NRMP) reports.

The advantage of AMA Masterfile data is that it covers the universe of physicians in US. The physician is included in the Masterfile when they enter US medical school or in case of International Medical Graduates, a US residency program. All education and employment details are verified from the primary source. Each physician is tracked till death, including periods when they are inactive. This ensures that I correctly identify the physicians in the treatment and counterfactual groups –- taking into account –- their periods of inactivity. [28]

For training physicians, I mainly use the matched resident count data collapsed to state-academic year-speciality-program year level. I prepare this dataset by combining data from Graduate Medical Education (GME) Track, Association of American Medical Colleges and National Resident Matching Program (NRMP) for the academic years 1994-95 to 2018-2019.[29] This dataset allows me to analyse doctors' initial speciality choices, which may be different from their final specialities. This is because doctors switch between specialities during the first few years of their training program. The initial speciality is a more relevant indicator of a physician's interest in that speciality, as compared to their final speciality, which maybe affected by various factors –- including –- workplace conditions. I add to this dataset the number of slots each participating institution is willing to offer for each speciality by compiling

[25] The process of linking the datasets is described in Appendix D.

[26] I mainly use the website https://medicalschoolhq.net/med-school-reviews/ for this purpose.

[27] For some osteopathic medical schools, I use the average debt mentioned in the website https://choosedo.org/

[28] The usual reasons for inactivity of physicians in my analysis sample are: failure to renew the license, employment in a non-medical industry, and temporarily out of the labor force. I find that 2.2% of physicians have remained inactive in the treatment sample, whereas in the control sample the corresponding figure is 2.5%.

[29] GME Track provides rich data on active matched residents for 190 + specialities.

information from NRMP reports for the academic years 1994-95 to 2018-19.

### 3.1.2 Physician access, utilisation and outcomes data

First, I obtain data on access to primary care physicians and quality of primary care for the Medicare enrollees at county level from Dartmouth Atlas of Healthcare for the years 2000-2017. Next, to test the claim of whether the policy reduced excess utilisation of hospitals or emergency rooms (ER) for Medicare beneficiaries, I use county level data obtained from Area Health Resource file and aggregated Medicare claims for the years 2000-2017.[30]

While the above data allows me to observe hospital and ER utilisation only among Medicare beneficiaries, there are many people in underserved areas who are either uninsured or covered by Medicaid. To explore how these non-Medicare populations' admissions to hospitals are affected due to higher likelihood of access to preventive care, I leverage rich patient-hospital level data from California covering the years 1999-2017.There are some distinct advantages of this individual data where patient records can be linked over time as compared to the above county level data. First, it enables me to differentiate between first time admission and readmission of patients to a California hospital. Moreover, the source of admission information present in the data permits me to distinguish transfers of patients between inpatient hospitals and admission 'directly from home/not a hospital'. [31] Second, performing the analysis with hospital fixed effects allow me to rule out concerns that treated and control hospitals may be inherently different. Third, I can determine whether my treatment effects are driven by changing composition of patients being admitted to treated hospitals. This is possible because of the availability of demographic information and detailed information on diagnosis of an admitted patient's conditions.

The National Vital Statistics system serves as my source for county level annual mortality data for the years 1999-2017. These data contain detailed information on age, sex, race, cause of death, county of occurrence and county of residence for every death in the US. The CDC assigns cause of death by examining death certificates collected from across the country.

[30]The years for which data was missing in AHRF or Dartmouth Atlas were compiled from individual Medicare claims data, aggregated to county-year level.

[31]Note that, transfer of patients between hospitals may be related to hospital quality/ capability, lack of capacity among others and cannot be easily attributed to receiving adequate preventive care and not requiring hospital/ER visit.

### 3.1.3 County level covariates

I use labor force participation and unemployment data at county level from Local Area Unemployment Statistics published by the Bureau of Labor Statistics. I obtain county population data from the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute. Population data are available by age, sex, race and ethnicity. Finally, I use Area Health Resources File to obtain the other county level covariates of interest-for example, population density, per capita income, poverty rate, uninsured rate, other health professionals like nurse practitioners and physician assistants as well as other providers of service like home health agency, and rural health clinic.

To capture treated physicians' preference for amenities, I collect county level amenities data for the pre-treatment year 1994 from various sources.[32] Following Diamond (2016), I consider six broad categories of amenities-retail amenities, crime, environment, transportation amenities, education amenities, health amenities. Data on retail amenities, transportation infrastructure, transportation and public utilities come from County Business Patterns. I observe county level spending per pupil from National Center for Education Statistics, local level education and health spending as well as spending on parks and recreation from US Census of Governments, crime and police officers data from Uniform Crime Reports,Air Quality Index from Environmental Protection Agency. [33] Finally, I obtain natural amenities data from US Department of Agriculture (USDA)-Economic Research Service. According to USDA, the natural amenities scale is "a measure of the physical characteristics of a county area that enhance the location as a place to live." The scale is calculated based on warm winter, winter sun, temperate summer, low summer humidity, topographic variation and water area, which are environmental qualities generally preferred by people.

## 3.2 Sample construction and summary statistics

The lowest geographic unit of analysis used in this paper is a county, because that is the level of treatment/policy implementation across the various states. All the treated states

[32]Note that my sample starts in 1995 and I consider only those states in the sample which have implemented the policy from 1995 onwards.

[33]Data on full time police officers is for the year 1995 from documents on law enforcement personnel provided by UCR.

have their treated areas in rural counties, with the exception of New Jersey, Rhode Island, Washington DC and Delaware. For this reason, I first classify the counties as rural/urban and restrict the main sample to rural counties. Following Albouy et al(2018), I define counties as "rural" if 1) more than 50% of the population live in a rural area within the county *or* 2) the population density is under 64 per square mile for the entire county and the total population of the county is less than 50,000.[34] I believe this definition suits my purpose, because it covers all treated counties when I restrict my sample to states having treated areas in rural counties. The counties not grouped as rural as per the above definition, form the urban counties. I also complement my main sample analysis with a robustness exercise where I include all treated counties, both in rural and urban areas.

After defining rural/urban counties, I collect the list of treated counties from the respective state health department websites or through requests to the health departments. For details, see Appendix A.

Figure 3 illustrates how the treated counties are spatially distributed across various states. While 16 states passed the policy within 1997, 16 states implemented it between 1997 and 2007 and 17 states rolled it out between 2007 and 2017. The number of treated counties varies across states, as indicated by the blue coloured areas. The states having the highest number of treated counties include Mississippi, North Dakota, South Dakota, Texas, Virginia and Wisconsin.

I further impose the following sample restrictions to precisely define the eligible group of practicing physicians. I consider only non-federal physicians, who practice either in office or hospital, and whose primary activity is direct patient care. Hence, I drop those physicians who perform only administration, teaching and research activities. Additionally, I exclude physicians who are doing internship in hospitals. The broad eligible medical specialities of physicians include family medicine, internal medicine, general pediatrics and general obstetrics and gynecology. The set of eligible specialities varies by state. [35]

[34] The federal government uses two major definitions of rural areas. The first definition is by US Census Bureau, which does not follow city or county boundaries. The second definition is by Office of Management and Budget (OMB) which classifies all counties that are not part of a Metropolitan Statistical Area as rural. While the former definition includes some suburban areas as rural, the latter definition includes several rural areas in metropolitan counties.

[35] Note that, these specialities are the primary specialities of the physicians.Family medicine and internal medicine consists of general family medicine and general internal medicine which require 3 years of training and

Table 1 shows pre-treatment summary statistics for the county level covariates in my sample. Columns (2) and (3) report means of covariates in treated and untreated counties respectively. Column (4) reports the within state difference in means and corresponding p values. The treated and untreated counties seem quite similar individually on most covariates in the baseline period. Importantly, substitute and complementary health professionals to physicians like registered nurses, physician assistants and nurse practitioners are similar in treated counties relative to untreated counties, with all p values of the differences being insignificant. Similarly, health infrastructure like skilled nursing facilities and community mental health center also seem balanced across treated and untreated counties. However, treated counties are composed of lower proportion of population aged 25-29 and have lower fraction of white people. Moreover, the joint balance test rejects the null hypothesis that the treated counties are randomly chosen (p value:0.0367). These pre-policy differences are not inherently problematic for identification, as long as they don't result in differential trends, which I show in the next section.

# 4 Empirical Strategy

## 4.1 Identification

The main goal of the paper is to exploit both within-state variation across counties as well as time variation to estimate the causal effect of the policy on physicians' entry, exit as well as health access and outcomes of the people. I use generalised difference-in-differences and event study models to estimate the average treatment effects and dynamic treatment effects of the policy. The treated group is composed of those counties where eligible physicians receive funding under the policy. The untreated group comprises of those counties within the state where the policy was not implemented. In my main specifications, I compare treated and untreated counties within a state, before and after the policy. In an ideal experiment, the treatment would be randomly assigned to observably similar counties. However, in a non-experimental framework, the identification strategy depends on the assumption that without the policy, outcomes in treated and untreated counties would have followed similar paths over time. Although this assumption cannot be directly tested, I show its validity below.

certain other specialities which require further training. For detailed list of eligible specialities, see Appendix B.

I use an "event study" framework to show that there is no difference in the dynamics of outcome variables between treated and control counties prior to the policy. Lack of such differences provides suggestive evidence that the treatment is orthogonal to determinants of outcomes, strengthening the causal interpretation of my estimates. Instead, suppose treated physicians were already increasingly entering treated counties relative to control counties in the years preceding the policy, [36] then I cannot attribute their entry to the policy.

Even though preexisting trends are absent, it is possible that the policy and outcomes are systematically correlated with unobserved shocks. This can be a confounding factor if the unobserved shocks occur in the same county and time period as the policy implementation and are not captured by the extensive set of county level covariates and fixed effects I include in the model. As I compare treated and untreated counties within treated states in a given year, and the policy has both granular space and time variation, it is less likely that such confounding factors will affect the causality of my estimates.

### 4.2 Empirical specification

I start my analysis with eligible practicing physicians by estimating "event study" models to test for the presence of confounding pre-trends and to capture the evolution of treatment effects over time. I consider the following baseline specification:

$$Y_{ct} = \sum_{k=-6, k\neq -1}^{+9} \alpha_k \mathbb{I}(t = D_c + k) + \beta X_{ct} + \gamma_c + \delta_{s(c)t} + \epsilon_{ct} \tag{1}$$

where c denotes county and t denotes year. $Y_{ct}$ denotes the log of outcome of interest in each county-year cell. I use inverse hyperbolic sine transformation to deal with meaningful zeroes in the outcome variable. $D_c$ is the year of policy implementation in county c in state s. The coefficients of interest are $\alpha_k$'s on the interaction between indicator for policy dummy $D_c$ and the indicator function $\mathbb{I}(t = D_c + k)$, where k indexes the time elapsed after roll out of the policy. I show the evolution of outcomes for 6 years before and 9 years after the policy, with the endpoints being binned for years outside this window. k=-1 is the reference year, with all coefficients interpreted relative to the year before the policy. I include

[36]This differentially higher entry before the policy can take place because of favorable economic conditions or better infrastructure or higher salaries in treated counties

the following time varying county level covariates ($X_{ct}$) in equation (1) like unemployment rate, per capita income, % population in different age groups (20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65+), % population black, % population white, % population male, % population Hispanic. These covariates increase precision and control for additional time varying differences between treated and counterfactual counties. $\gamma_c$ and $\delta_{st}$ indicate county and state-year fixed effects respectively. The county fixed effects capture time invariant county specific differences in outcomes. As economic conditions and other policies and perspectives towards migration may differ by state, counties may not be comparable across states. To construct appropriate counterfactuals, I will compare treated counties to untreated counties in the same state, as these counties are likely to be more similar along unobservable characteristics. To do this, I include state-by-year fixed effects, which removes state specific shocks and results in comparison within a state and year. The presence of never treated counties in the analysis allows me to overcome the issues raised by the recent methodological literature when estimating treatment effects in a staggered difference-in-differences framework- mainly, the issue of negative weights attached to some treatment units when averaging heterogenous treatment effects in a standard two way fixed effects regression. [37] Standard errors are clustered at the county level to account for the error terms being correlated within counties. In all estimations, I exclude the two states that have not implemented the policy.

To capture heterogeneity in the entry of treated physicians by proximity to largest metro county in the state, I take the distance between centroid of each treated county to the centroid of largest metro county in that state.[38] I divide the distance (in miles) variable into four bins- [0,50], (50, 90], (90, 130], $> 130$ and consider the following specification:[39]

[37]These papers include Goodman Bacon (2021), Abraham and Sun(2020), De Chaisemartin and D'Haultfoeuille (2020), Callaway and Sant'Anna (2020), Borusyak and Jaravel(2017), Athey and Imbens(2021), Imai and Kim (2021)

[38]I define largest metro county as 'counties in metro areas of 1 million population or more' or 'counties in metro areas of 250,000 to 1 million population' or 'counties in metro areas of fewer than 250,000 population' as per Rural-urban continuum code of USDA-ERS, depending on the state.

[39]Details on the calculation of 'distance' variable and data sources used in this regard can be found in Appendix G.

$$Y_{ct} = \sum_{k=-5,k\neq-1}^{+9} \alpha_{1k}\mathbb{I}(t = D_c + k)\mathbb{I}(Distance \leq 50)_c + \sum_{k=-5,k\neq-1}^{+9} \alpha_{2k}\mathbb{I}(t = D_c + k)$$
$$\mathbb{I}(50 < Distance \leq 90)_c + \sum_{k=-5,k\neq-1}^{+9} \alpha_{3k}\mathbb{I}(t = D_c + k)\mathbb{I}(90 < Distance \leq 130)_c +$$
$$\sum_{k=-5,k\neq-1}^{+9} \alpha_{4k}\mathbb{I}(t = D_c + k)\mathbb{I}(Distance > 130)_c + \beta X_{ct} + \gamma_c + \delta_{s(c)t} + \epsilon_{ct} \quad (2)$$

The coefficients of interest are $\alpha_{1k}, \alpha_{2k}, \alpha_{3k}, \alpha_{4k}$ on the interaction between indicator for policy dummy $D_c$ , the indicator function $\mathbb{I}(t = D_c + k)$ and the dummies for various distance bins. [40] All the other variables are same as in equation 1. The coefficients $\alpha_{xk}$'s denote how much farther from the largest metro county are marginal physicians willing to move as a result of the temporary funding they receive. Put differently, these coefficients indicate the extent to which eligible physicians in treated counties value proximity to large metro counties.

Complementary to equation (2), I also estimate two similar specifications –- one, with a continuous measure of the distance variable; and the second one with a dummy for below the median distance.

My next specification examines whether treated physicians stay in the treated county after the service period, when the positive shock to their non-wage income is turned off. Key to answering this question is a long panel of individual physician data tracking their precise workplace location. The regression I estimate is :[41]

$$Y_{ict} = \sum_{k=-6,k\neq-1}^{+9} \alpha_k\mathbb{I}(t = D_c + k) + \beta X_{ct} + \gamma_c + \delta_{s(c)t} + \epsilon_{ict} \quad (3)$$

The outcome variable $Y_{ict} = \mathbb{I}$ (individual i is still in treated county c they started in year t one year after minimum service period of that state). The coefficients of interest are $\alpha_k$'s which can be interpreted as follows. It indicates the likelihood of physician i, who starts in treated county c in state s at time t, to stay in that county 1 year after service period , relative to the

[40] For example, $\mathbb{I}(50 < Distance \leq 90)_c$ takes the value 1 if the distance between treated county and nearest largest metro county within the state is between 50 and 90 miles, 0 otherwise.

[41] Since I collapse the dataset to one observation per physician to look at the individual's retention behavior 1 year/5 years after minimum service period in that state, I don't include individual fixed effects in this specification.

likelihood of a physician who starts in c in t*-1 (t* being the year of implementation of policy in c) and stays there 1 year after service period (first difference) and relative to this difference in likelihood in control counties within s (second difference). For instance, if Illinois passed the policy in 1996 and the service period is 3 years, then the coefficient on t=0 suggests the likelihood of a physician who started practicing in treated county c in 1996 to stay in that county in 2000, relative to likelihood of physician who started in that county in 1995 and stayed there in 1999 (first difference) and relative to this difference in control counties within Illinois. (second difference). Separately, I also consider the dependent variable to be $Y_{ict} = \mathbb{I}$ (individual i is still in treated county c they started after 5/10 years). It is possible that eligible physicians may not move away from treated counties immediately after the minimum service period, because renewal of funding is allowed in some states.

Complementing the above event study specifications, I also estimate a pooled difference-in-differences model. The equation I estimate is:

$$Y_{ct} = \alpha D_{ct} + \beta X_{ct} + \gamma_c + \delta_{s(c)t} + \epsilon_{ct} \tag{4}$$

The indicator variable $D_{ct}$ takes the value 1 if county c in state s has the policy in place at time t and zero otherwise. The main coefficient of interest $\alpha$ represents the causal effect of the policy on outcome y. Note that, under the policy I consider, all treated counties within a treated state receive treatment at the same time a state is treated, and continue to remain treated during my sample period. Moreover, I consider a constant composition of treated counties within a treated state in my analysis. Because of these advantages, I am able to clearly identify 'treated' and 'untreated' places due to these policies, even in the long run. This is in contrast to various reimbursement based policies, where the treated areas are based on federal shortage area designation –- whose status changes over time –- and where it may be harder to find appropriate 'control' groups. [42]

After analysing the choices of practicing physicians, I move on to training physicians' responses. I explore effects of the policy on their choice of speciality through the following specification:

$$Y_{st} = \sum_{n=-6, n\neq -1}^{+9} \alpha_k \mathbb{I}(t = D_s + k) + \beta X_{st} + \gamma_s + \delta_t + \epsilon_{st} \tag{5}$$

[42] I also take into account all county boundary changes over the period 1995-2017.

where s indicates state and t year. The outcome variable is log number of first year training physicians. I estimate equation (5) separately for primary care/treated specialities as well as specialist/untreated specialities. $D_s$ indicates the year of policy adoption in state s. Included in equation 5 are state level covariates ($X_{st}$) to reduce standard errors of estimates and state and year fixed effects. In this equation, I cluster standard errors by state to take into account the possibility of error terms being correlated within states.

Finally, to explore policy induced access to treated physicians, ER visits, quality of health care, health spending and mortality outcomes, I estimate event study specifications similar to equation (1) and pooled difference-in-differences specification similar to equation (4).

# 5 Results

## 5.1 Effects of the policy on entry of practicing physicians

In this section, I causally estimate whether the policy succeeded in incentivizing the eligible practicing physicians to move to the treated counties. For this purpose, I first assess the pre-trends and evolution of treatment effects over time using event study specification described in equation (1). Figure 4 shows the results. The outcomes at event time t are measured relative to the year before the policy ($t = -1$).

The event study figures in (4a) and (4b) show that for both MDs and DOs, the coefficients $\alpha_k$ for the outcome log physicians per 100,000 population are flat and insignificant prior to the policy. These figures suggest that treated physicians did not disproportionately enter treated counties relative to untreated counties in the same state, prior to the policy. The lack of pre-trends within treated states strengthens the causality of my estimates. After implementation of the policy, there is a notable increase in the number of MDs and DOs in treated counties. The fact that the effect sizes grow over time indicates there is persistent entry of new physicians in the treated counties. In fact, for MDs, the long run effect sizes are twice as large as the short run ones. These effect sizes suggest that outflows of physicians, if any, are smaller than the inflows, so that net inflow at any point in time is positive. [43] Any delays in the entry of MDs –- as seen in the event study –- may be due to the transition

[43] I examine the stock of physicians as an outcome, instead of the corresponding flows –- because –- I am underpowered to detect the effects, given the unit of treatment and the demanding specification.

period of the eligible physicians between their medical school/internship and their first job.

This stickiness in physicians' location can be due to the presence of large moving costs, which comprise both pecuniary and non-pecuniary ones. The pecuniary costs include the direct costs of renting a truck, high home-ownership rates, opportunity costs of time spent on moving and costs of getting settled in a new location. The non-pecuniary costs involve getting familiar with a new setting, away from the professional and personal ties formed during the stay in the treated county. The physicians' decisions to continue staying in the treated county after the contract period coupled with no evidence on involuntary separation from their employers –- may be viewed as an improvement in the match quality between physicians and their jobs, due to these policies. Separately,I examine the outmigration decisions of eligible physicians after their service period ends, in Section 5.4. [44]

In Table 2, I report the pooled difference-in-difference estimates obtained from equation (4). As seen in columns (1) and (3), the policy increases number of MDs and DOs per 100,000 population by 5.1% and 7.8% respectively, in treated counties relative to untreated counties *within a treated state.* The addition of county level controls to the model changes the estimate only slightly, while improving the precision of estimates. The preferred estimates in columns (2) and (4) imply an increase of 2.3 MDs and 0.6 DOs per 100,000 population in treated counties. I then disaggregate the main effects of the policy into short run ("within five years of policy") and long run ("> 5 years of policy") effects, as shown in columns (5)-(8). In case of MDs, there is a short run increase of 3.5% in their number per 100,000 population in treated counties which grows to 7.1% in the long run. Importantly, the short run and long run effects are statistically different from each other (p value: 0.0073 with controls), indicating a policy induced steady entry of MDs in treated areas. In contrast, although the policy is associated with a highly statistically significant 6.62% increase in number of DOs in the short run and a 9.2% increase in long run, there is no evidence of continued entry of DOs in treated counties (p value: 0.4540 with controls). Collectively, these estimates indicate that temporary non-wage income policies are effective in moving eligible physicians to designated treated areas.

After establishing that physicians enter the treated counties, the natural question is

[44]The growth in effect sizes over time is not due to changing composition of treated states in each event study time period. I show this using a balanced panel of states in each event time period t in Figure A21.

whether physicians from the untreated rural counties sort into the treated counties to benefit from the policies. However, **??** shows almost no evidence of treatment spillovers to rural untreated counties, thereby suggesting no violation of the Stable Unit Treatment Value Assumption as well as negligible double counting of the treatment effect. A plausible reason may be that those jobs are filled to capacity by higher quality physicians on a priority basis. Additionally, the physicians in the non-treated areas may have strong preferences to stay there, or they may be ineligible under the policy –- which can also explain their lack of movement.

### 5.2 What factors drive the entry of practicing physicians?

In this section, I perform several heterogeneity tests to better understand the drivers of the entry effect. *First*, I examine the heterogenous effects on entry of treated physicians by speciality. The intuition behind this exercise is that the policy aims to make lower paying specialities more attractive for physicians in the treated areas. Hence, it is worth exploring whether these incentives are sufficient to relocate the lower income physicians to a treated county.

Figure 5 quantifies the results. I observe near zero and insignificant pre-trends for all the four specialities- general family and internal medicine, family and internal medicine speciality, general pediatrics and general obstetrics and gynaecology. This indicates that none of the eligible specialities were differentially entering the treated counties before receiving the benefits. Even after rollout of the policy, there is negligible entry of general family and internal medicine physicians in the treated counties till 6 years after implementation. Starting from the 7th year, their entry picks up with long run effect sizes of around 5% per 100,000 population. In contrast, I see an immediate 2.3% increase in number of eligible family and internal medicine speciality physicians per 100,000 population which rises to 5.3% in the long run. Similarly, the policy causes a sharp 4% increase in number of general pediatricians per 100,000 population in the short run, which continues to grow over time. The physicians specialising in general obstetrics and gynaecology also show instantaneous entry responses to the policy, with effect sizes ranging from 3.9% in the short run to around 9% in the long run. These estimates provide suggestive evidence that the generosity of the policy was more successful in moving higher income eligible physicians to the treated counties. [45] In other words, the benefit amount was not enough to move general family and internal medicine physicians to treated areas in the short run (within 5 years of policy). [46] There may be other speciality specific workplace attributes including work-life balance, perceived goodness-of-fit, size of patient caseload, and salary among other factors –- which are priority considerations for these physicians.

The difference-in-difference results imply a 2.3% (baseline mean: 41.7) increase in the number of general family medicine and internal medicine physicians(Table A1) This is driven by long run effects. This is accompanied by significant increases in the number of general obstetrics/gynaecology physicians, general pediatrics physicians, and eligible family medicine and internal medicine speciality physicians. Thus, the overall response of physicians to the policy –- though not uniform across the specialities –- is distributed among the four broad eligible ones.

*Second*, I investigate their differential entry decisions by distance to the nearest largest metro county using the event study model in equation (2). [47] This enables me to determine whether these positive labor supply policies cause physicians to locate near big cities or whether they move in to the remote counties, where their demand may be higher. Proximity to large metro counties has the advantages of easier access to urban consumption amenities, transport facilities, more opportunities for networking with other physicians and more employment opportunities for spouses, among other benefits. See Figure A2 for the distribution of treated counties relative to the big cities.

In Figure 6, I find that treated and control counties in the same state follow similar trends in the pre-policy period for all the four distance bins. After the policy, there are significant effect sizes in the distance bins closer to the largest metro counties -[0,50], (50,90] and (90,

[45] The exception may be general pediatricians who on average earn lower than general family medicine at a national level (Doximity).

[46] Note that, general family and internal medicine physicians earn among the lowest in the category of eligible physicians. The average annual compensation of general family medicine physicians in 2018 was \$242,000 while it was \$264,000 for general internal medicine physicians (Doximity physician compensation report , 2019). However, these are average compensation across metro areas and may not be indicative of the salaries of treated physicians in my sample.

[47] It is important to remember that these results are for distances between centroids, not actual travel distances.

130]. These effect sizes do not diminish over time, suggesting that these treated counties continue to receive inflows.[48] Consider the [0,50] distance band- The policy results in 8% and 14% increase in number of physicians per 100,000 population immediately and in long run respectively. In the next distance band (50,90]- effect sizes are slightly smaller in magnitude, ranging from 6% immediately after the policy to 9.5% in the long run. The third distance band (90, 130] has effect sizes in the range of 5%-6.8%. The policy leads to statistically insignificant entry of treated physicians as one moves more than 130 miles away from the largest metro county. However, these effects are imprecise and do not rule out entry of the physicians in $(130, \infty)$.

I report the pooled difference-in-difference estimates in Table A2 with five distance dummies.[49] The effects in the five distance bins are statistically different as indicated by the p value of joint test of equality of all five coefficients (p value: 0.0011). I reject equality of effects between physicians in treated counties within 130 miles and beyond 130 miles (p value: 0.0049 for '$<= 50$ miles' and '$130 - 170$ miles', p value: 0.0012 for '$<= 50$ miles' and '$> 170$ miles', p value: 0.0022 for '$50 - 90$ miles' and '$130 - 170$ miles', p value: 0.0046 for '$50 - 90$' miles and '$> 170$ miles', p value: 0.0118 for '$90 - 130$ miles' and '$130 - 170$ miles', p value: 0.0044 for '$90 - 130$ miles' and '$> 170$ miles'). This provides suggestive evidence of eligible physicians exhibiting preference for treated areas which are relatively closer to big cities. Importantly, using individual records, I find that physicians' preferences for proximity to big nearby cities are largely driven by causal effects of the policy, rather than their preferences to practice closer to their medical school. In fact, $\sim 68\%$ of the overall treatment effect is driven by the policy. [50]

*Third*, I evaluate heterogeneity in move-in decisions of eligible physicians by size of the benefits. For this purpose, I interact the indicator for policy dummy $D_c$ and the indicator function $\mathbb{I}(t = D_c + k)$ with generosity of the policy. I define generosity of the policy in county c at time t as log of total benefit amount in (c, t) cell divided by the minimum service period

[48] I get similar results when I use different specifications with the distance variable.

[49] The thresholds for distance bins correspond to the quartiles of the distance variable, with the last quartile divided into two distance bins of (130, 170] and $>$ 170 miles to clearly show the decaying effects on entry of physicians. For alternative distance bins, refer to (Figure A27).

[50] Results available upon request.

in that cell.[51]

Even after the policy, there is negligible entry of physicians in more generously treated counties (Figure A3). This may be because places with higher benefit amount among treated counties may have lower wages, lower amenities, institutions with higher average debt levels or other undesirable features which act as a barrier to their entry. An important caveat is that $log\ benefits_{ct}$ variable may be correlated with other unobserved state and county level covariates at time t which limits causal interpretation of the estimates $\pi_k$. According to the point estimate in table A3, the policy results in a 0.007% increase in the number of physicians due to a 1% increase in benefits among treated counties. This effect is statistically insignificant, but imprecise, with the 95% confidence intervals allowing me to rule out increases in number of physicians more than 0.048%. The range of elasticity estimates of the policy is consistent with Falcettoni (2020)'s estimate of the overall income elasticity of labor supply for physicians. Given the imprecision and cautionary nature of these estimates, I am unable to conclude definitively that physician inflows were not driven by benefit generosity.

*Fourth*, I examine the heterogeneous entry responses of treated physicians by pre-treatment exposure of treated counties to amenities. For natural amenities – I define high amenity counties as those with a positive amenities scale, while low amenity counties have a negative scale. I also construct an index for non-natural amenities for the pre-treatment year 1994. To create the index, I concentrate on the six broad categories of retail environment, environmental quality, transportation infrastructure, educational amenities, health infrastructure, and crime.[52] The in-migration of eligible physicians to higher amenity treated counties, whether natural or non-natural, supports their preference for these counties complementing their preference for benefits offered by the policy.[53]

### 5.3 Who is screened in by the policy?

In this section, I characterise the compliers of the policy in terms of their state of training, medical school debt, state of medical school and state of birth. This analysis focuses on

[51] I consider an alternative definition of generosity: log of total funding amount per physician in (c,t) cell divided by the minimum service period in that cell, where funding amount per physician is $min\{benefit, loan\}$. Results are unchanged.

[52] See Appendix H for details on construction of the index.

[53] See Figure A4 and table A4.

incoming physicians using individual data, collapsed to county-year level, following equation (1). In each of the event studies below, the effects are shown for different cohorts who have started their job coinciding with years since implementation of the policy.

I begin with the state of training. With no evidence of pre-existing trends, I find a significant and persistent increase in the fraction of physicians who choose to practice in their state of training after the policy is implemented (Figure A5d).The effect can be interpreted as follows: the policy's financial incentives inducing eligible physicians to enter treated counties in proximity to big cities increased the attractiveness of the job and these benefits outweighed the opportunity costs of moving out of state and searching for a new job or forming new professional connections. [54]These findings are consistent with descriptive facts in the literature i.e. 54.2% of doctors in the United States practice in their state of training (AAMC, 2018) as well as the state-wise breakup of this number.

I next examine the role of medical school debt among the compliers. The policy causes a significant rise in the proportion of ranked US medical school, and foreign medical school physicians.(Figure A5a, Figure A5b) [55] From these findings, I conclude that the benefit amount of the policy attracts some high debt foreign physicians as well as physicians from lower ranked US medical schools. At the same time, it deters some higher debt physicians from unranked US medical schools (Figure A5c). [56] This is likely because of the design of the policy that physicians will not be provided their entire loan amount if the benefit amount of the state in which they practice is less than their accumulated debt. These results provide suggestive evidence of the targeting effects of the policy. [57]

[54]There is anecdotal evidence that generally, in small states like New Hampshire, full time employment opportunities for doctors are few and crossing state lines are common. Long-distance moves between states may be facilitated because licenses can be easily transferred between states (Johnson and Kleiner, 2019). Hence, only 38.8% of doctors who train in New Hampshire practice there(AAMC). It is likely that the policy has successfully incentivised some eligible physicians to stay in these states after training.

[55]Figure A6 presents the relationship between medical school rank and average medical school debt.

[56]Some unranked medical schools have average debt higher than $295,000 like A.T. Still University of Health Sciences-Mesa, Pacific Northwest University of Health Sciences and Campbell University (Wallace).

[57]In contrast to training results, I find that majority of physicians who practice in treated counties are from out-of-state medical schools (Figure A7).Furthermore, among those who practice in treated counties after the policy, I find a slightly higher proportion of out-of-state birthplace physicians than out-of-state medical physicians.

## 5.4 Effects of the policy on exit of practicing physicians

The goal of the policy is not only to hire new eligible physicians to treated counties, but also to retain them , even after the minimum service period. While literature documents that physicians tend to move to underserved areas, their retention remains a problem (American Medical Association, American Academy of Family Physicians (AAFP)). Due to lack of long panel of individual employment data that spans the timeline of adoption of policy, causal estimates of turnover of eligible physicians in treated counties are lacking. It is possible that after the policy lapses, the cost of staying in the treated county increases sufficiently for the marginal physician, so as to outweigh the benefits offered by the policy. This additional cost can include commuting for jobs by spouses, ability to manage emergencies,difficult medical situations and busy outpatient practices without consultants or advanced technology among others (AAFP). If these costs are sufficiently large, eligible physicians may be willing to move out of the treated county,but stay within the state. To explore how the policy affects different cohorts of eligible physicians' probability of staying in a treated county after their contract period, I follow specification (3).

I first show how likelihood of staying one year after the minimum service period responds to the policy, where the service period varies by state. The estimates of Figure 7 indicate an economically small but statistically significant 1.15 percentage point (baseline mean: 0.814) decrease in likelihood of staying in a treated county. But, this decline does not persist for the recent cohorts. [58] Important to note here is that, the net inflows of physicians is positive, implying that a temporary government program has persistent effects on physicians' location choices. I find some suggestive evidence using Census and American Community Survey data that these retention patterns are driven by married physicians, or more precisely working couples.

## 5.5 Effects of the policy on speciality choice of training physicians

Do temporary positive income policies affect the initial speciality choices of training physicians? In other words, do these policies induce more medical students to join the treated specialities, thereby expanding the supply of eligible specialities? To answer this question, I

[58] Figure A8 shows qualitatively similar results for propensity of moving away from a treated county 5 years and 10 years after the contract period.

first examine how the policy affects the number of first year residents in treated specialities.[59]

The estimates in Figure 8 are based on event study specification 5. I consider the specification at a broader level of the entire state instead of focusing on the main sample of rural counties within a state because of the following reason. For example, it is possible that eligible physicians may complete their training in Chicago and choose to move to a treated county in Illinois because of the policy. I may ignore this group of eligible physicians if I confine my analysis to the rural sample within a state during the residency stage. [60] Figure 8 shows no pre- policy trends, evidence that treated states did not experience disproportionate change in first year residents in treated specialities relative to control states. Immediately after roll-out of the policy, there is a 2% increase in number of residents joining a treated speciality in treated states. This effect grows over time. One explanation can be that marginal medical school students, upon graduating, anticipate that the policy will sufficiently improve the lifetime income prospects of practicing in the treated specialities.This can increase their willingness to join a treated speciality. The difference-in-difference estimates imply a 2.2% (baseline mean: 396) increase in number of primary care/eligible residents in treated states after the policy (Table A6) .

The above results provide evidence that the policy increases the number of matched residents entering the treated specialities in treated states. This outcome can capture trainees' preference for specialities (labor supply) or residency programs' preference for trainees (labor demand). To disentangle the labor supply and demand channels, I look at how the residency programs respond to the policy in terms of slots they are willing to offer in treated specialities in each participating institution. I show in Figure A9 that timed with the policy, there is negligible and insignificant increase in residency positions that programs are willing to offer. Thus, I can rule out the labor demand channel of programs expanding to accommodate the growth of matched residents. With unfilled slots available in treated specialities in the pre-

[59]To be clear, AMA data also allows me to observe physicians' initial speciality and switching between specialities during their internship. I find qualitatively similar results on initial speciality decisions of the eligible physicians, using the larger AMA data.

[60]Note that, one may not make a direct comparison between increase in training physicians and increase in practicing physicians in a state after the policy is passed. Suppose Indiana passes the policy before Illinois. It may happen that when Illinois passes the policy, there is an increase in internal medicine residents in already treated state Indiana who are willing to move to Illinois to practice.

treatment period, I conclude that increase in the number of matched residents signals a labor supply response - an implicit interest in treated specialities.[61]

Moving on, I consider the obvious question: are any untreated speciality residents induced to practice in treated specialties as a result of the policy's monetary incentives? Indeed, I find evidence of substitution between treated primary care and untreated speciality residents: a 1.2% decline in count of first year untreated residents in treated states, which grows over time (Figure 8). The difference-in-difference estimate indicates a 2.43% (baseline mean: 595) decrease in the number of ineligible speciality trainees (Table A6).

The higher entry of residents in the treated specialities after the policy obscures significant heterogeneity across the specialities. To uncover this heterogeneity, I estimate specification 5 separately for four groups of matched training physicians- General Family medicine and Internal Medicine, General Obstetrics/Gynaecology, General Pediatrics, Family and Internal Medicine speciality. General family and internal medicine residents exhibit around 6% increase in their number for two years after the policy (Figure A10). Beyond two years, their entry persistently declines and becomes insignificant over the years. This pattern is in line with prior descriptive evidence documenting decreasing interest in these specialities. Recall that, my results also showed increase in practicing general family and internal medicine physicians in treated areas within treated states after 6 years of the policy. A possible explanation to reconcile these findings is that the increase in practicing physicians comes from those who completed their training out-of-state. On the contrary, there are sustained increases in the number of first year general obstetrics/gynaecology, general pediatrics and certain family and internal medicine speciality trainees.

## 5.6 Effects of the policy on physician access, quality of health care and health outcomes of people

Primary care physicians comprise a first point of contact of the population to the health care system.They refer patients to higher levels of care and ensure care coordination through information transfer to specialists (Starfield, 1994).

One goal of the policy in incentivising newly trained physicians to move to underserved

[61]In case of treated primary specialities, the capacity constraint is not as binding as in case of higher paying surgical specialities.

areas is to increase people's access to those physicians so that they face lower number of procedures prescribed by specialists and hence reduced costs. However, ex-ante it is not clear whether patients' demand for the new physician will increase when they enter a treated county. On one hand, seeing established physicians may be more beneficial for patients. The reason can either be purely psychological such as beliefs that common racial background promotes communication and willingness to follow the physician's advice (Alsan et al, 2019). Else it can also be related to patient's health, as pointed out by research on continuity of care (David and Kim, 2018; Agha et al, 2019). In this scenario, patients may not visit the new physician. [62] On the other hand, new physicians may be more conveniently located or more cost effective, thereby increasing their demand among people, especially high risk ones or those lacking a relationship with existing physician. Additionally, new innovations in healthcare have made it possible to transfer detailed patient records to the new physician, thereby reducing patients' preferences for particular physicians. At the same time, standardisation of medications prescribed by physicians may reduce the importance of skills/expertise of any particular physician (Goldin and Katz, 2016).

To empirically determine patients' allocation of care between newly entering physicians and emergency room (ER) or hospital use, I follow event study specification identical to equation (1). Essentially, I compare the following outcomes: number of visits to physicians, take-up of preventive services, adverse events like emergency department (ER) visits, hospital admissions and mortality between treated and untreated counties within treated states, before and after the policy. Visits to physicians, ER visits and hospital admissions seem to trend similarly for treated and untreated counties prior to policy. The plotted estimates in Figure 9a indicate a 2% increase in per capita medicare enrollees having at least one ambulatory visit to a physician in year 1 after roll-out. Because it may take time to build a relationship with a new physician, the increase in visits occurs with a one-year delay.[63] Importantly, the effects do not drop off over time, suggesting a continued increase in utilisation of new physicians' services.

To further assess the likelihood of adverse advents when people of treated counties are

[62] Patients are willing to travel large distances to continue being treated by their original physicians (Sabety, 2021)

[63] The dependent variable is age, sex and race adjusted. Ambulatory conditions are treatable in office/outpatient settings by physicians.

able to access preventive care, I quantify how ER visits evolve timed with the policy in Figure 9b. The results show a notable 4% decline in emergency department visits after the policy, which further drops to 6% in years 6 and beyond. [64] Figure A11 graphs the evolution of preventable hospital stays rate, coincident with timeline of policy. There is a 5% decrease in hospital admissions right after the policy and the magnitude expands over time. [65] I propose two possible explanations for the results in Figure 9b and Figure A11. First, patients are being treated preventatively or for a chronic condition by the physician, so it is not increasing the severity of the condition, leading to the need for ER visit. Second, before the policy, patients were utilizing ER and hospitals for primary care treatable conditions while after the policy, they visit the newly entering physicians to receive the treatment. [66]

The following set of results emerge from the pooled difference-in-difference estimates reported in Table 3. *First*, there is a modest and statistically significant 2.1% (baseline mean: 0.081) increase in per capita Medicare enrollees having at least one ambulatory visit to physician in treated counties relative to control counties within a state, as shown in column 2. *Second*, the policy causes a 4.98% (baseline mean: 0.096) reduction in per capita emergency department visits of Medicare beneficiaries, consistent with them receiving more preventive care. *Third*, I find a 10.2% (baseline mean: 80.19) decline in preventable hospital stays rate of Medicare enrollees.

I supplement the above analysis of hospital admissions at a county-year level with longitudinal patient level data from California. The advantages of this dataset over the aggregated Medicare data are discussed in section 3.1.2. To keep the analysis as comparable as possible to the aggregate data, I mainly focus on *index hospital admissions* for Medicare patients. The increased visits to entering physicians can be reflected in either new patients not requiring hospital care or patients who were going to hospital for preventive care before the policy but are now reducing their readmissions. I focus on first time admissions instead of readmissions to more cleanly identify the effects of the policy. This is because, hospital readmission is often

[64] The fall in ER visits is not driven by closure of facilities or reduction of hospital capacity.

[65] Preventable hospital stays is number of hospital admissions/discharges for ambulatory conditions per 100,000 fee-for-service adult Medicare enrollees.This measure is age-adjusted.

[66] Some of the leading conditions for which adults visited ER in rural areas were sprains and strains, contusions, abdominal pain, headache and back problems. On the other hand, children largely visited ERs for upper respiratory infections, contusions, ear infections, sprains, strains and open wounds. (HCUP, 2011)

regarded as an indicator of poor quality care during hospital stay (MEDPAC 2007). It has been heavily used by economists as proxy for quality of hospital care (Cutler (1995); Ho and Hamilton (2000); Kessler and Gepert(2005)).

I estimate:

$$Y_{ht} = \sum_{k=-4, k\neq -1}^{+7} \alpha_k \mathbb{I}(t = D_{c(h)} + k) + \phi Z_{ht} + \beta X_{c(h)t} + \gamma_h + \delta_t + \epsilon_{ht} \tag{6}$$

In this equation, Y is log number of admissions of payer category p in hospital h in year t. The coefficient of interest is $\alpha_k$, which shows the interaction between policy dummy $D_c(h)$ and the indicator variable $\mathbb{I}(t = D_{c(h)} + k)$. Therefore, I essentially compare admissions in treated hospitals with control hospitals, before and after policy. Z is mean characteristics of patients admitted to hospital h in year t (average age, share white, share female).[67] X denotes county level covariates similar to equation 1. I also include hospital fixed effects , year fixed effects and cluster standard errors by hospital. I weight the regressions by $T + (1 - T) * \frac{p}{(1-p)}$ where T is the indicator for treatment and p is the estimated propensity score. This weighting method provides a consistent estimator of average treatment effect of the policy. [68]

The identifying assumption in specification (8) is that the only reason for changing admissions in treated hospitals relative to matched control hospitals is the policy.

Recall that, our treatment is at a county level.To address pre-policy differences between treated and control counties, I construct a set of control counties from untreated counties in California that most closely resemble the treated counties. For this purpose, a propensity score model is fitted which predicts the treatment status of a county. I estimate the model using a rich set of covariates selected by the double lasso procedure.

Figure 10a-d and Figure A12 plot the event studies for log number of admissions of the following four major payer categories-Medicare, Medicaid, private insurance and uninsured as well as unscheduled admissions. I observe that treated and control hospitals in California trend similarly prior to the policy. Consistent with the aggregated county level results on hospital admission, I find a modest yet significant decline of 0.67%-1.02% in admissions of

[67] I have checked that these patient demographics are unaffected by the policy. Furthermore, I find no changes in the case mix index of treated hospitals timed with the policy, implying that the severity of illness of patients admitted to these hospitals has not changed.

[68] As a robustness check, I also weight the regressions by baseline total admissions multiplied by $T + (1 - T) * \frac{p}{(1-p)}$ and find qualitatively similar results.(Gruber and Kleiner, 2012).

Medicare patients in treated California hospitals after the policy (Figure 10a). Emergency admissions also witness a small but significant decline in treated hospitals, consistent with the drop in ER visits seen in county-level data (Figure A12). In addition, Figure 10b shows a sustained drop in admissions of Medicaid patients after the policy –- greater than the decline in Medicare admissions. [69] The difference-in-difference estimates in Table 4 suggest that the policy's effect on the extensive margin of hospital admissions in California depends on the insurance coverage of the patient.

To understand the change in government spending due to shifting away from ER or hospital based care to timely preventive care,I begin by determining whether the policy affects hospital reimbursement.[70] The estimates in Figure A13a and Table A7 imply a 1.44% (baseline mean: $4153.35) reduction in hospital reimbursements per Medicare enrollee in treated counties after the policy.

Next, I turn to explore the effects on physician reimbursement. One major advantage of the Dartmouth physician reimbursement data is that it provides a better measure of the value of services provided, as it is based on the actual price physicians receive which may be higher than the price reimbursed by Medicare. The variation in these two amounts may arise due to copayments and coinsurance paid by insurers or patients (Dartmouth report, 2010). I find that the reimbursement per enrollee remains flat after the policy, as shown in event study Figure A13b. The negligible effect on physician reimbursements in treated counties provides suggestive evidence of eligible physicians entering treated counties only due to non-wage benefits provided by the policy.

Finally, I examine how the policy affects quality of primary care received by Medicare patients. I look at take-up of the following preventive care services: mammogram, blood lipids test, eye examination, hemoglobin a1c test. I find no evidence of degradation of quality of primary care in treated counties (Figure A14 and Table A8). Overall, the policy reduces

[69] This result seems consistent with Medicaid patients having additional difficulty in finding physicians ready to treat them (Candon et al 2018; Oostrom, Einav and Finkelstein, 2017). Furthermore, doctors are denied reimbursement for 25% of initial Medicaid claims they submit (Dunn et al 2021).

[70] Both hospital and physician reimbursements are price, age, race and sex adjusted. The price adjustment adjusts not only for the cost of living, it also removes additional payments by Medicare to facilities for resident training and disproportionate share hospital (DSH) program for hospitals that accept a higher fraction of low income patients.

healthcare costs with minimal effects on quality.

### 5.7 Effects of the policy on mortality of the people

In this section, I examine another metric of adverse event: mortality. I expect the main channel through which the policy can plausibly affect mortality outcomes is through improved access to eligible providers.[71] Empirically, the net effect on mortality depends on the magnitude and direction of impact of the policy on access to physicians for preventive care or chronic illness treatment, quality of primary care received and patterns of utilisation of hospitals and emergency departments. While I find increased visits to newly entering physicians and reduced admission to hospitals and ER on account of the policy, I find no evidence of significant improvement in quality of protective care by physicians. This makes it worth investigating whether places where physicians move in after the policy experience larger declines in mortality relative to other places.

In my main analysis, I focus on mortality of elderly population aged 65+, disaggregated by cause of death. This is because I show higher access to physicians for the Medicare population. [72] I follow Sommers et al (2014) to classify amenable conditions using ICD-10 codes available in the mortality data.

In Figure A15a, all cause mortality shows a clear decrease in treated counties in the post-policy period, starting with a notably significant reduction of approximately 2 deaths per 10,000 population. The reduction in mortality remains fairly constant over the post-policy period. These lower aggregate mortality rates for the elderly are primarily driven by amenable causes- those which are considered to be more responsive to timely medical care (Figure A15b). I observe statistically significant declines in cardiovascular mortality of 0.6 deaths per 10,000 population immediately after the policy, which is sustained during the post policy years (Figure A16a). People suffering from cardiovascular diseases largely benefit from access to correct medication like statins. Having more providers may increase the scope for effective treatment of patients with cardiovascular disease and reduce their risk of heart

[71] Note that, while factors like risky health behaviours, insurance coverage rates of the population , accurate diagnosis and effective treatment may affect health outcomes and in the extreme case, mortality, these are not influenced by the policy.

[72] In Figure A17, Figure A18, I show the mortality patterns of all adults in treated counties timed with the policy, disaggregated by detailed cause of death.

attacks or stroke. Similarly, Figure A16b shows evidence of modest declines in mortality of elderly people due to drug and alcohol poisoning. Family physicians play an important role in the treatment of alcohol and drug withdrawal in the elderly population (American Family Physician, 2000). Positive responses on this disease margin may indicate that family physicians are capable of identifying and treating older patients with these problems. Other causes of death exhibit little response in treated counties following the rollout of the policy.

Even though my point estimates on all-cause mortality are large relative to the reduction of 15 deaths per 100,000 population for the age group 20-64 years due to Medicaid expansions (Borgschulte and Vogler, 2020), they are included in their 95% confidence intervals. Overall, our estimates are in line with the separate literature on Medicaid expansions/health insurance coverage (Sommers et al (2014); Sommers (2017); Swaminathan et al(2018); Miller et al(2021); Goldin et al(2021), Khatana et al(2019); Borgschulte and Vogler(2020).) However, our estimates are not directly comparable, as the above literature considers an insurance based policy with different affected population, time period of analysis, and definition of treatment and control areas. Instead, our elderly mortality reductions of approximately 2.35% more closely align with the literature on Medicare and mortality (Huh and Reif (2017); Card et al (2009); Kaestner et al (2019); Finkelstein and McKnight(2008)).[73]

After characterising the policy induced overall mortality effects, I examine heterogeneity of the effects by pre-treatment physician availability. I focus on the Medicare population to maintain consistency with the earlier results. The estimates in Table A22 imply that among treated counties, adding a physician to a place with more physicians at baseline reduces overall mortality by an additional 0.1251 deaths per 10,000 population relative to a place with fewer physicians at baseline. This effect is largely accounted for by amenable cause mortality.

## 6 Robustness and Alternative Specifications

I carry out a number of placebo exercises to rule out alternative explanations. In Appendix C, I also present a range of robustness checks and alternative specifications to strengthen the credibility of my results. I summarise them here. First, I begin with falsification tests. The policy incentivised newly trained physicians to enter treated counties. Consistent with this,

[73]While our estimates are larger than studies by Finkelstein and McKnight(2008) ; Kaestner et al(2014), they are smaller in magnitude than Card et al(2009).

Figure A19a and Table A9 show flat and insignificant entry of older physicians aged 45 and above after the policy. Additionally, the policy was tailored to specific categories of eligible physicians. Thus, Figure A19b and Table A9 show almost null entry effects for specialists, who were not subjected to the policy. In Section 5.6, I further show that the entry of physicians in the treated areas is not driven by changes in physicians' salary incentives or reimbursement rates. These policies were targeted to state-level jobs, and hence federal physicians were not eligible for them. I find evidence of negligible entry of these physicians in the treated counties (See Figure A20)[74] Each of these placebo tests support that confounding factors may not drive my estimates.

Then, I move on to decomposing the difference-in-differences estimator into five groups of 2X2 estimators as per Goodman-Bacon(2018). Results are shown in Table A10. Around 94%-95% of the baseline estimate is explained by comparisons between treated and never treated counties. This finding is consistent with the fact that the difference-in-difference estimates are largely similar to the estimates of event studies, which are relatively more robust to problems that arise when units are treated in a staggered manner. Table A11 shows the Goodman Bacon decomposition for training physicians' speciality choice estimates. I also check the robustness of my main estimates for practicing physicians, and health outcomes using the doubly robust estimator of Callaway and Sant'Anna (2021). This approach enables me to estimate the average treatment effect on the treated (ATT), separately for each cohort of treated counties 'g' and each year t. Reassuringly, the estimates are highly similar to my baseline results.[75]

Next, I present results from several alternative specifications : balanced panel fixed effect models(Figure A21-Figure A26), county population weighted specifications (Table A12). The estimates from the above specifications are very close to the main sample. Some states like New Jersey, Rhode Island, Washington DC and Delaware have their treated counties in urban areas. When I include these states in my analysis, the estimates are slightly larger, and presented in Table A13. Finally, to address the threat of physicians sorting into specialities either due to their own preferences , or due to some policies, or due to varying selection procedures and job attributes of different specialities –- I estimate the hiring rates of physicians

[74] In this analysis, I restrict my sample period to 1995-2009, to avoid confounding the effects with the Affordable Care Act's broad-based set of reforms that affected federal physicians.

[75] Results are not provided here due to space constraints, but available on request.

within a speciality and find qualitatively similar results, in fact mildly larger effect sizes.[76]

## 7 Conclusion

In this paper, I attempt to isolate the causal effects of temporary labor supply programs for new physicians on their migration flows, speciality choice and the health outcomes of the people in the areas where these physicians move in –- as well as unpack the channels explaining the effects. For this purpose, I combine the staggered rollout of the state and local PCP loan repayment programs over the last four decades, in hundreds of counties across 49 US states, with individual longitudinal data that tracks the universe of physicians' employment details from their medical school to mid-career. I find that the program has persistent effects, increasing the number of MDs and DOs by 4.9% and 7.8% respectively, in treated counties relative to untreated counties, within a treated state. My estimates of labor supply elasticity for physicians with respect to the benefits offered by the program rule out effects more than 0.48% due to a 10% rise in benefits among the treated counties.

Importantly, the inflows of newly hired physicians are largely driven by higher-paying eligible specialities who prefer to locate closer to the big cities within the state. Thus, the effects of these policies are unequal across the counties.

I also document a small decline in a physician's probability of remaining in the treated county one year after the minimum contract period of that state. Similar patterns are observed both 5 years and 10 years after the contract period. These estimates are economically small and noisy –- and therefore –- I conclude stickiness in their location choices, even when the program expires. Indeed, their net inflows continue to be significantly positive and grow over time.

While I observe an increase in physicians in treated areas relative to untreated areas, this may simply be a reallocation of physicians within the state, with no aggregate welfare effects. But I demonstrate an increase in trainees entering the higher paying eligible specialties in treated states after the policy. This is achieved by a more than proportionate decline in the trainees of ineligible specialities. To derive the welfare estimates of such programs, it is necessary to consider how the initial supply and composition of the applicant pool of physicians change, in addition to the usual redistribution of physicians across space and time.

[76]Results available on request.

Finally, I show evidence of significant benefits to the elderly and most vulnerable people residing in the treated counties. There is a persistent 2% increase in per capita Medicare enrollees having at least one ambulatory visit to a physician in the treated counties. Furthermore, the policy causes a significant fall in ER visits and hospital admissions of the elderly, which is also confirmed in a hospital level analysis for the state of California. A back of the envelope calculation suggests that $\sim \$149,160$ is saved in a treated county due to the decline in ER visits and hospital admissions of Medicare enrollees alone, with additional savings arising from such declines for the non-elderly population, and mortality reductions for the elderly population. The Marginal Value of Public Funds (MVPF) of the program is 0.6, indicating that every $1 of net government spending provides $0.6 of benefits to the beneficiaries of the policy.[77]

More broadly, the findings of this paper suggest that temporary supply-based incentives aimed at correcting market failures and information problems do work. They can elicit persistent responses from liquidity constrained physicians. There are economically meaningful health improvements in both the elderly and non-elderly population. Simultaneously, there appears to be unequal effects of these programs across the counties, as well as muted entry responses for general family and internal medicine physicians –- which must be addressed for improving the program design, delivery and effectiveness.

The paper's results have a few limitations. For policy purposes, it is important to understand whether the gains in the treated counties due to the program came at the expense of other parts of the country. The estimates in this paper inform us of the local mobility and health effects, which may not necessarily mirror the aggregate effects. Hence, I am currently unable to draw conclusions about the efficiency aspect of these labor supply programs. A next logical step is to consider the long run income implications of physicians' movement across locations, as well as, the quality of physicians attracted by these programs –- which can shed light on the equity-efficiency tradeoff of these programs.

[77]MVPF of a $1 rise in cash benefits is given by $\frac{1}{1+FE}$,where FE is the fiscal externality of the program per dollar increase in spending on each infra-marginal beneficiary of the program. For details, see Finkelstein and Hendren(2020).

# References


Agha, L., Frandsen, B. and Rebitzer, J.B., 2019. Fragmented division of labor and healthcare costs: Evidence from moves across regions. *Journal of Public Economics*, 169, pp.144-159.

Albouy, D., Farahani, A. and Kim, H., 2018. The Value of Rural and Urban Public Infrastructure, *Working paper*

Bolduc, D., Fortin, B. and Fournier, M.A., 1996. The effect of incentive policies on the practice location of doctors: a multinomial probit analysis. *Journal of labor economics*, 14(4), pp.703-732.

Callaway, B. and Sant'Anna, P.H., 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics.*

Candon, M., Zuckerman, S., Wissoker, D., Saloner, B., Kenney, G.M., Rhodes, K. and Polsky, D., 2018. Declining Medicaid fees and primary care appointment availability for new Medicaid patients. *JAMA internal medicine*, 178(1), pp.145-146.

Card, D., Dobkin, C. and Maestas, N., 2009. Does Medicare save lives?. *The Quarterly Journal of Economics*, 124(2), pp.597-636.

Carrillo, B. and Feres, J., 2019. Provider supply, utilization, and infant health: evidence from a physician distribution policy. *American Economic Journal: Economic Policy*, 11(3), pp.156-96.

Costa, F., Nunes, L. and Sanches, F.M., 2019. How to Attract Physicians to Underserved Areas? Policy Recommendations from a Structural Model, *Working Paper*, first draft 2015.

Dal Bó, E., Finan, F. and Rossi, M.A., 2013. Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, 128(3), pp.1169-1218.

Deserranno, E., 2019. Financial incentives as signals: experimental evidence from the recruitment of village promoters in Uganda. *American Economic Journal: Applied Economics*, 11(1), pp.277-317.

Dunn, A., Gottlieb, J.D., Shapiro, A., Sonnenstuhl, D.J. and Tebaldi, P., 2021. A Denial a Day Keeps the Doctor Away, *Working paper*

Falcettoni, E., 2018. The determinants of physicians' location choice: Understanding the rural shortage. *Working Paper*, Available at SSRN 3493178.

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

Finkelstein, A. and Hendren, N., 2020. Welfare analysis meets causal inference. *Journal of Economic Perspectives*, 34(4), pp.146-67.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Huh, J., 2021. Medicaid and provider supply. *Journal of Public Economics*, 200, p.104430.

Hurley, J.E., 1991. Physicians' choices of specialty, location, and mode: A reexamination within an interdependent decision framework. *Journal of Human Resources*, pp.47-71.

Kaestner, R., Schiman, C. and Alexander, G.C., 2019. Effects of prescription drug insurance on hospitalization and mortality: evidence from Medicare Part D. *Journal of Risk and Insurance*, 86(3), pp.595-628.

Khatana, S.A.M., Bhatla, A., Nathan, A.S., Giri, J., Shen, C., Kazi, D.S., Yeh, R.W. and Groeneveld, P.W., 2019. Association of Medicaid expansion with cardiovascular mortality. *JAMA cardiology*, 4(7), pp.671-679.

Kruk, M.E., Johnson, J.C., Gyakobo, M., Agyei-Baffour, P., Asabir, K., Kotha, S.R., Kwansah, J., Nakua, E., Snow, R.C. and Dzodzomenyo, M., 2010. Rural practice preferences among medical students in Ghana: a discrete choice experiment. *Bulletin of the World Health Organization*, 88, pp.333-341.

Lee, S., 2010. Ability sorting and consumer city. *Journal of Urban Economics*, 68(1),20-33.

Miller, S., Johnson, N. and Wherry, L.R., 2021. Medicaid and mortality: new evidence from linked survey and administrative data. *The Quarterly Journal of Economics*.

Oostrom, T., Einav, L. and Finkelstein, A., 2017. Outpatient office wait times and quality of care for Medicaid patients. *Health Affairs*, 36(5), pp.826-832.

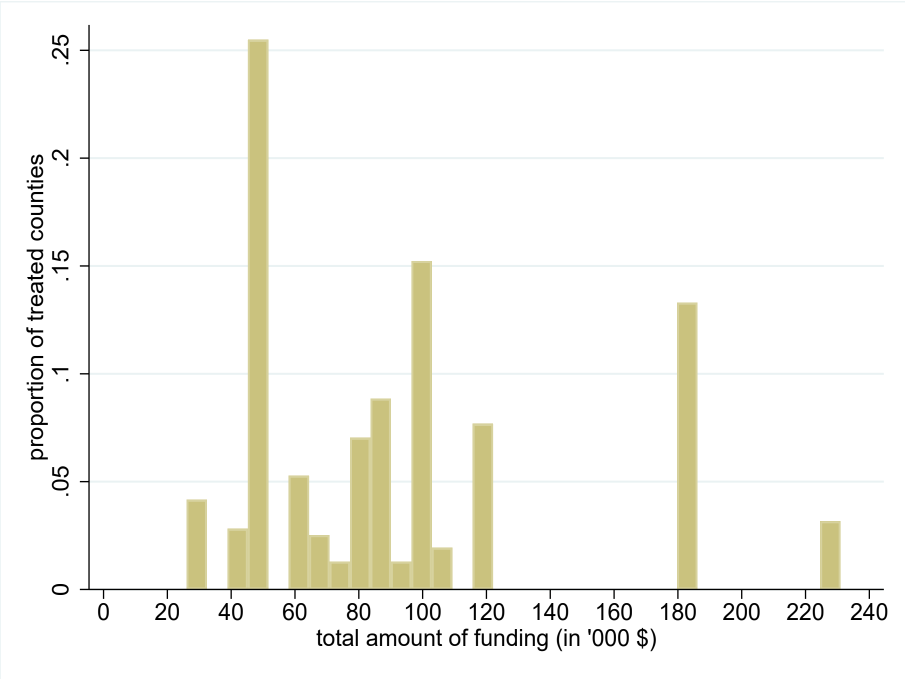
Sabety, A.S., 2021. The value of relationships in health care. *Working paper.* https://www.adriennesabety. com/research.

Sommers, B.D., 2017. State Medicaid expansions and mortality, revisited: a cost-benefit analysis. *American Journal of Health Economics*, 3(3), pp.392-421.

Starfield, B., 1994. Is primary care essential?. *The lancet*, 344(8930), pp.1129-1133.

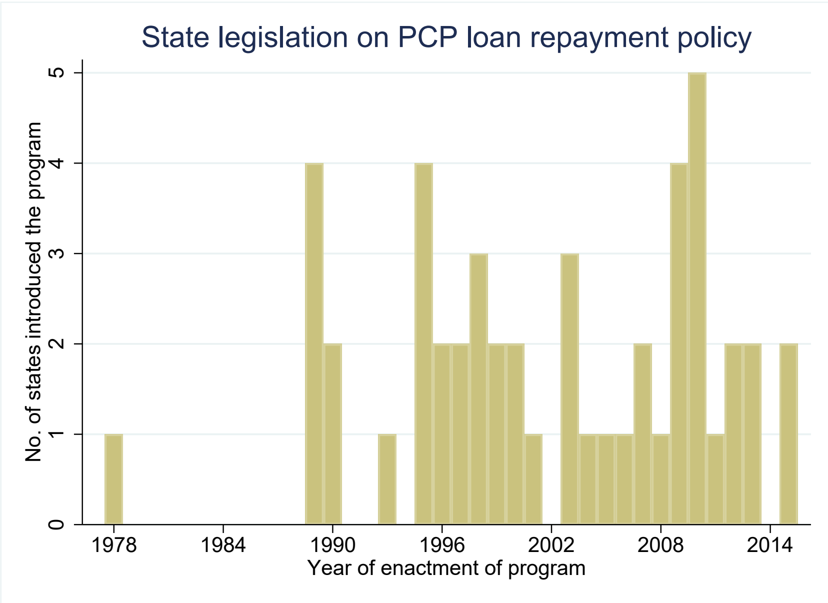
Wasserman, M., 2019. Hours constraints, occupational choice, and gender: Evidence from medical residents, *Working paper.*

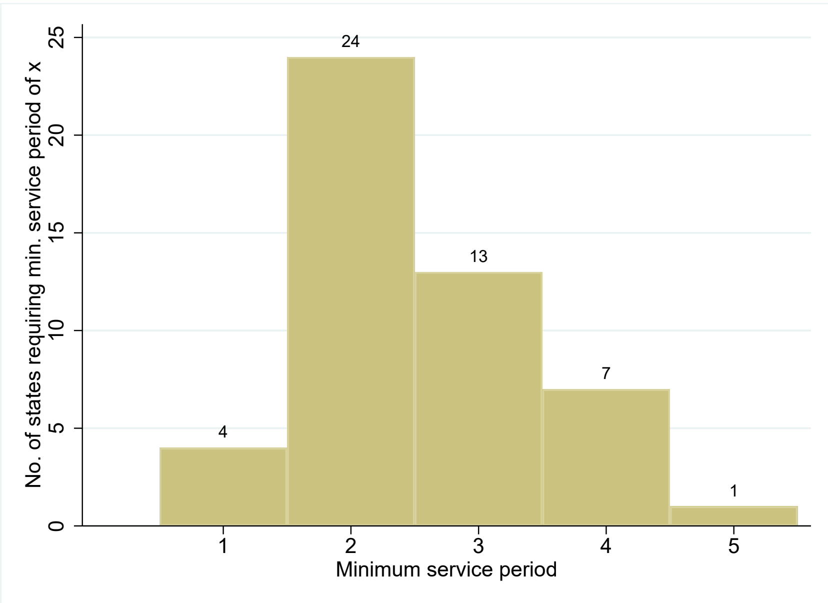
Figure 1: Funding offered to physicians under the policy




*Notes:* The above figure shows the distribution of funding amount offered in various counties for the duration of minimum service period. Data source: State health department websites.

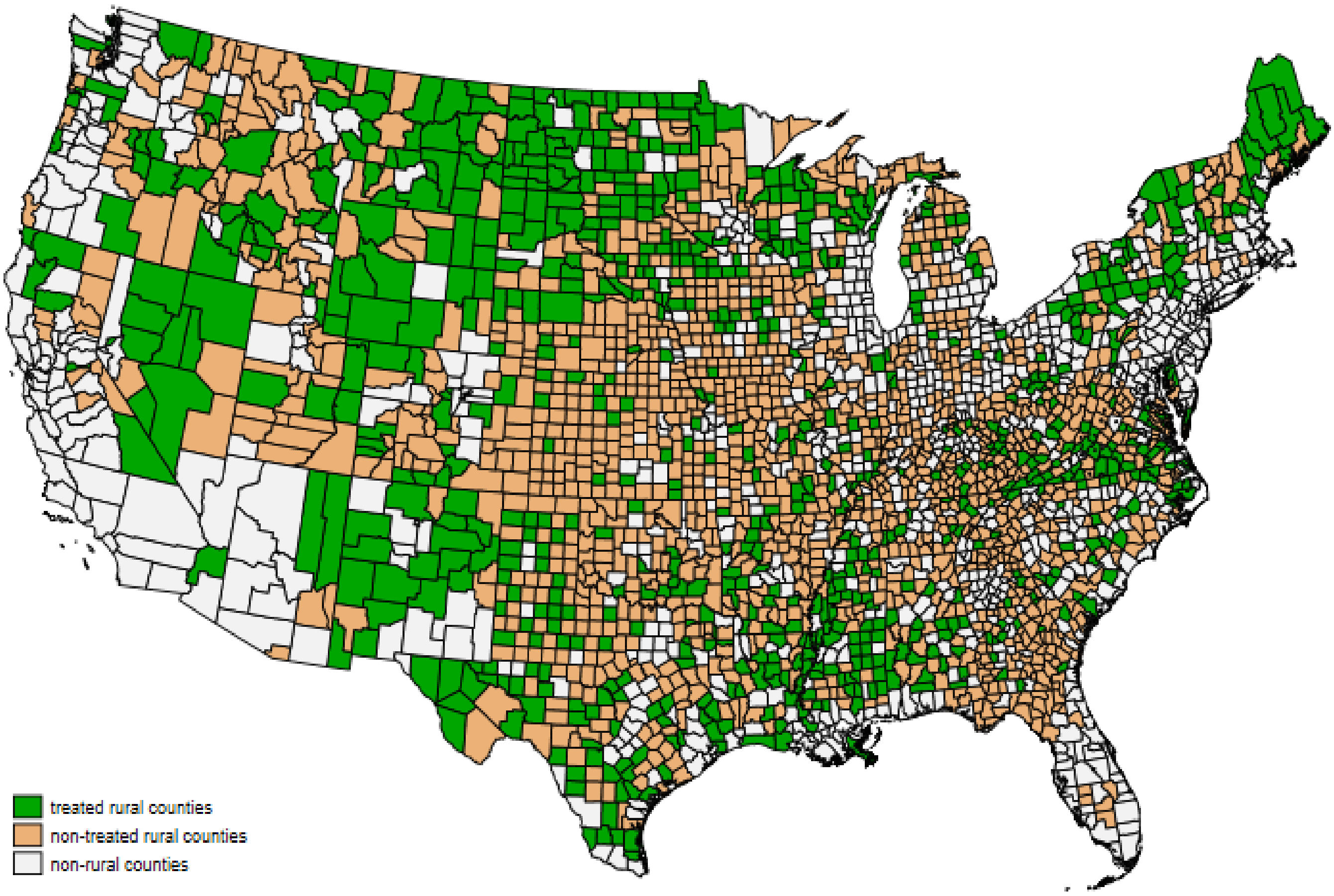
Figure 2: Timeline of adoption and minimum obligation period of policy



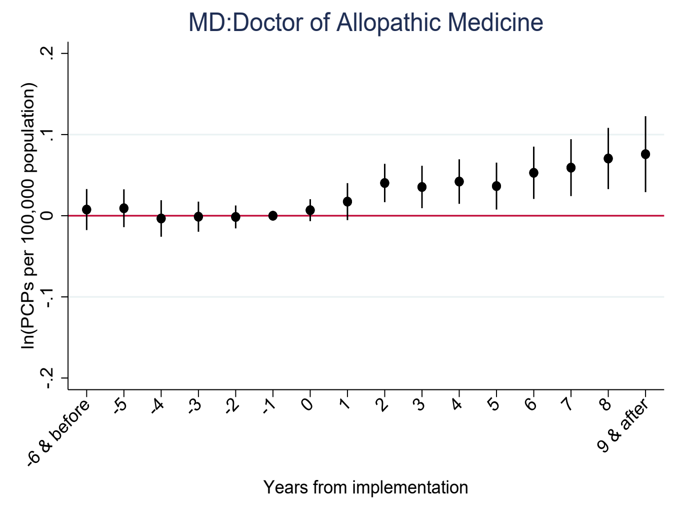

*Notes:* The top figure illustrates the number of states that have introduced the policy only for eligible physicians in a staggered manner over the period 1978-2015. The bottom figure shows the distribution of minimum service period that is required by the policy. Data Source: Original legal documents, state health department websites.

Figure 3: Spatial distribution of treated counties

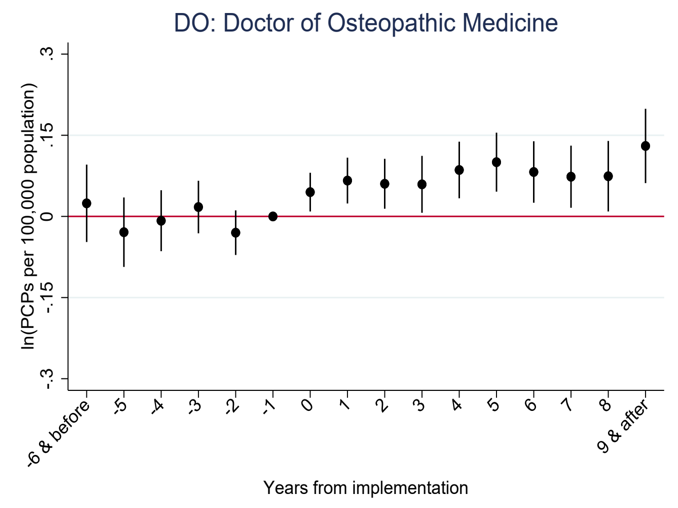



*Notes:* Even though Hawaii and Alaska are not shown in this map, they are included among the treated states.

Figure 4: Effects of the policy on entry of eligible practicing physicians



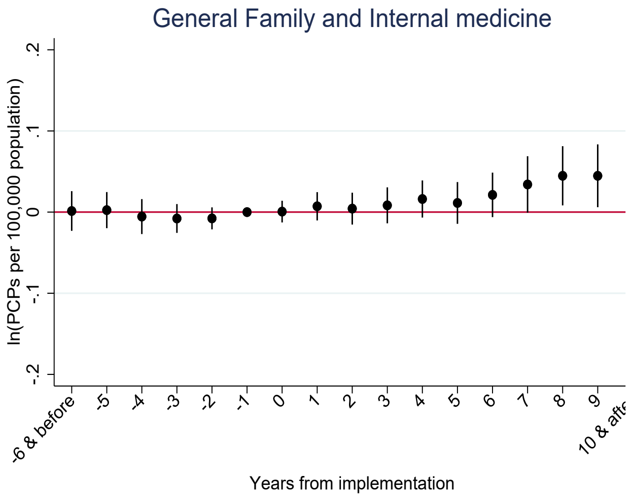

(a) Pre-treatment mean: 47.3



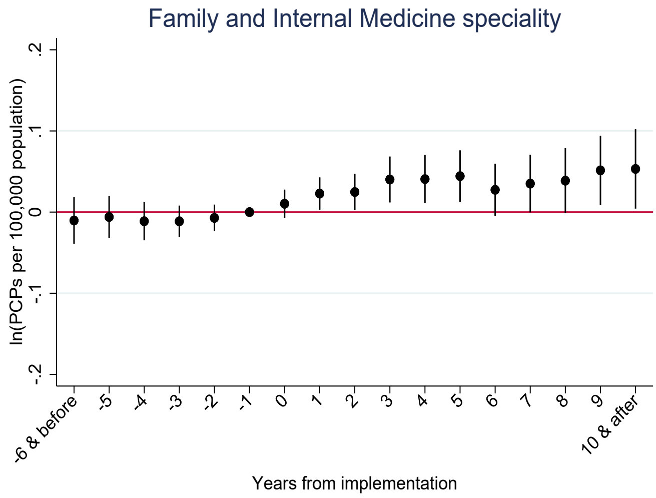

(b) Pre-treatment mean: 7.58

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Pre-treatment mean is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

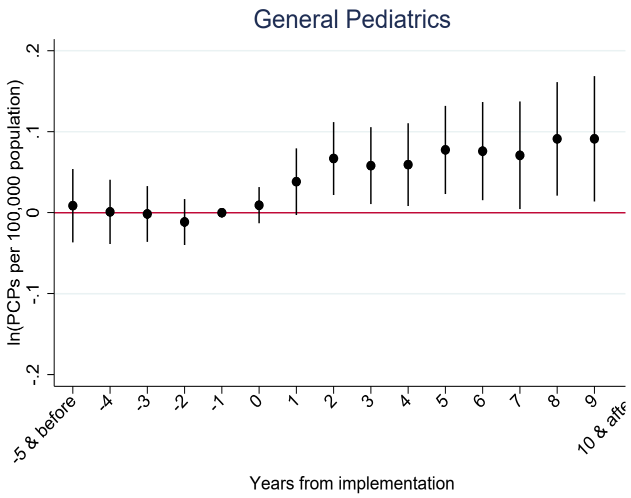
Figure 5: Effects of the policy on entry of eligible practicing physicians by speciality



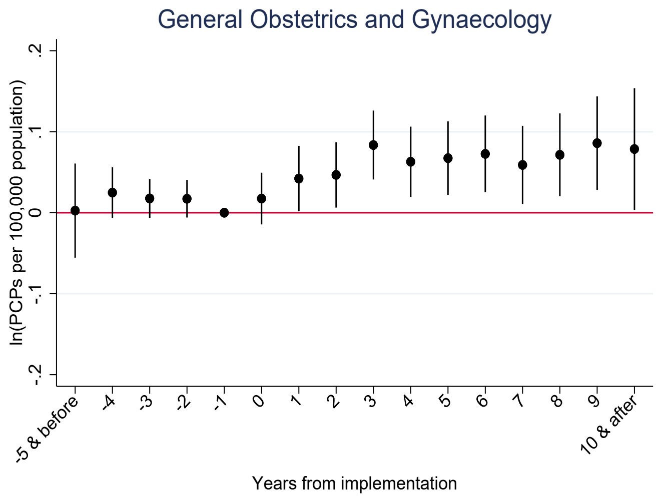

(a) Pre-treatment mean:41.7




(b) Pre-treatment mean:4.5



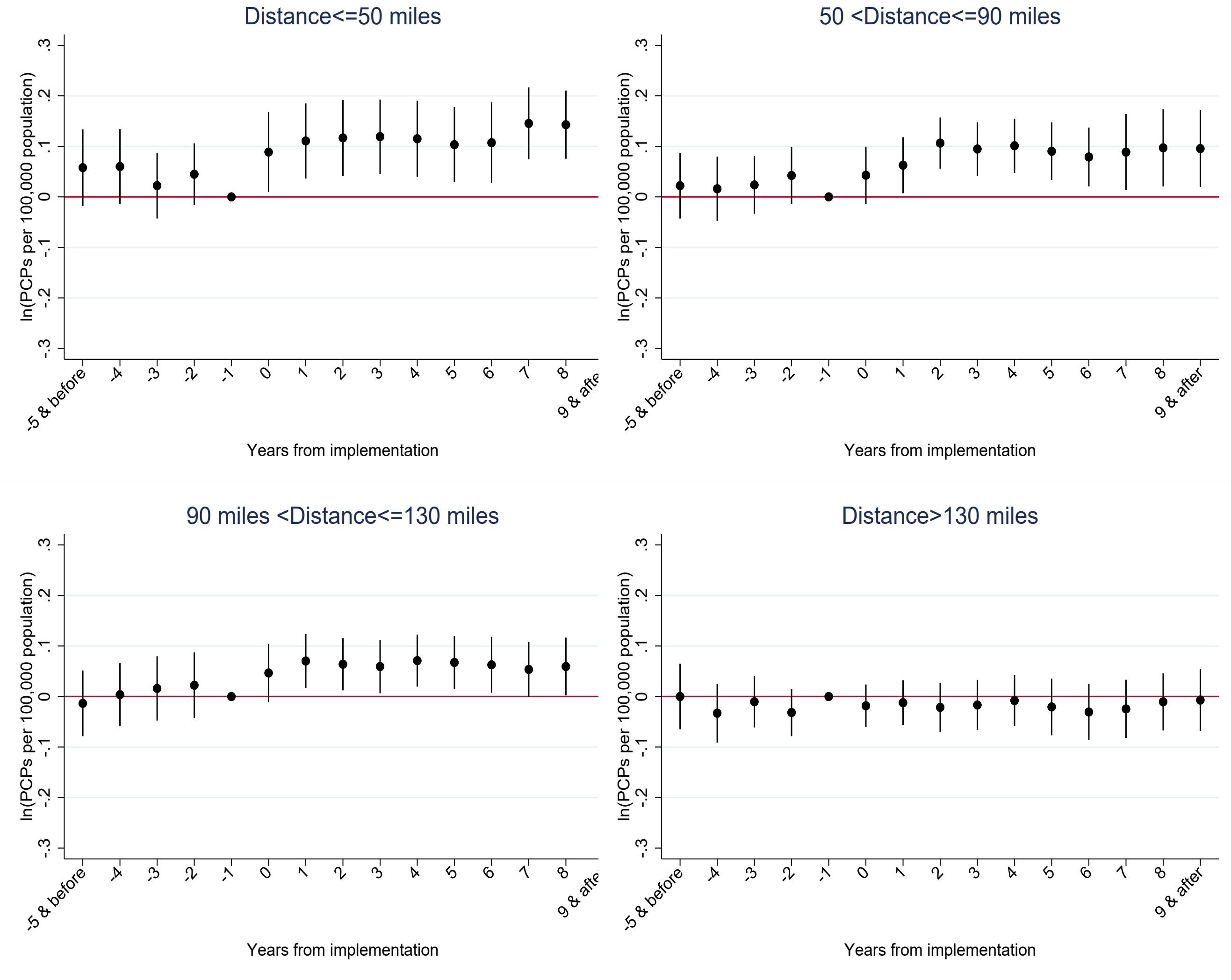

(c) Pre-treatment mean:5.3




(d) Pre-treatment mean:4.5

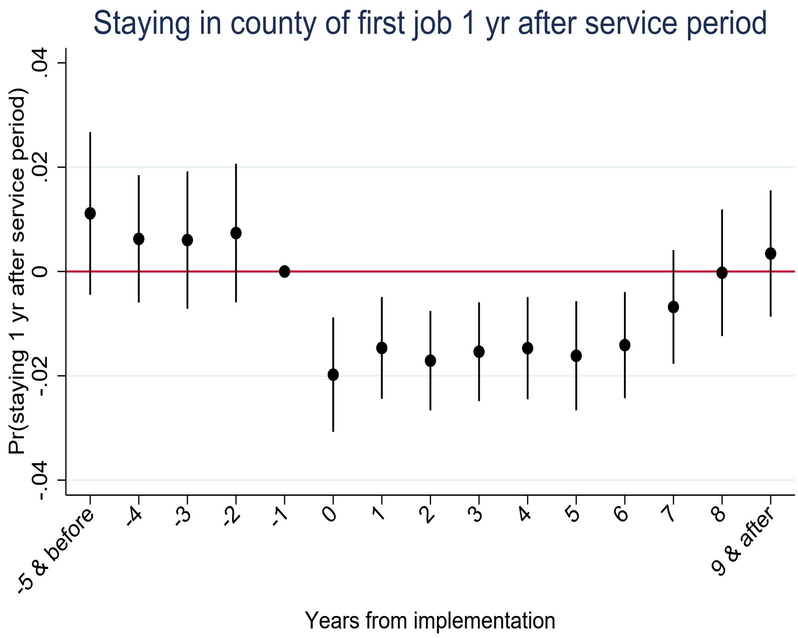
*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log physicians per 100,000 population for the four broad eligible specialities at a county-year level. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Pre-treatment mean is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

Figure 6: Effects of the policy on entry of eligible practicing physicians by distance to nearest largest metro county:distance bins



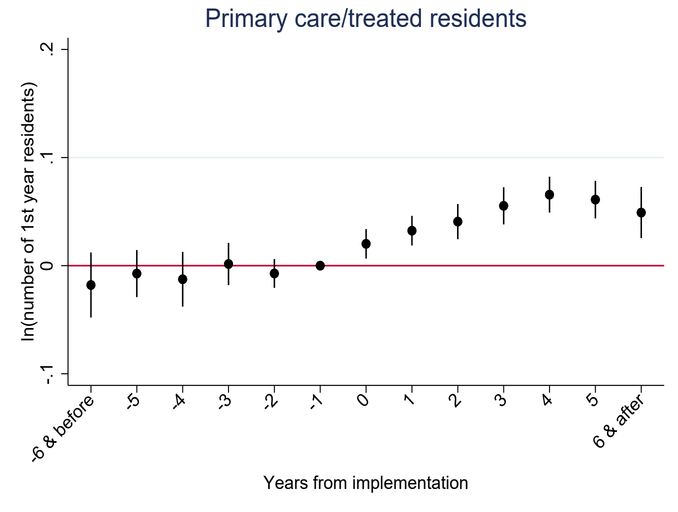

*Notes:* The above figures plot the event study coefficients $\alpha_{1k}, \alpha_{2k}, \alpha_{3k}$ and $\alpha_{4k}$ from equation (2). Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Distance variable represents the distance between a county and nearest largest metro county within the state. See text and Appendix G for the choice of cutoffs and construction of the distance variable. Vertical lines represent 95% confidence intervals. The specification includes county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

Figure 7: Effects of the policy on retention in county of first job one year after the minimum contract period

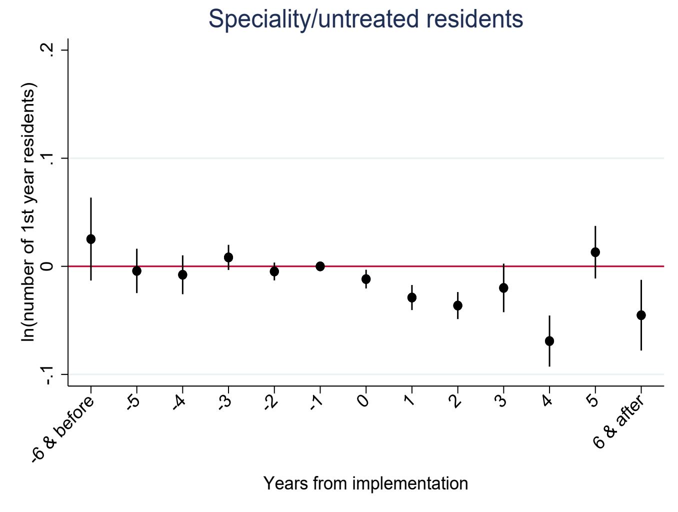



*Notes:* The above figure plots the event study coefficients $\alpha_k$ from equation (3). Outcome variable is likelihood of an eligible physician staying in a treated county they started 1 year after the minimum service period of that state. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: AMA physician masterfile, 1996-2017.

Figure 8: Effects of the policy on choice of specialisation of training physicians: first year residents



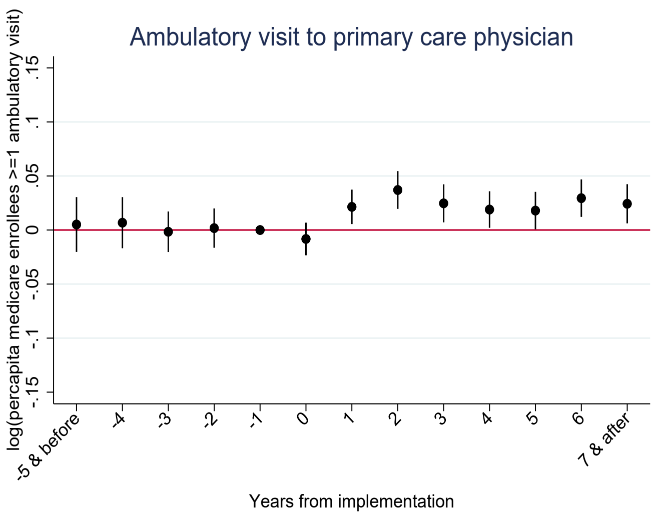

(a) Pre-treatment mean: 396



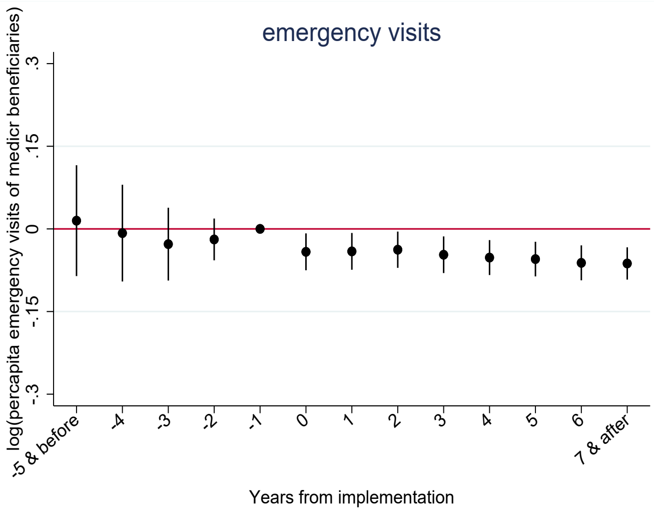

(b) Pre-treatment mean: 595

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (5). Outcome variable is log number of first year residents in treated specialities (panel a) and untreated specialities(panel b) at a state-year level. See Appendix B for definition of treated and untreated specialities. Vertical lines represent 95% confidence intervals. All specifications include state level controls, state fixed effects and year fixed effects. Standard errors are clustered at state level. Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

Figure 9: Effects of the policy on access to physicians and ER visits of Medicare beneficiaries



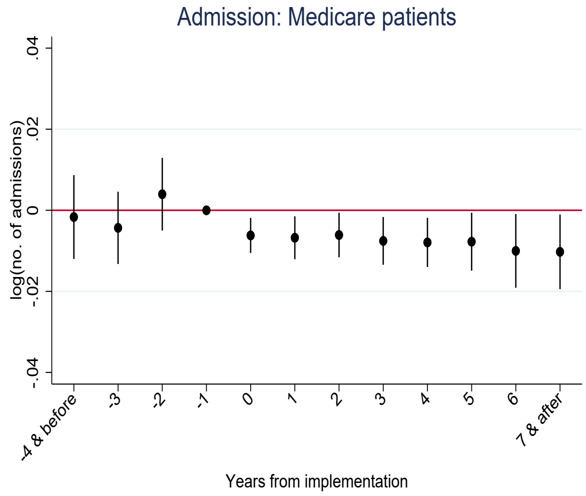

(a) Pre-treatment mean: 0.081



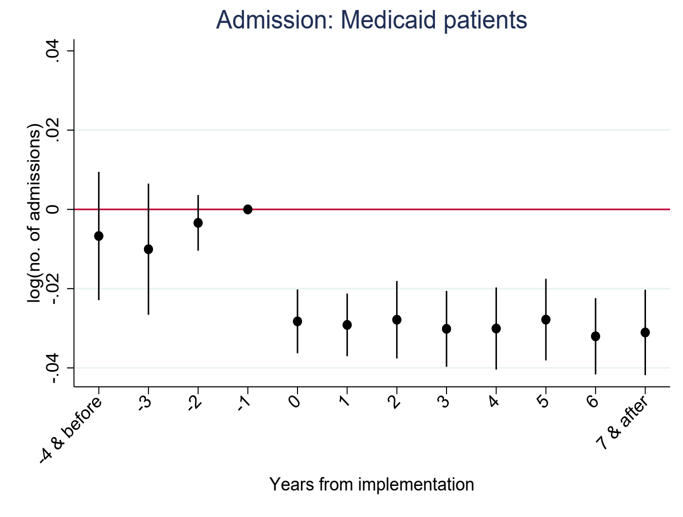

(b) Pre-treatment mean:0.096

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log per capita Medicare enrollees having at least one ambulatory visit to a physician in the top panel and log per capita ER visits of Medicare beneficiaries in the bottom panel. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File, Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017

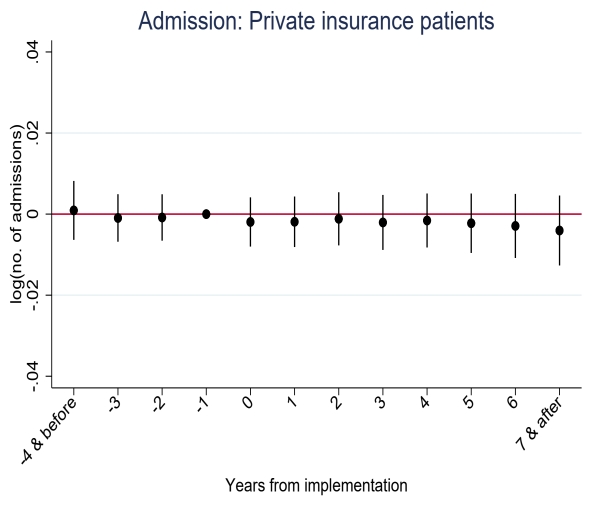
Figure 10: Effects of the policy on hospital admissions in California by payer category: Patient level records



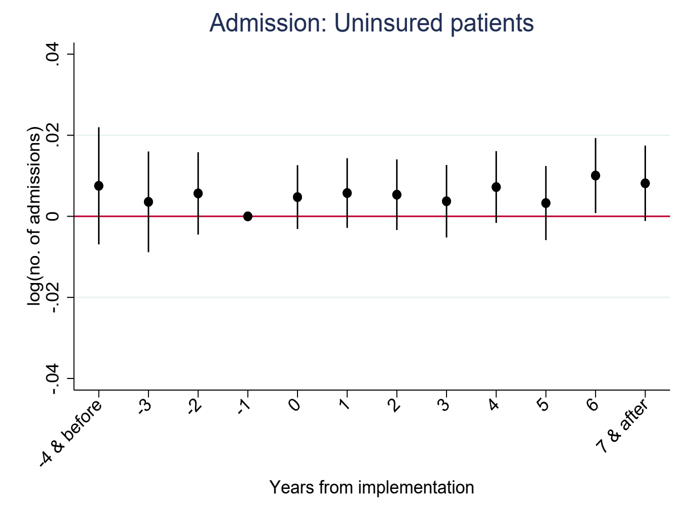

(a) Pre-treatment mean:285.5




(b) Pre-treatment mean:231.5




(c) Pre-treatment mean: 153.5




(d) Pre-treatment mean: 31.75

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (6). Outcome variable is log number of admissions of Medicare patients (panel a), Medicaid patients (panel b), private insurance patients (panel c) and uninsured patients (panel d). Vertical lines represent 95% confidence intervals. All specifications include hospital level controls, county level controls, hospital fixed effects and year fixed effects. Standard errors are clustered at hospital level. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017

Table 1: Baseline means of county level covariates

| Variables | Treated counties Mean (SD) (1) | Untreated counties Mean(SD) (2) | Within state diff (p value) (3) |
|---|---|---|---|
| County level covariates | | | |
| % population aged 25-29 | 5.57(1.23) | 5.83(0.92) | -0.110(0.062) |
| % population aged 30-34 | 5.35(1.10) | 5.45(0.76) | -0.061(0.155) |
| % population aged 35-39 | 7.82(0.87) | 7.83(0.77) | -0.010(0.792) |
| % population aged 40-44 | 7.26(0.67) | 7.36(0.83) | -0.006(0.869) |
| % population aged 45-49 | 6.65(0.74) | 6.60(0.56) | 0.022(0.485) |
| % population aged 50-54 | 5.45(0.52) | 5.44(0.64) | 0.036(0.228) |
| % population aged 55-59 | 4.76(0.63) | 4.72(0.54) | 0.091(0.577) |
| % population aged 60-64 | 4.59(0.82) | 4.55(0.71) | 0.094(0.534) |
| % population aged 65-69 | 4.512(0.89) | 4.506(0.99) | 0.051(0.292) |
| % population aged 70-74 | 4.05(1.03) | 4.10(0.92) | -0.015(0.763) |
| % female | 49.71(2.40) | 49.89(2.47) | -0.385(0.794) |
| % white | 89.33(17.65) | 90.79(13.06) | -1.433(0.036) |
| % hispanic | 5.27(12.80) | 4.29(8.31) | 0.738(0.117) |
| Real Median income (per $10,000) | 2.79(0.60) | 2.83(0.51) | -0.066(0.120) |
| Poverty rate (in %) | 16.39(6.61) | 16.18(5.27) | 0.800(0.288) |
| Unemployment rate (in %) | 5.83 (2.28) | 5.81(2.08) | 0.218(0.141) |
| Uninsured rate (in % & $<$ 65 yrs) | 19.03(6.36) | 18.65(5.63) | 1.133(0.116) |
| No.of Skilled Nursing facilities | 1.96(1.65) | 2.09(1.74) | -0.596(0.144) |
| No.of Home health agency | 0.92(1.01) | 0.87(0.99) | 0.605(0.226) |
| No.of Rural health clinics | 1.46(1.79) | 1.44(1.86) | 0.141(0.121) |
| Federally qualified health centers | 1.13(1.90) | 0.98(1.46) | 0.103(0.124) |
| No.of Community mental health center | 0.05(0.26) | 0.07(0.34) | -0.017(0.196) |
| Advanced practice registered nurses | 7.81(10.89) | 7.79(12.66) | 1.221(0.739) |
| Nurse practitioners | 5.76(9.01) | 5.38(7.23) | 0.326(0.903) |
| Physician assistants | 3.48(6.75) | 4.82(7.32) | -2.952(0.116) |
| Joint Balance test (p value) | | | 0.0367 |
| Number of counties | 786 | 1428 | |

*Note:*Columns (1) and (2) report unweighted means and standard deviations of county level covariates in the pre-treatment period. I consider year before implementation of the policy as the pre-treatment period. Column (3) reports within state differences between treated and untreated counties and the corresponding p values. Sample is restricted to rural counties. The joint balance test row reports the p value from jointly testing whether the covariates in the above panel predict the treatment.

Table 2: Effects of the policy on entry of eligible physicians

Dependent variable: log(physicians per 100,000 population)

| | MDs | | DOs | | MDs | | DOs | |
|---|---|---|---|---|---|---|---|---|
| | (Baseline) | (Controls) | (Baseline) | (Controls) | (Baseline) | (Controls) | (Baseline) | (Controls) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| policy dummy | 0.0508*** | 0.0491*** | 0.0776** | 0.0783** | | | | |
| | (0.0142) | (0.0140) | (0.0333) | (0.0332) | | | | |
| Within 5 years of policy | | | | | 0.0356*** | 0.0354*** | 0.0664*** | 0.0662*** |
| | | | | | (0.0134) | (0.0133) | (0.0217) | (0.0216) |
| > 5 years of policy | | | | | 0.0717*** | 0.0705*** | 0.0897*** | 0.0921*** |
| | | | | | (0.0194) | (0.0192) | (0.0291) | (0.0290) |
| Mean dependent variable | 47.3 | 47.3 | 7.58 | 7.58 | 47.3 | 47.3 | 7.58 | 7.58 |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value for test of equality | | | | | 0.0064 | 0.0073 | 0.2681 | 0.4540 |
| Observations | 48,964 | 48,964 | 48,964 | 48,964 | 48,964 | 48,964 | 48,964 | 48,964 |

Notes: Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. The policy dummy takes the value 1 if county c in state s has implemented the policy at time t and 0 otherwise. Rows 2 and 3 report the coefficients on policy dummy for "$\leq$ 5 years of policy" and "> 5 years of policy" respectively. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated counties. The p values correspond to the test of equality of short run ("within 5 years of policy") and long run("> 5 years of policy") coefficients on policy dummy. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Effects of the policy on access to physicians and hospital admissions of Medicare beneficiaries

| | Access to physicians | | ER visits | | Hospital admissions | |
|---|---|---|---|---|---|---|
| | (Baseline) | (Controls) | (Baseline) | (Controls) | (Baseline) | (Controls) |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| policy dummy | $0.0189^{**}$ | $0.0206^{***}$ | $-0.0485^{***}$ | $-0.0498^{***}$ | $-0.1026^{***}$ | $-0.1015^{***}$ |
| | (0.0083) | (0.0081) | (0.0170) | (0.0169) | (0.0354) | (0.0353) |
| Mean dependent variable | 0.081 | 0.081 | 0.096 | 0.096 | 80.19 | 80.19 |
| Controls | No | Yes | No | Yes | No | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 38,320 | 38,320 | 38,320 | 38,320 | 38,320 | 38,320 |

Notes: Outcome variable is log per capita Medicare enrollees having at least one ambulatory visit to physician in cols(1) and (2), log per capita ER visits of Medicare beneficiaries in cols(3) and (4) and log preventable hospital stays rate of Medicare beneficiaries in cols(5) and (6) at a county-year level. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File, Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Effects of the policy on hospital admissions in California by payer category: Patient level records

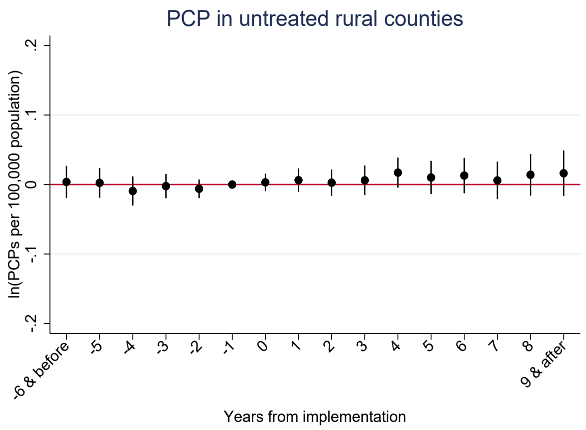
| | (Medicare) | (Medicaid) | (Private insurance) | (Uninsured) | (Unscheduled) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| policy dummy | $-0.0077^{**}$ | $-0.0292^{***}$ | -0.0021 | 0.0059 | $-0.0058^{**}$ |
| | (0.0034) | (0.0040) | (0.0034) | (0.0044) | (0.0024) |
| Mean dependent variable | 285.5 | 231.5 | 153.5 | 31.75 | 582.5 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Hospital FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 5548 | 5548 | 5548 | 5548 | 5548 |

Notes: Outcome variable is log number of admissions of Medicare patients (col 1), Medicaid patients (col 2), private insurance patients (col 3), uninsured patients (col 4) and log number of unscheduled admissions (col 5) at a hospital-year level. All specifications include hospital and year fixed effects. Standard errors are clustered at hospital level and reported in parenthesis. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017 .

*** p<0.01, ** p<0.05, * p<0.1

## For Online Publication
## Appendix Figures and Tables

Figure A1: Are there treatment spillovers to untreated rural counties?




*Notes:* The above figure plots the event study coefficients on the rural untreated counties in an equation that includes rural treated, rural untreated and propensity score selected similar urban untreated counties.Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level.

Figure A2: Distribution of treated counties relative to nearest largest metro county

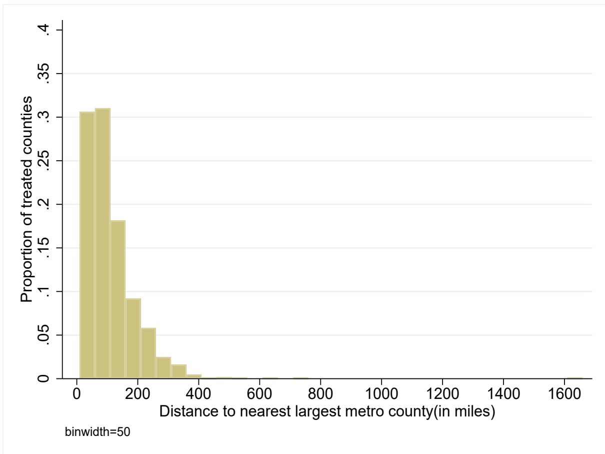
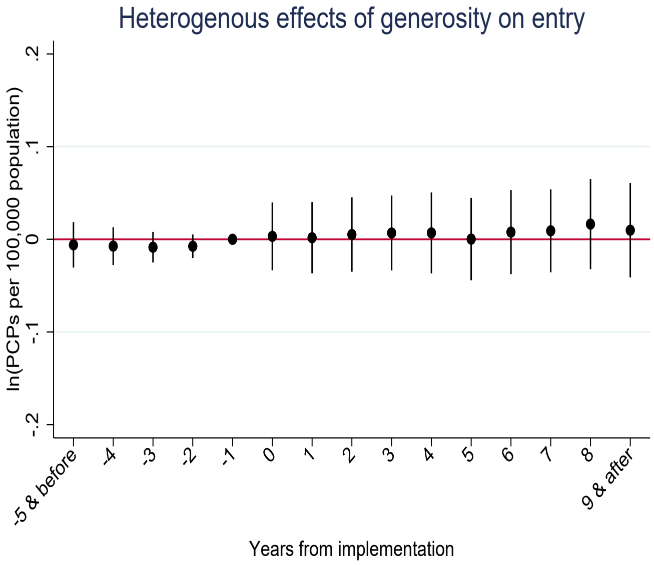
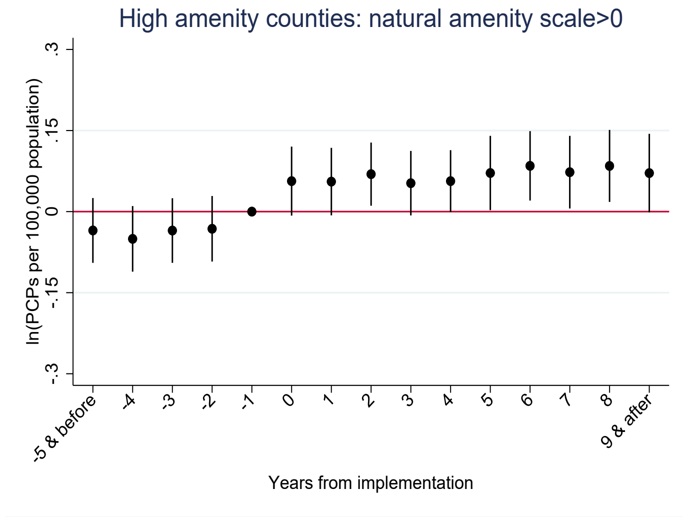


Figure A3: Effects of the policy on entry of eligible practicing physicians by size of benefits

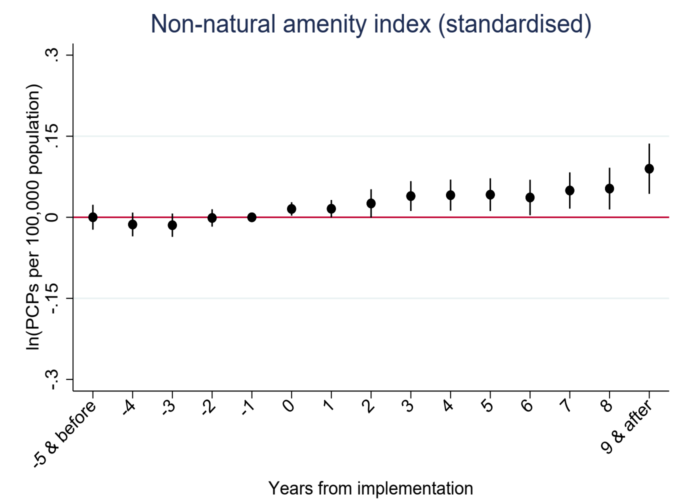



*Notes:* The above figure plots the event study coefficients of the interaction term: indicator for policy dummy $D_c$, the indicator function $\mathbb{I}(t = D_c + k)$, and generosity of the policy. Outcome variable is log physicians per 100,000 population at a county-year level. physicians include both MDs and DOs. The generosity of the policy in county c year t is defined as log of total benefit amount in (c,t) cell divided by minimum service period in that cell. Vertical lines represent 95% confidence intervals. The specification includes county level controls, county and state-by-year fixed effects. Standard errors are clustered at the county level. Pre-treatment mean is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

Figure A4: Effects of the policy on entry of eligible practicing physicians: Heterogeneity by natural and non-natural amenities



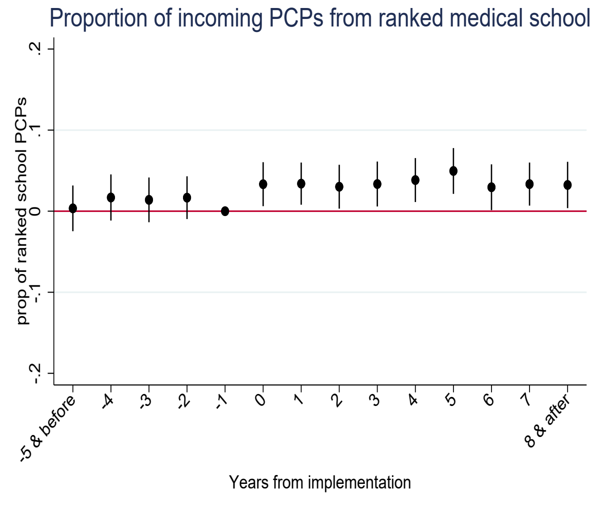

(a)



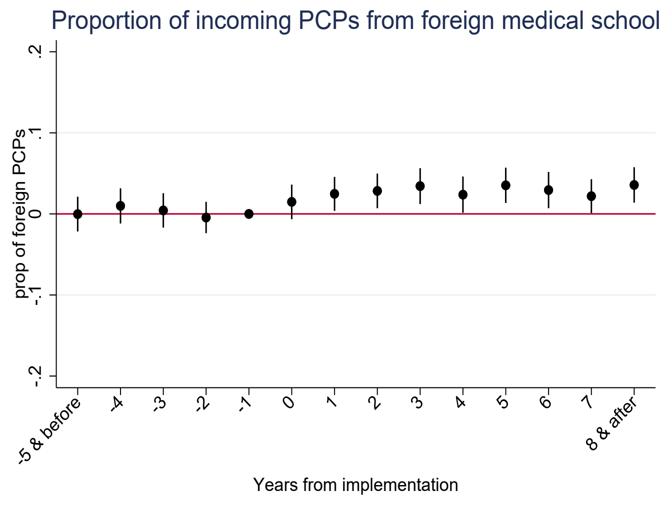

(b)

*Notes:* The above figures plot the event study coefficients of the interaction terms: indicator for policy dummy $D_c$, the indicator function $\mathbb{I}(t = D_c + k)$, and indicator for high amenity counties in panel(a); indicator for policy dummy $D_c$, the indicator function $\mathbb{I}(t = D_c + k)$, and standardised non-natural amenity index in 1994 in panel (b). Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

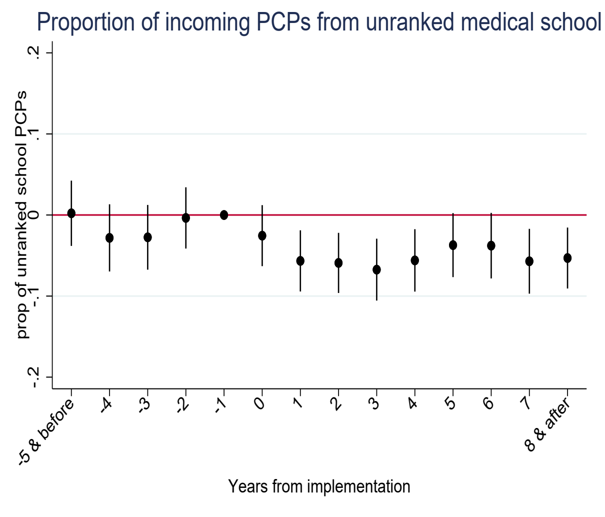
Figure A5: Who are the compliers of the policy?



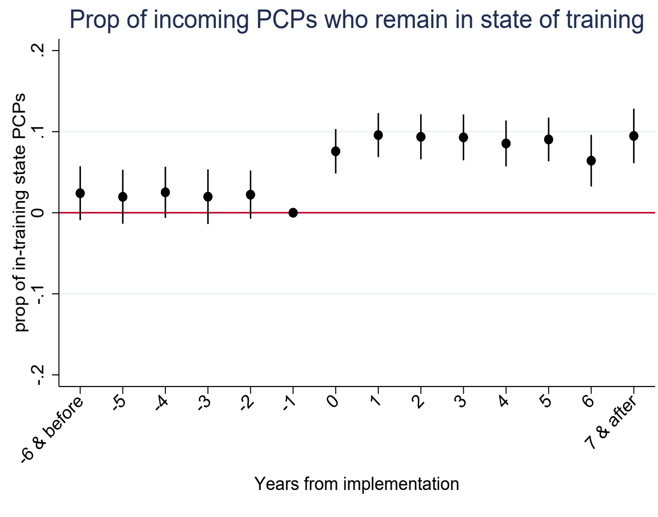

(a) Pre-treatment mean: 0.162




(b) Pre-treatment mean: 0.076



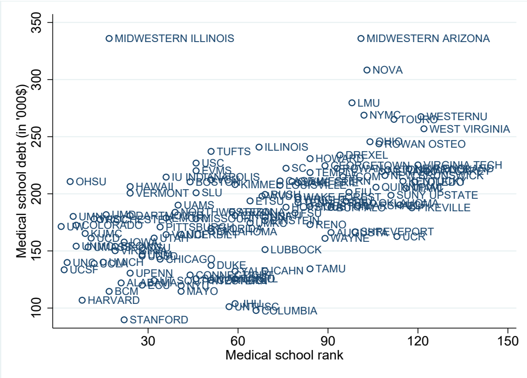

(c) Pre-treatment mean: 0.759




(d) Pre-treatment mean: 0.794

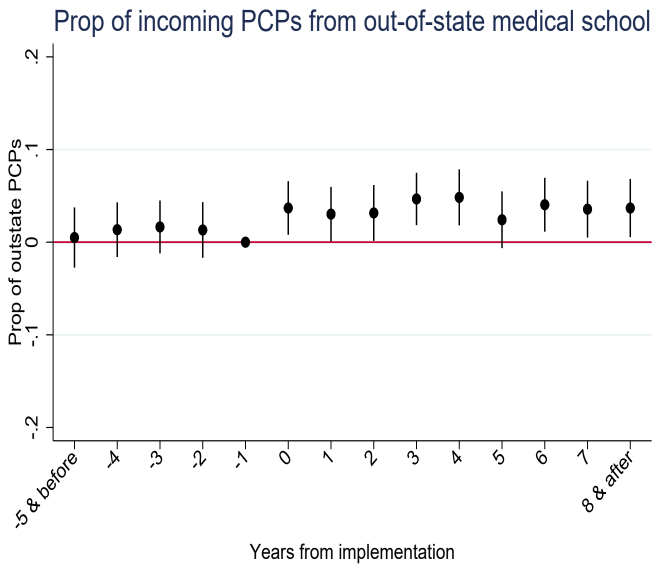
*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is proportion of incoming physicians from ranked US medical school (panel a), proportion of incoming physicians from foreign medical school (panel b) , proportion of incoming physicians from unranked US medical school (panel c) and proportion of incoming physicians who start their job in their state of training (panel d). All outcome variables are aggregated to county-year level using individual longitudinal data. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: AMA physician masterfile, 1996-2017.

Figure A6: Relationship between medical school rank and medical school debt



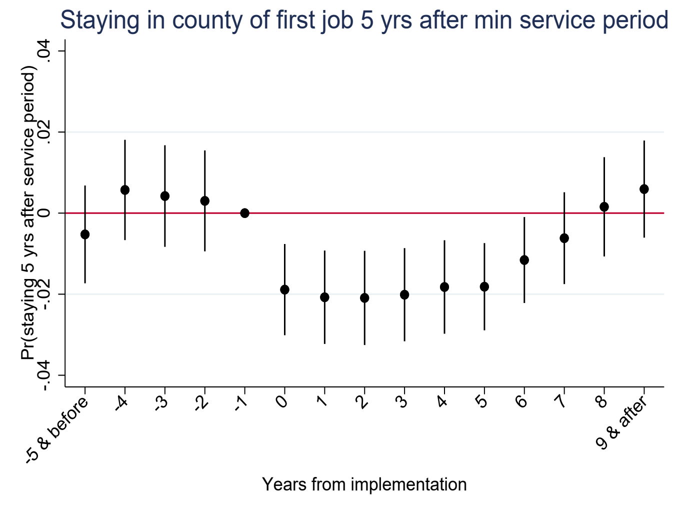

*Notes:* Main data source: US News and World Report primary care ranking, 2019

Figure A7: Who are the compliers of the policy: In-state versus out-of-state medical school

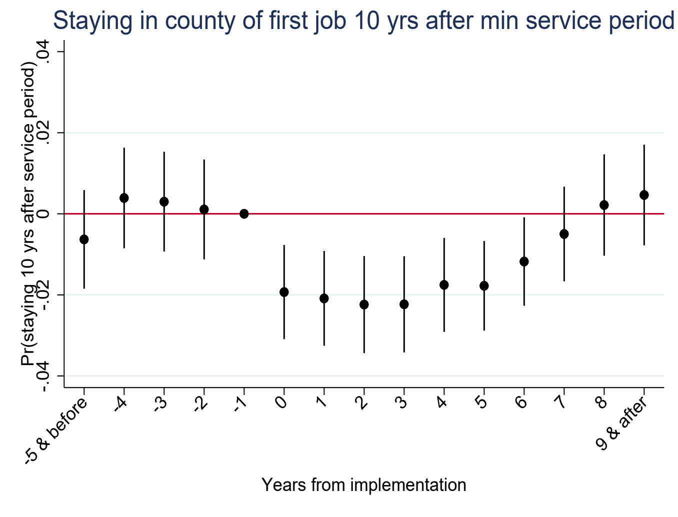



*Notes:* The above figure plots the event study coefficients $\alpha_k$ from equation (1). Outcome variable is proportion of incoming physicians from out-of-state medical school. Out-of-state medical school includes foreign medical school. All outcome variables are aggregated to county-year level using individual longitudinal data. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: AMA physician masterfile, 1996-2017.

Figure A8: Effects of the policy on retention in county of first job five years and ten years after service period




(a)



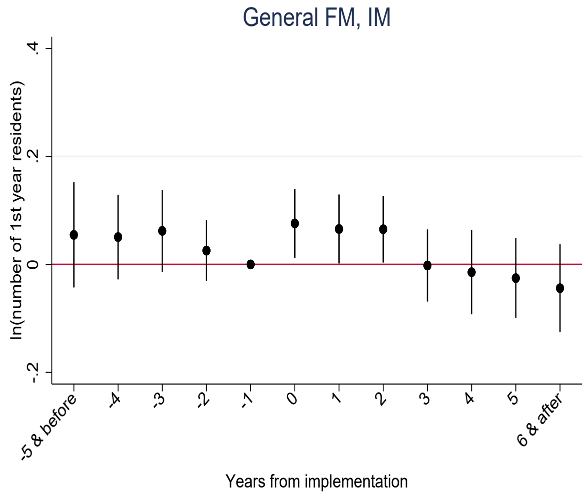

(b)

*Notes:* The above figure plots the event study coefficients $\alpha_k$ from equation (3). Outcome variable is the likelihood of an eligible physician staying in a treated county they started 5 years and 10 years respectively, after the service period of that state. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: AMA physician masterfile, 1996-2017.

Figure A9: Is the increase in number of matched residents in treated specialities driven by slot increases?

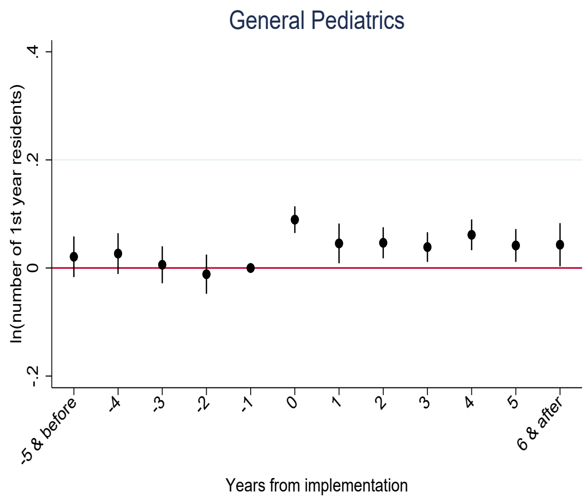

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (5). Outcome variable is log number of treated speciality slots that programs are willing to offer at a state-year level. Vertical lines represent 95% confidence intervals. All specifications include state level controls, state fixed effects and year fixed effects. Standard errors are clustered at county level. Main data source: National Resident Matching Program, AY 1994-95 to 2018-19.

Figure A10: Effects of the policy on speciality choice of first year training physicians: Heterogenous effects

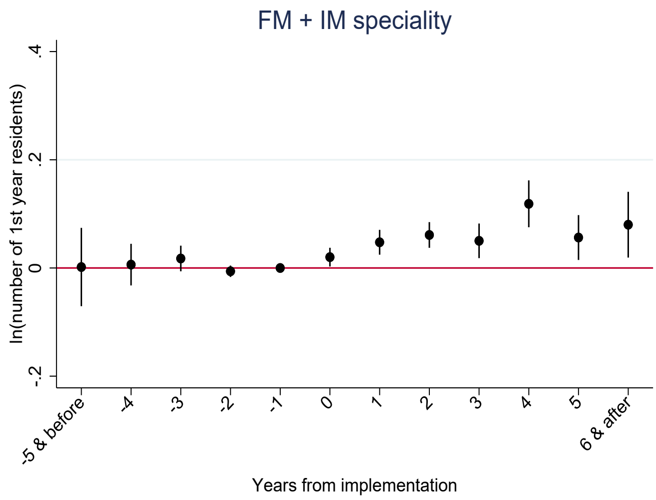



(a) Pre-treatment mean: 231


(b) Pre-treatment mean: 25



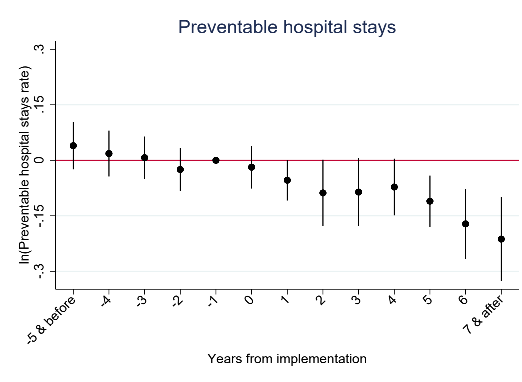

(c) Pre-treatment mean:61




(d) Pre-treatment mean:79

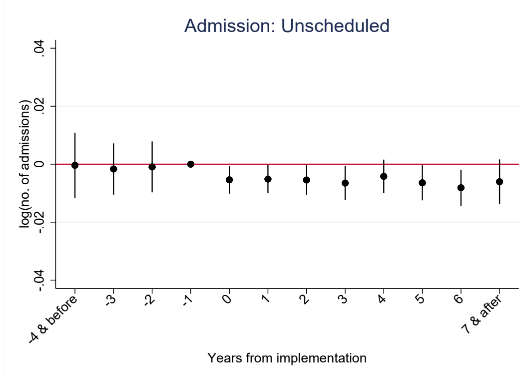
*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (5). Outcome variable is log number of first year residents for the four broad eligible specialities at a state-year level. Vertical lines represent 95% confidence intervals. All specifications include state level controls, state fixed effects and year fixed effects. Standard errors are clustered at county level. Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

Figure A11: Effects of the policy on preventable hospital stays




Pre-treatment mean:80.19

Figure A12: Effects of the policy on unscheduled hospital admissions in California: patient level records

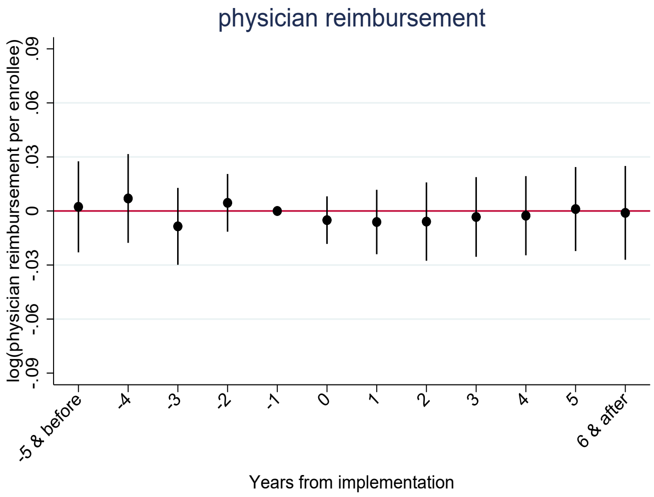



Pre-treatment mean: 582.5

*Notes:* Figure A11 plots the event study coefficients $\alpha_k$ from equation (1), and figure A12 plots $\alpha_k$ from equation (6). In the top figure, the outcome variable is log preventable hospital stays rate of Medicare beneficiaries, while in the bottom figure, the outcome is log number of unscheduled or emergency admissions. Main data source: Area Health Resource File, Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017 in figure A11. Main data source: OSHPD patient discharge data at the day level, 1999-2017 in figure A12.

Figure A13: Effects of the policy on hospital reimbursement and physician reimbursement for Medicare enrollees

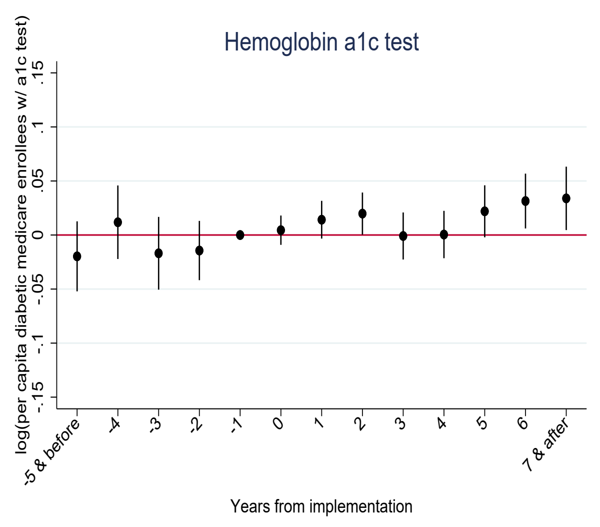

(a) Pre-treatment mean: 4153.35



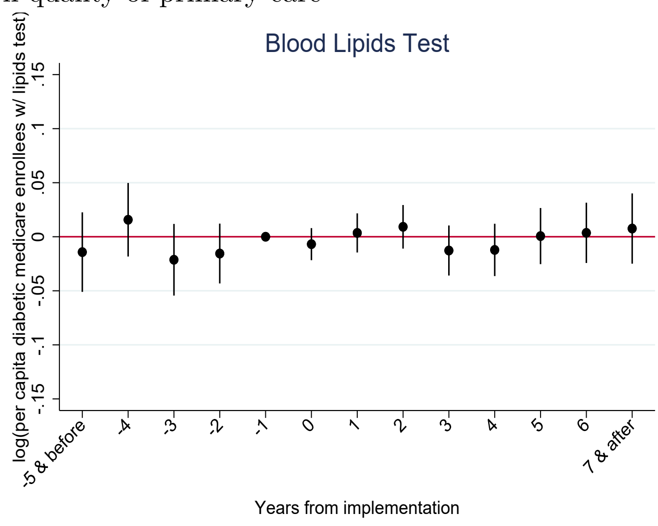

(b) Pre-treatment mean: 1846.24

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log hospital reimbursement per enrollee in the top panel and log physician reimbursement per enrollee in the bottom panel. Both hospital and physician reimbursements are price, age, race and sex adjusted. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017

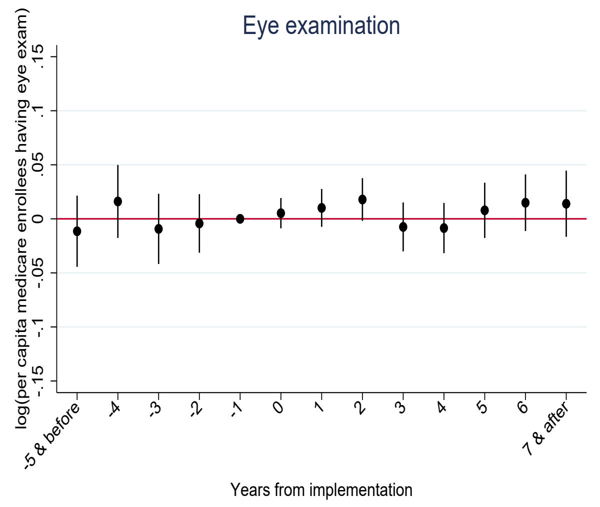
Figure A14: Effects of the policy on quality of primary care



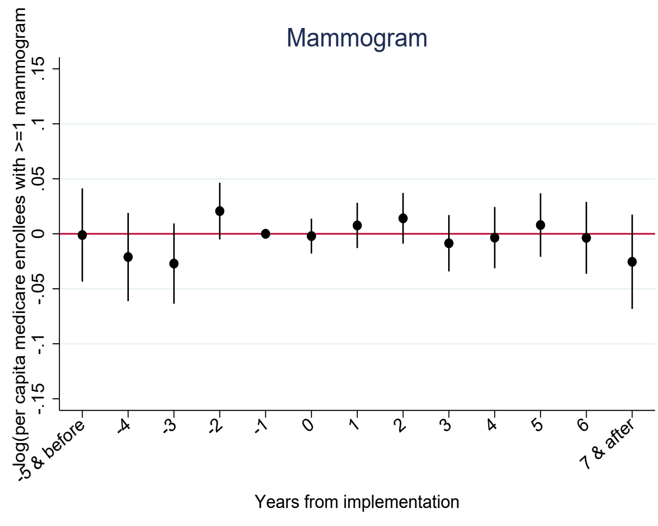

(a) Pre-treatment mean:0.0058




(b) Pre-treatment mean:0.0052



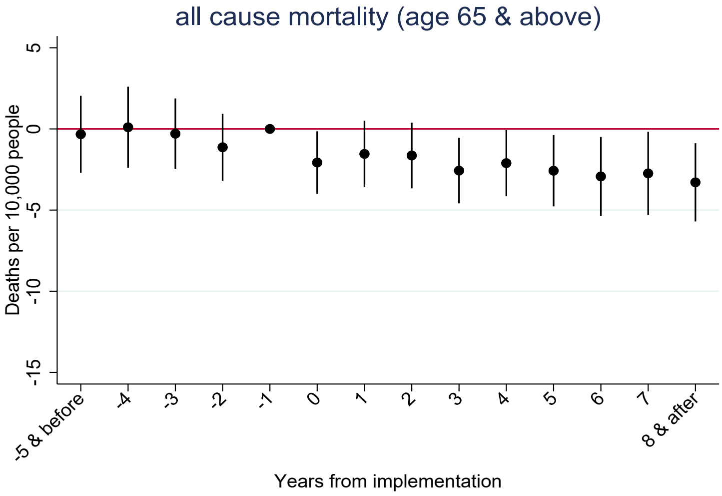

(c) Pre-treatment mean:0.0048




(d) Pre-treatment mean:0.0041

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log per capita diabetic medicare enrollees having a1c test (panel a), log per capita diabetic medicare enrollees having blood lipids test (panel b), log per capita medicare enrollees having eye examination (panel c), log per capita female medicare enrollees having at least one mammogram (panel d). Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File, Dartmouth Atlas of Healthcare, 2000-2017

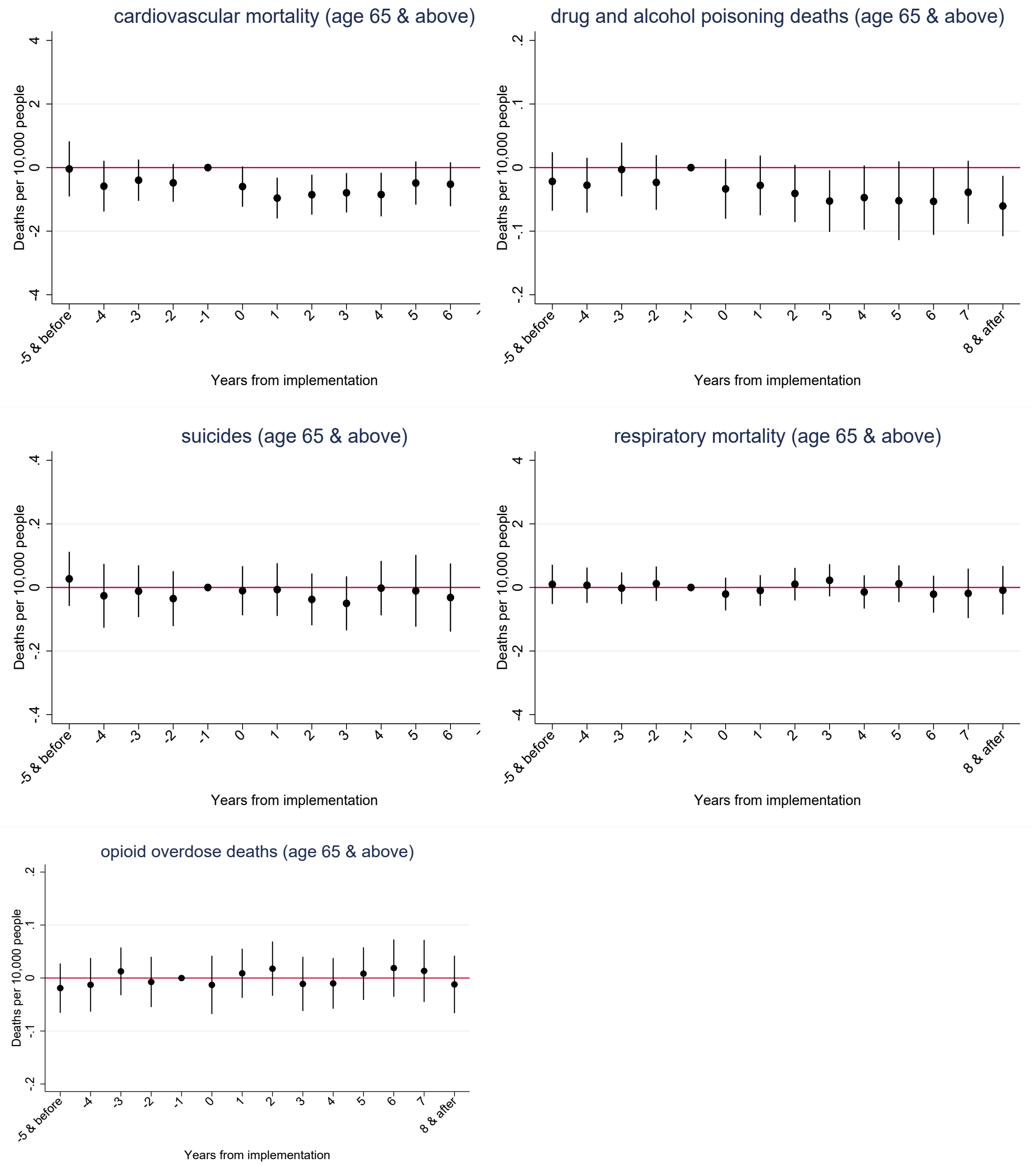
Figure A15: Effects of the policy on mortality of the elderly population





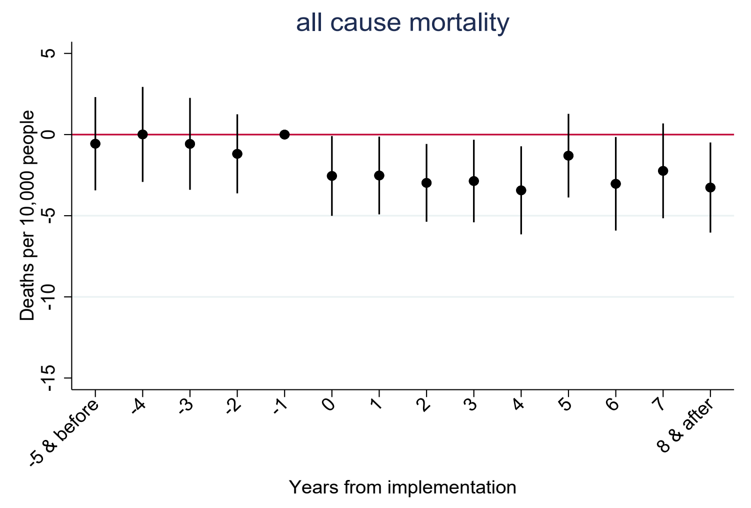
*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is number of deaths of elderly over 65 years per 10,000 population , for all cause mortality (panel a) and amenable cause mortality (panel b). The list of amenable health conditions included in the paper are reported in Table A17. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017

Figure A16: Effects of the policy on mortality of the elderly population by cause of death

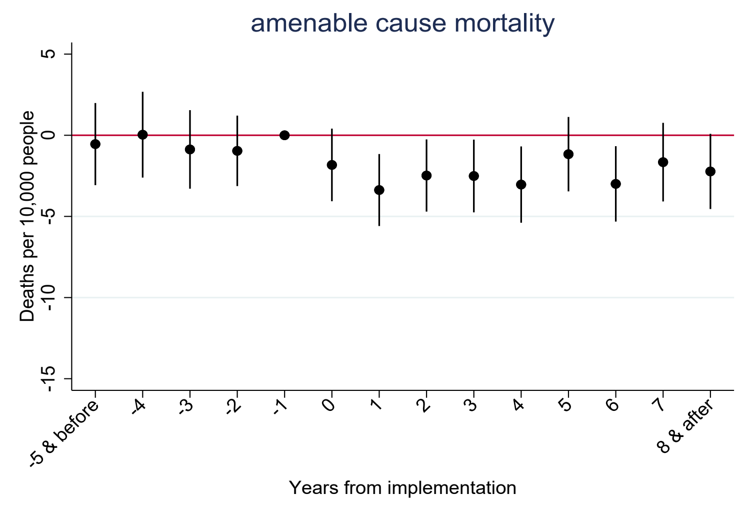



*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is number of deaths of elderly over 65 years per 10,000 population, for cardiovascular mortality (panel a) respiratory mortality (panel b), suicides (panel c), drug and alcohol poisoning (panel d), opioid overdose mortality (panel e). Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017

Figure A17: Effects of the policy on mortality of adult population



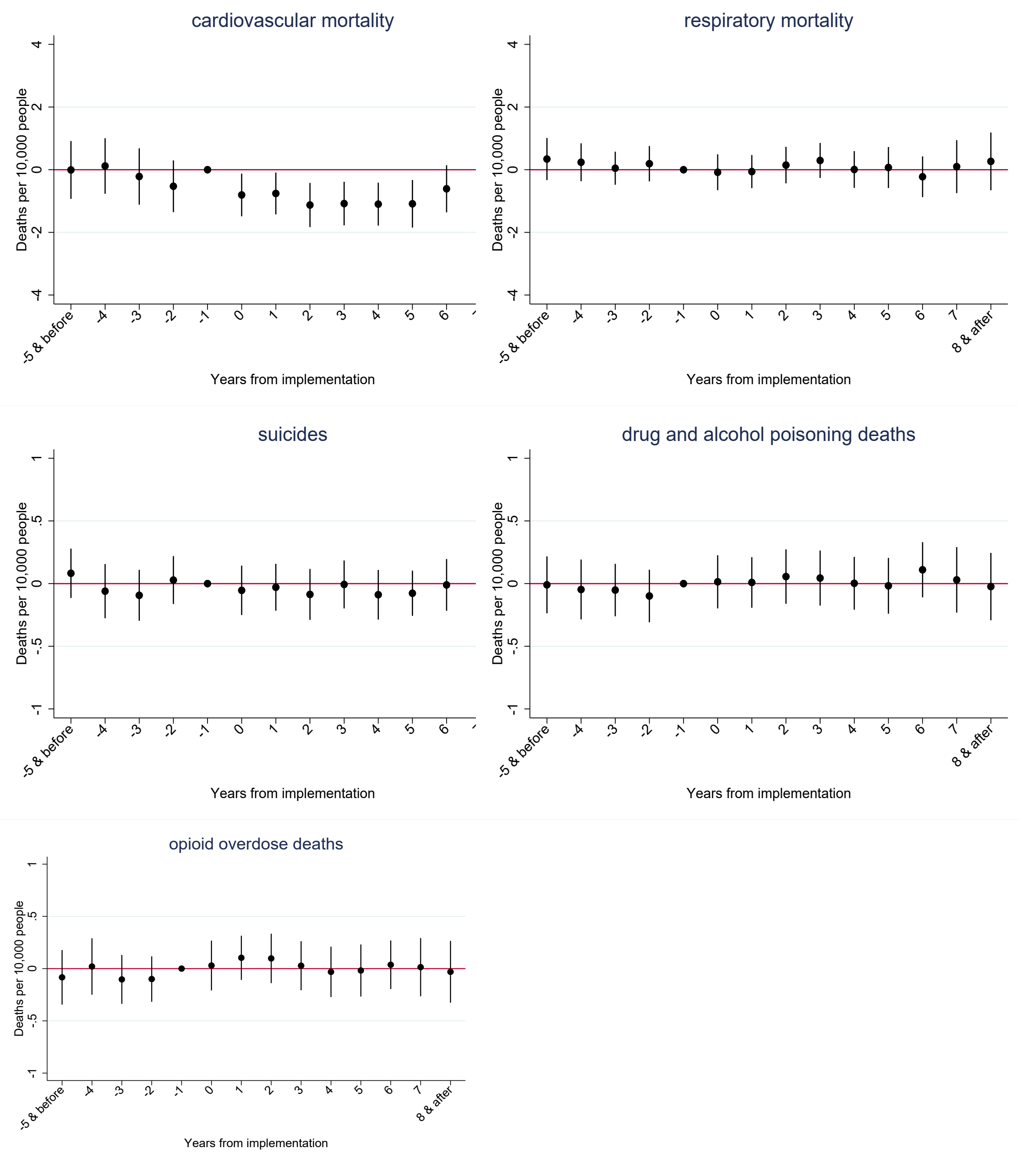

(a) Pre-treatment mean:




(b) Pre-treatment mean:

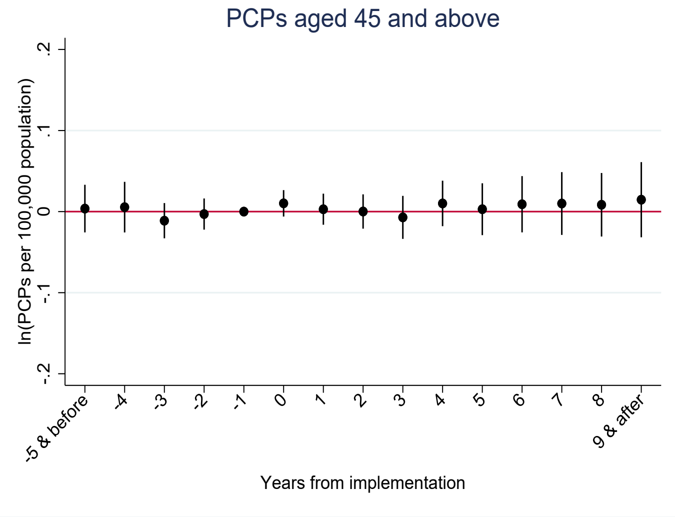
*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is number of deaths of all adults over 20 years per 10,000 population, for all cause mortality (panel a) and amenable cause mortality (panel b). The list of amenable health conditions included in the paper is reported in Table A17. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017

Figure A18: Effects of the policy on mortality of adult population by cause of death

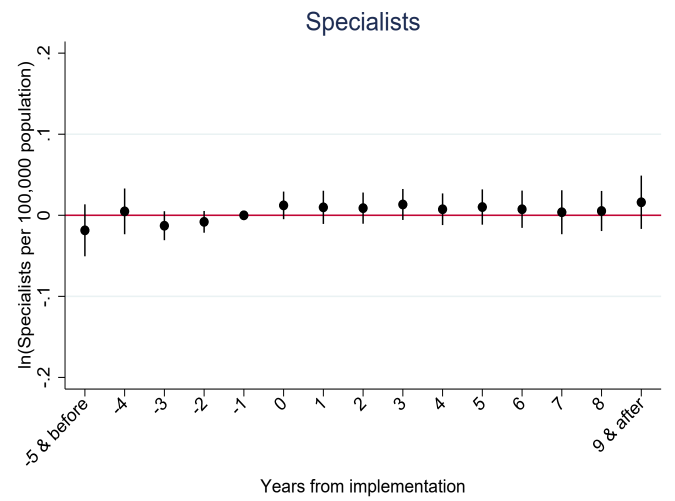



*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is number of deaths of all adults over 20 years per 10,000 population, for cardiovascular mortality (panel a) respiratory mortality (panel b), suicides (panel c), drug and alcohol poisoning (panel d), opioid overdose mortality (panel e). Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017

Figure A19: Effects of the policy on entry of ineligible practicing physicians and specialists: Falsification test




(a) Pre-treatment mean: 30.05




(b) Pre-treatment mean: 24.2

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log physicians above 45 years per 100,000 population in panel (a) and log specialists per 100,000 population in panel (b) at a county-year level. physicians and specialists include both MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Pre-treatment mean is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

Figure A20: Entry of federal PCPs- placebo analysis

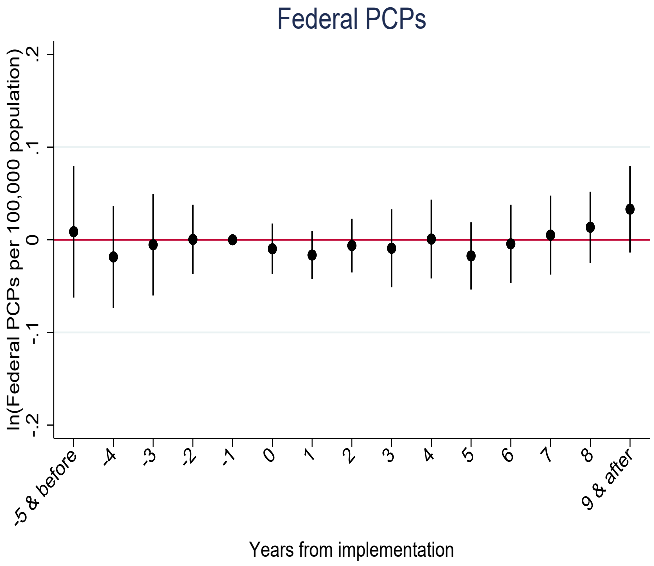
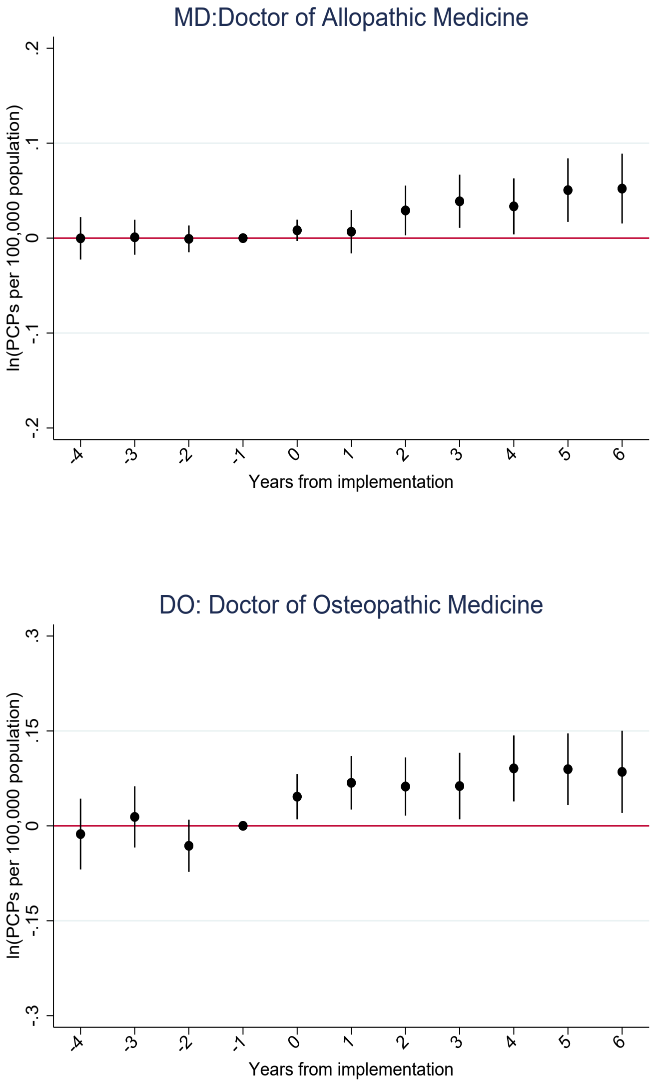
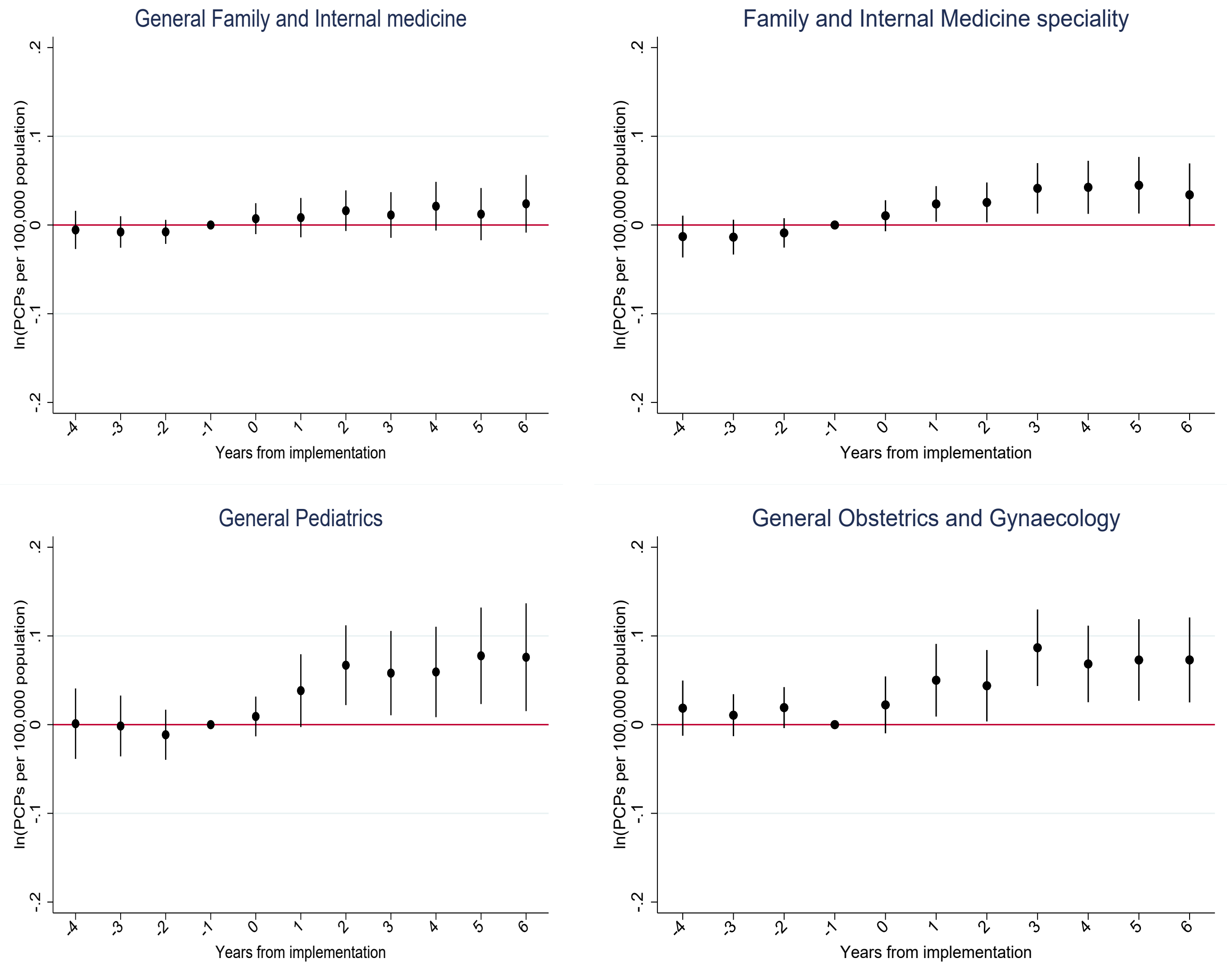


Figure A21: Effects of the policy on entry of eligible practicing physicians: Balanced Panel



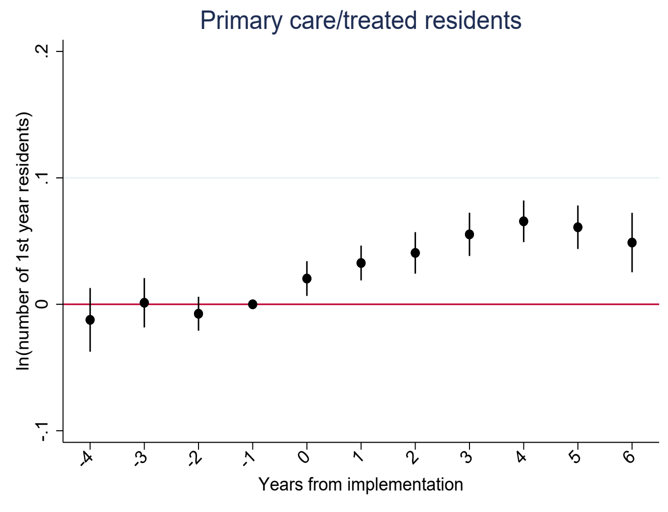

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. The event window ranges from -4 to +6, dropping unbalanced states (states with less than 4 years of pre-periods or 6 years of post-periods) and unbalanced years (event years outside the event window). Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

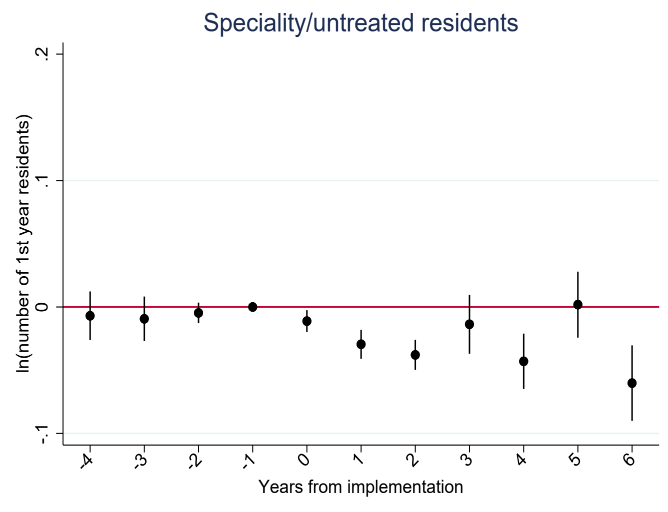
Figure A22: Effects of the policy on entry of eligible practicing physicians by speciality-Balanced panel




*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log physicians per 100,000 population for the four broad eligible specialities at a county-year level. Physicians consist of MDs and DOs. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. The event window ranges from -4 to +6, dropping unbalanced states (states with less than 4 years of pre-periods or 6 years of post-periods) and unbalanced years (event years outside the event window). Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

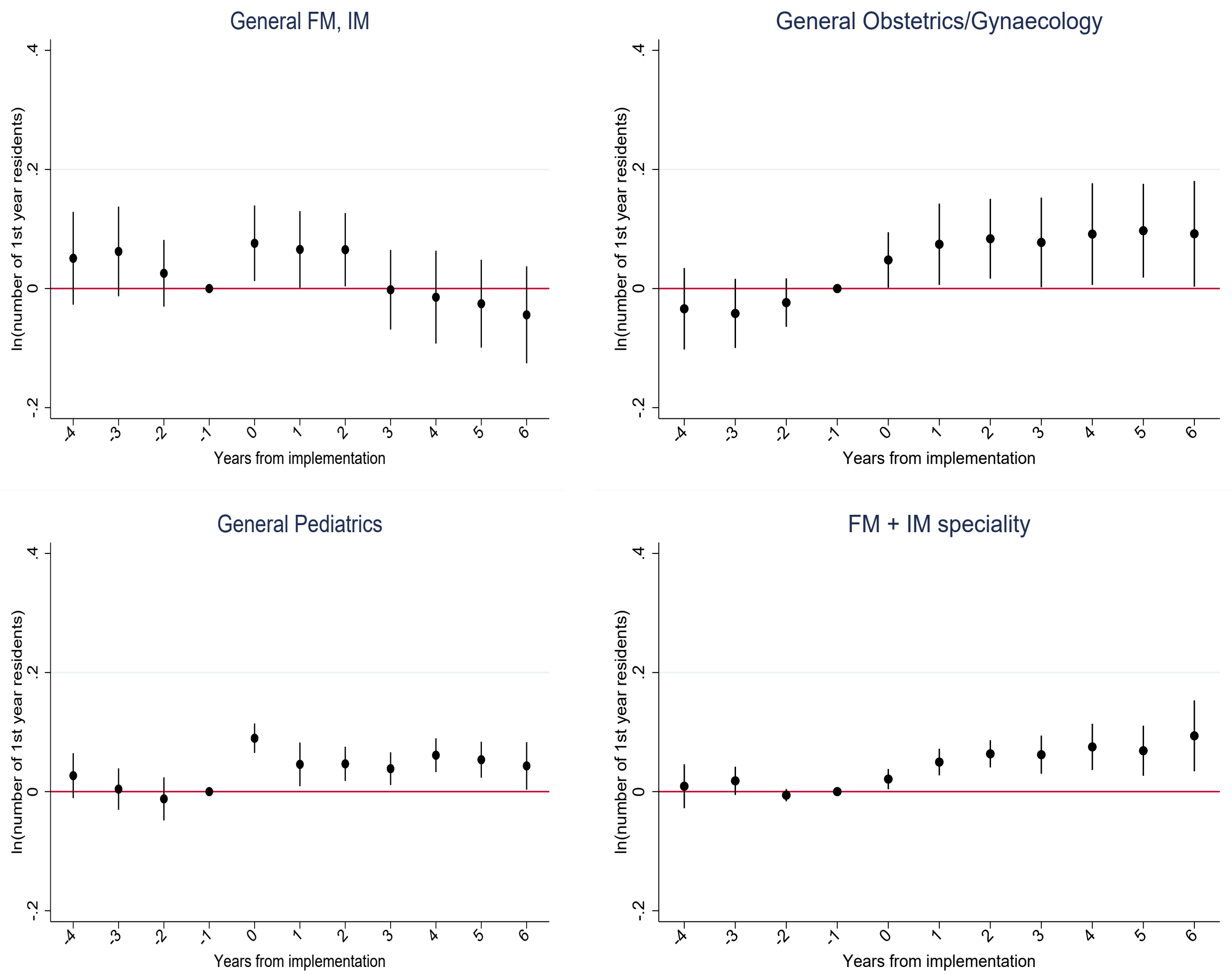
Figure A23: Effects of the policy on choice of specialisation of first year training physicians: Balanced panel



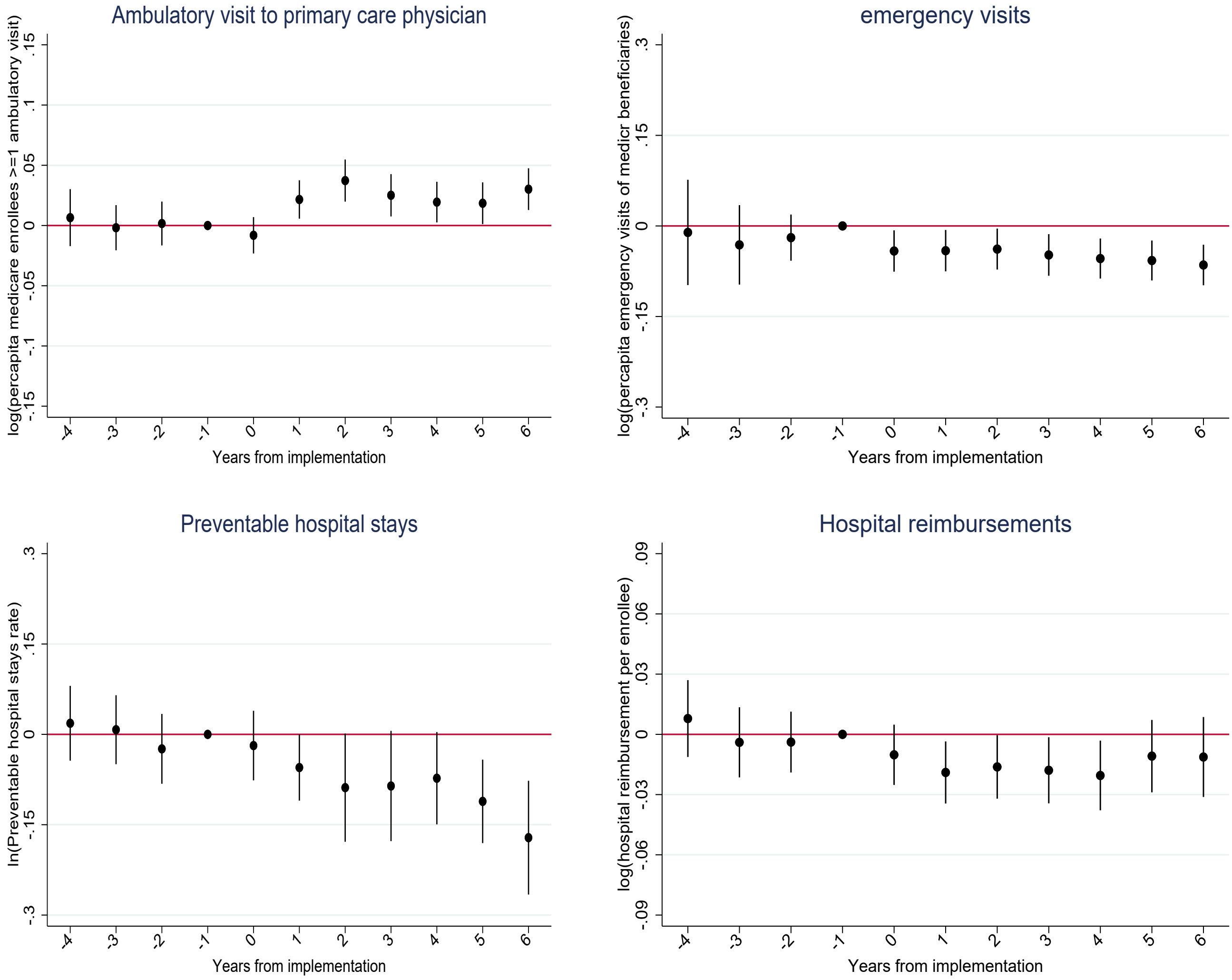

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (5). Outcome variable is log number of first year residents in treated specialities (panel a) and untreated specialities(panel b) at a state-year level. Vertical lines represent 95% confidence intervals. All specifications include state level controls, state fixed effects and year fixed effects. Standard errors are clustered at county level.The event window ranges from -4 to +6, dropping unbalanced states (states with less than 4 years of pre-periods or 6 years of post-periods) and unbalanced years (event years outside the event window). Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

Figure A24: Heterogenous effects of the policy on speciality choice of first year training physicians: Balanced panel



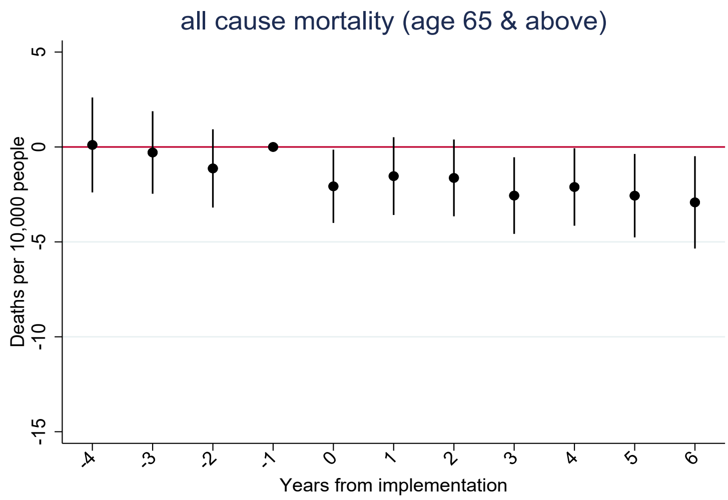

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (5). Outcome variable is log number of first year residents for the four broad eligible specialities at a state-year level. Vertical lines represent 95% confidence intervals. All specifications include state level controls, state fixed effects and year fixed effects. Standard errors are clustered at county level. The event window ranges from -4 to +6, dropping unbalanced states (states with less than 4 years of pre-periods or 6 years of post-periods) and unbalanced years (event years outside the event window). Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

Figure A25: Effects of the policy on access to physicians and adverse events for Medicare beneficiaries: Balanced panel




*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is log per capita Medicare enrollees having at least one ambulatory visit to a physician in panel a, log per capita ER visits of Medicare beneficiaries in panel b, log preventable hospital stays rate of Medicare beneficiaries in panel c, log hospital reimbursement per Medicare enrollee in panel d. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. The event window ranges from -4 to +6, dropping unbalanced states (states with less than 4 years of pre-periods or 6 years of post-periods) and unbalanced years (event years outside the event window). Main data source: Area Health Resource File, Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017

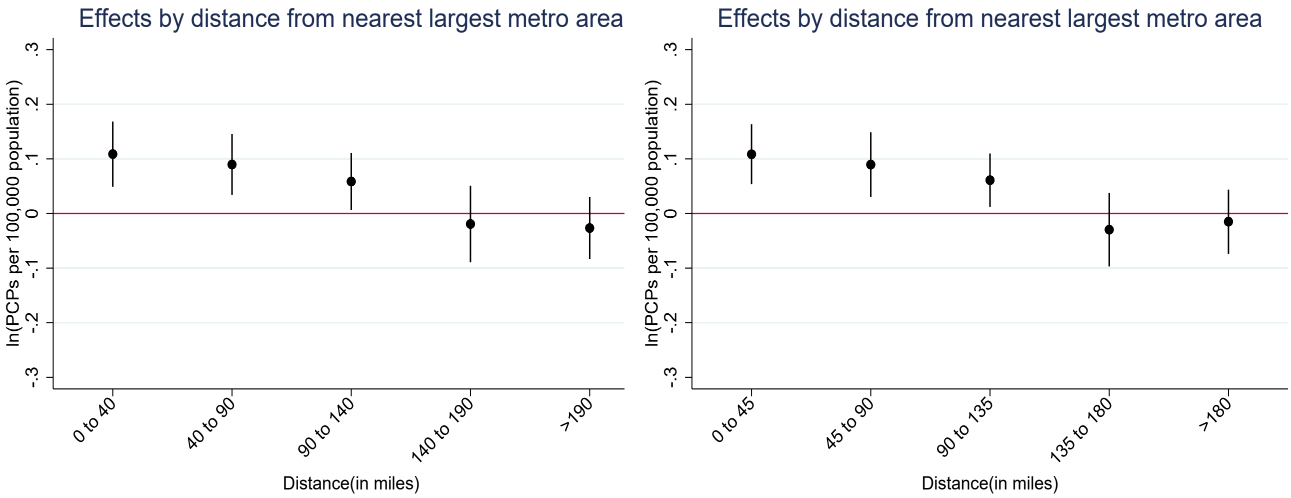
Figure A26: Effects of the policy on mortality of the elderly population-Balanced panel





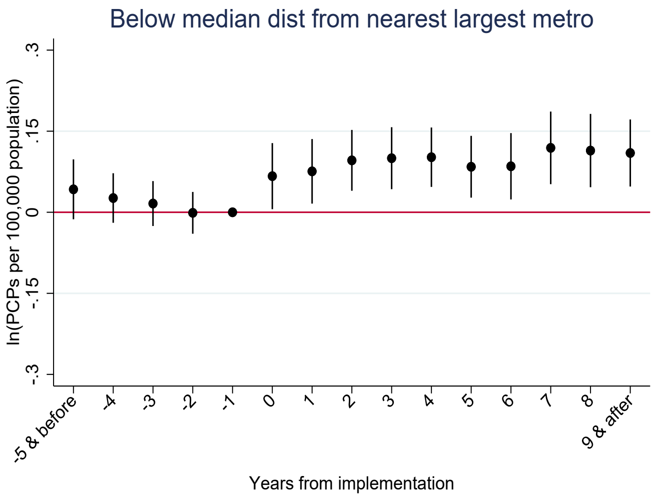
*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (1). Outcome variable is number of deaths of elderly over 65 years per 10,000 population , for all cause mortality (panel a) and amenable cause mortality (panel b). Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. The event window ranges from -4 to +6, dropping unbalanced states (states with less than 4 years of pre-periods or 6 years of post-periods) and unbalanced years (event years outside the event window). Main data source: National Vital Statistics System CDC, 1999-2017

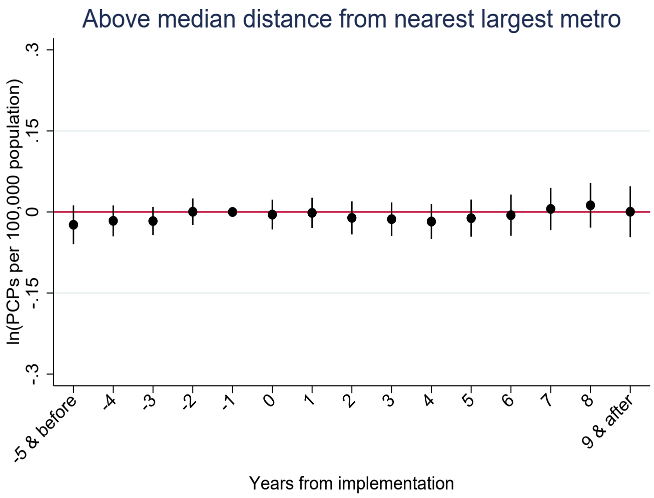
Figure A27: Effects of the policy on entry of eligible practicing physicians by distance to largest metro county in the state: Robustness to alternative distance bins




*Notes:* The above figures plot the pooled difference-in-difference coefficients corresponding to two alternative distance bin widths in panels a and b as compared to Table A2. Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Distance variable represents the distance between a county and nearest largest metro county within the state. Vertical lines represent 95% confidence intervals. The specification includes county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

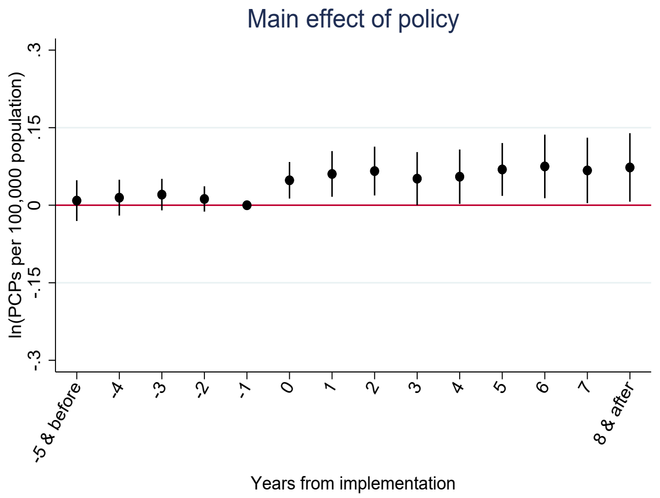
Figure A28: Effects of the policy on entry of eligible practicing physicians by distance to largest metro county in the state:median distance

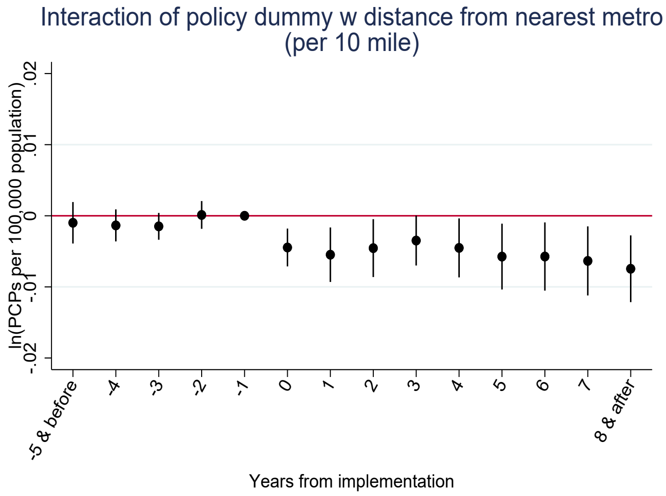



*Notes:* The above figures plot the event study coefficients from an equation similar to (2). Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Distance variable represents the distance between a county and nearest largest metro county within the state. See text and Appendix I for construction of distance variable. Vertical lines represent 95% confidence intervals. The specification includes county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

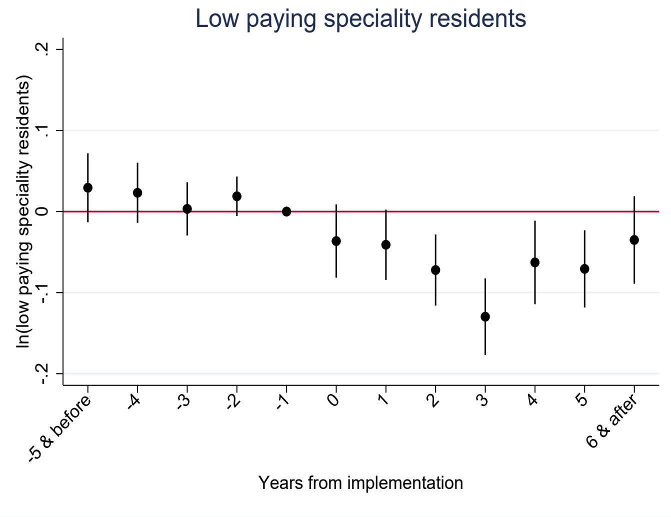
Figure A29: Effects of the policy on entry of eligible practicing physicians by distance to largest metro county in the state: continuous measure of distance



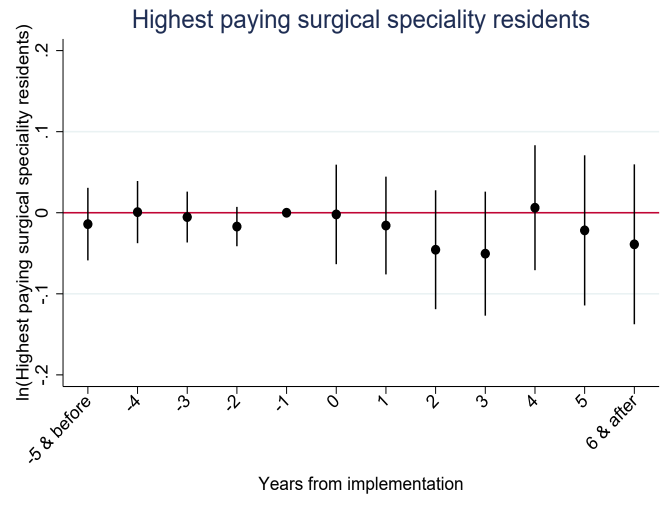




*Notes:* The above figures plot the event study coefficients from an equation similar to (2). Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. Distance variable represents the distance between a county and nearest largest metro county within the state. See text and Appendix G for construction of distance variable. Vertical lines represent 95% confidence intervals. The specification includes county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

Figure A30: Which specialities are likely to be substituted timed with the policy?



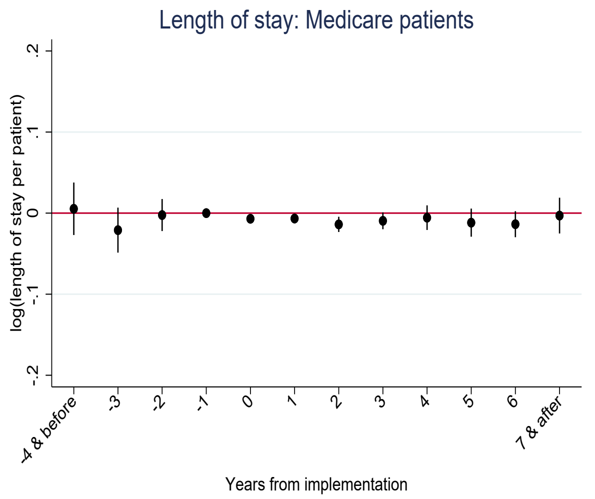

(a)



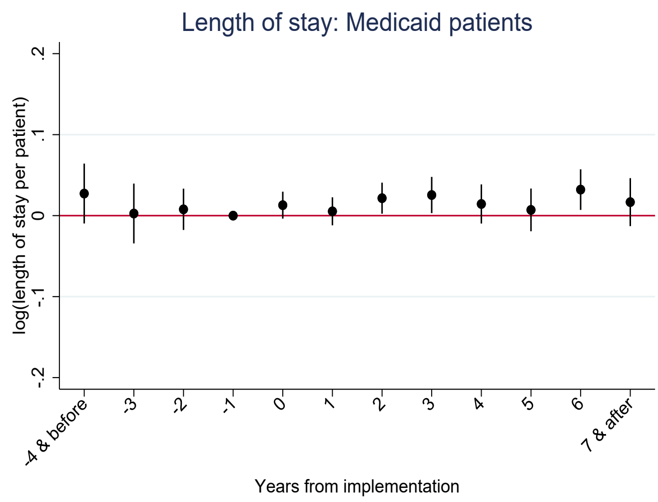

(b)

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (5). Outcome variable is log number of first year residents in low paying untreated specialities (panel a) and high paying untreated specialities (panel b) at a state-year level. Vertical lines represent 95% confidence intervals. All specifications include state level controls, state fixed effects and year fixed effects. Standard errors are clustered at county level. Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

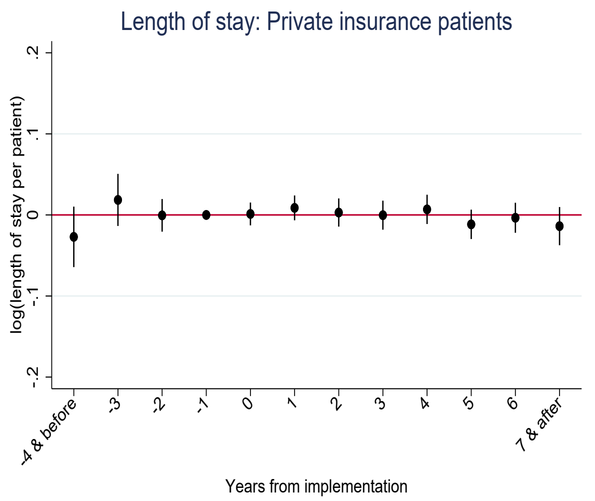
Figure A31: Effects of the policy on hospital length of stay in California by payer category: patient level records



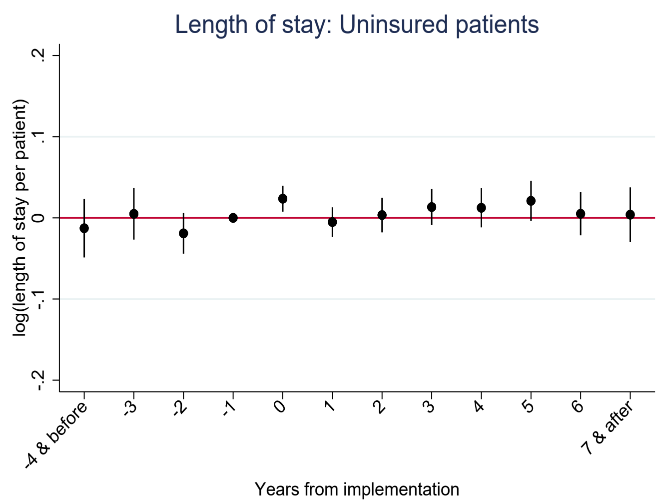

(a) Pre-treatment mean: 3.87




(b) Pre-treatment mean: 12.84



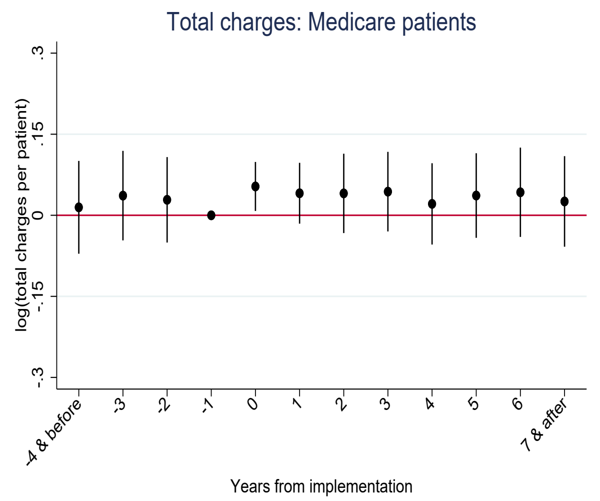

(c)Pre-treatment mean: 2.47



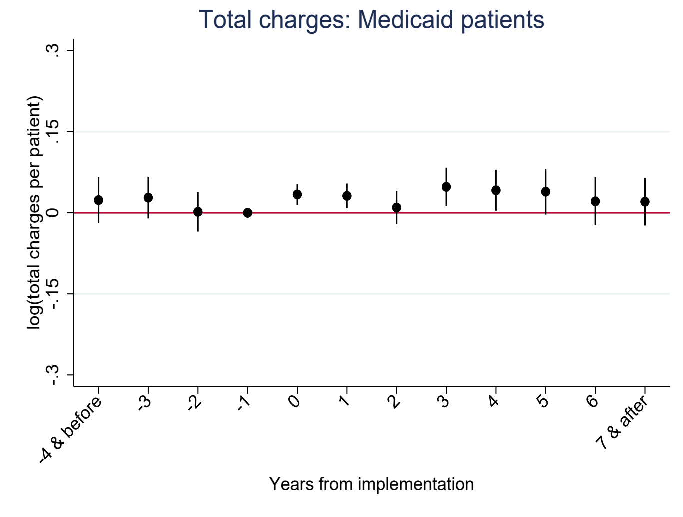

(d)Pre-treatment mean: 3.12

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (6). Outcome variable is log length of stay per patient for Medicare patients (panel a), Medicaid patients (panel b), private insurance patients (panel c) and uninsured patients (panel d). Vertical lines represent 95% confidence intervals. All specifications include hospital level controls, county level controls, hospital fixed effects and year fixed effects. Standard errors are clustered at hospital level. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017

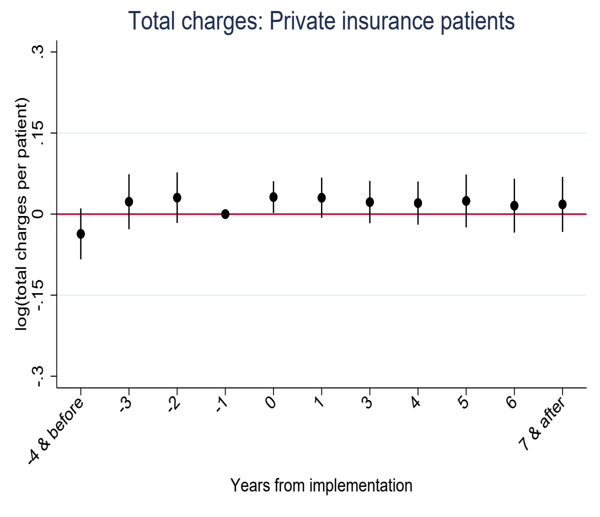
Figure A32: Effects of the policy on hospital charges in California by payer category: patient level records



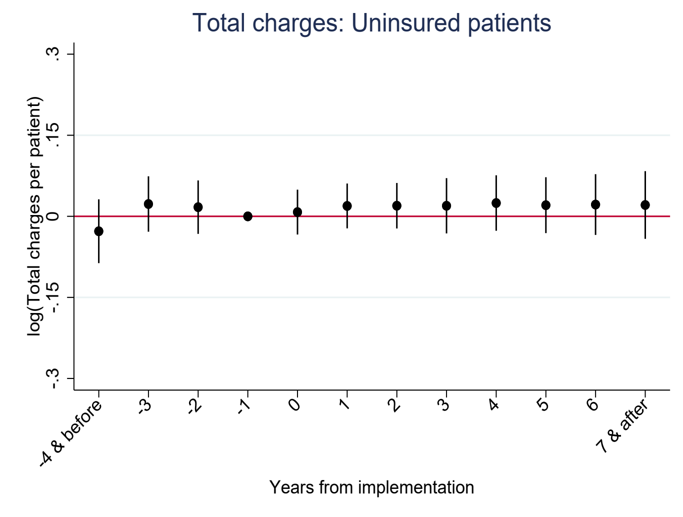

(a) Pre-treatment mean:16852.33




(b) Pre-treatment mean:13675.9



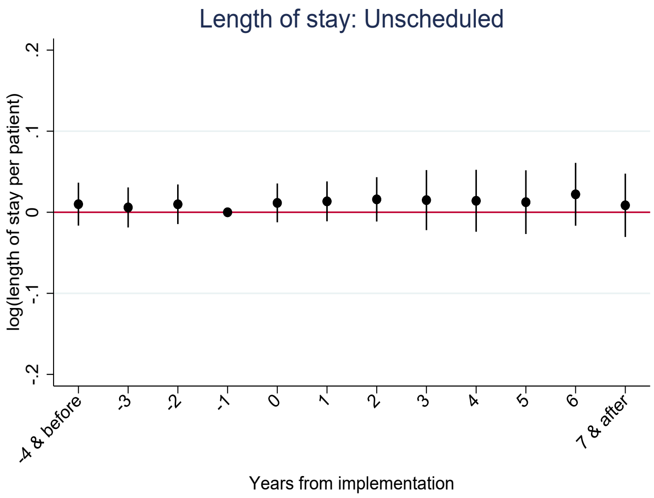

(c) Pre-treatment mean:11009.84



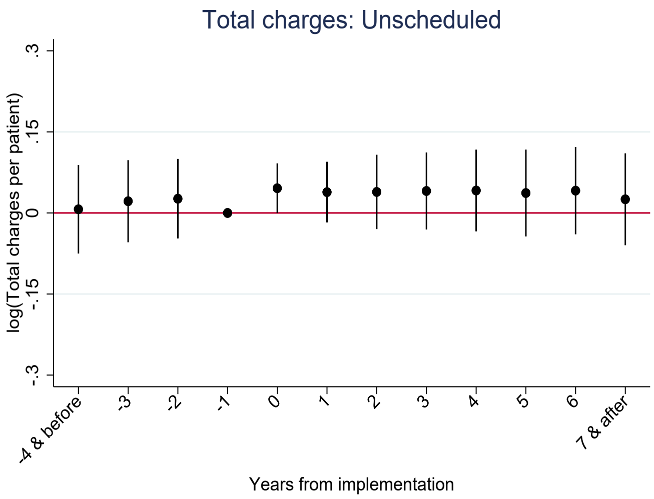

(d) Pre-treatment mean:12876.89

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (6). Outcome variable is log total charges per patient for Medicare patients (panel a), Medicaid patients (panel b), private insurance patients (panel c) and uninsured patients (panel d). Vertical lines represent 95% confidence intervals. All specifications include hospital level controls, county level controls, hospital fixed effects and year fixed effects. Standard errors are clustered at hospital level. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017

Figure A33: Effects of the policy on hospital length of stay and hospital charges for unscheduled admissions in California: patient level records



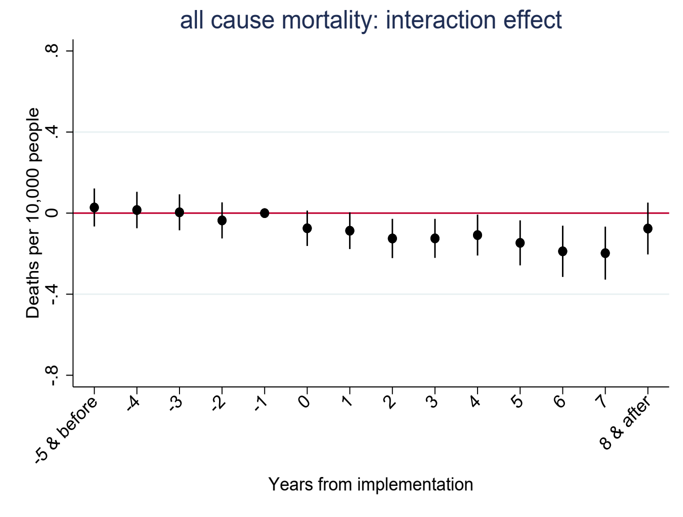

(a) Pre-treatment mean: 3.09



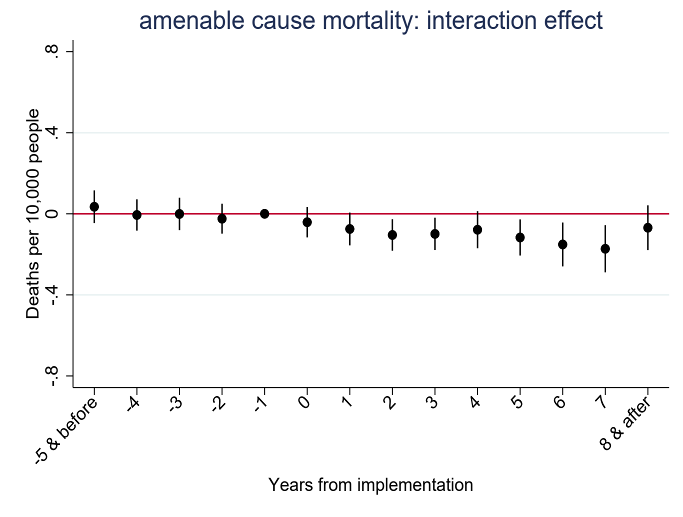

(b) Pre-treatment mean: 14074.81

*Notes:* The above figures plot the event study coefficients $\alpha_k$ from equation (6). Outcome variable is log length of stay per patient (panel a) and log total charges per patient (panel b) for unscheduled admissions. Vertical lines represent 95% confidence intervals. All specifications include hospital level controls, county level controls, hospital fixed effects and year fixed effects. Standard errors are clustered at hospital level. Regressions are weighted by $T + (1-T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017

Figure A34: Heterogenous effects of the policy on mortality of elderly population by baseline physician availability



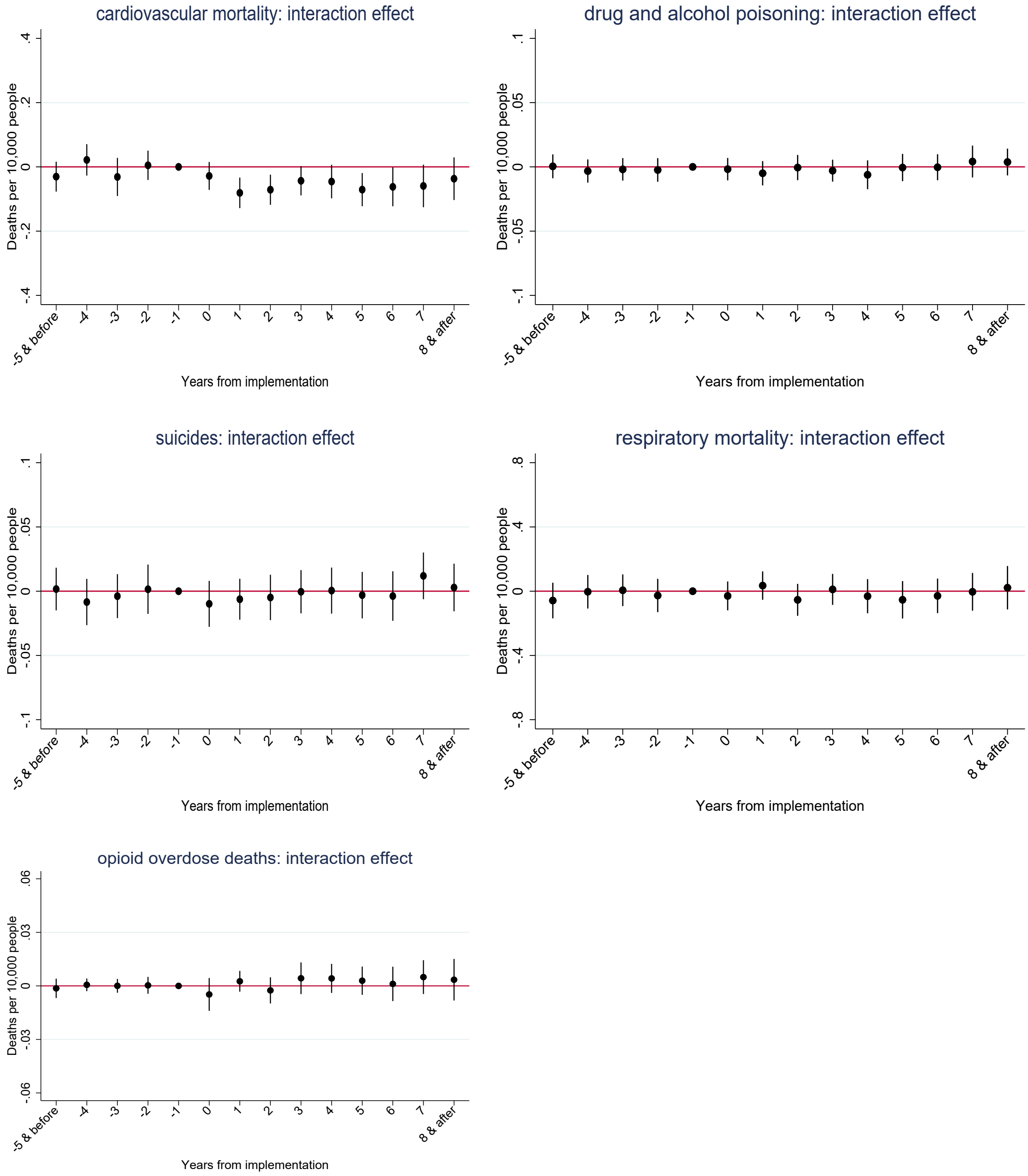

(a)




(b)

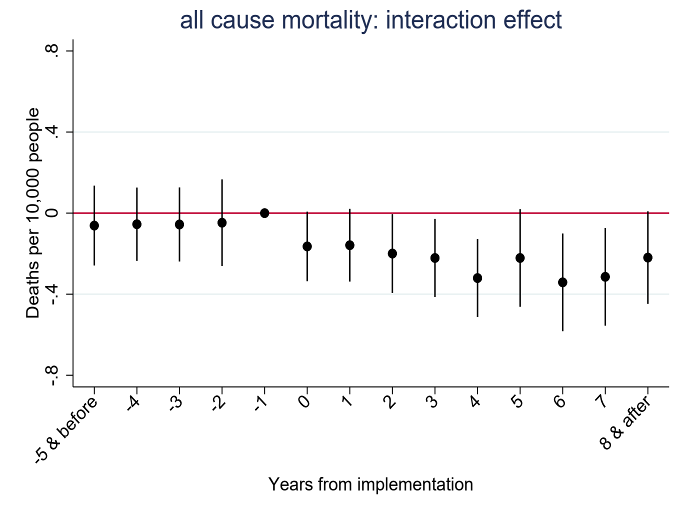
*Notes:* The above figures plot the event study coefficients $\phi_k$ from equation (8). Outcome variable is number of deaths of elderly over 65 years per 10,000 population , for all cause mortality (panel a) and amenable cause mortality (panel b). The list of amenable health conditions included in the paper is reported in Table A17. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017.

Figure A35: Heterogenous effects of the policy on mortality of elderly population by baseline physician availability: Breakdown by causes of death

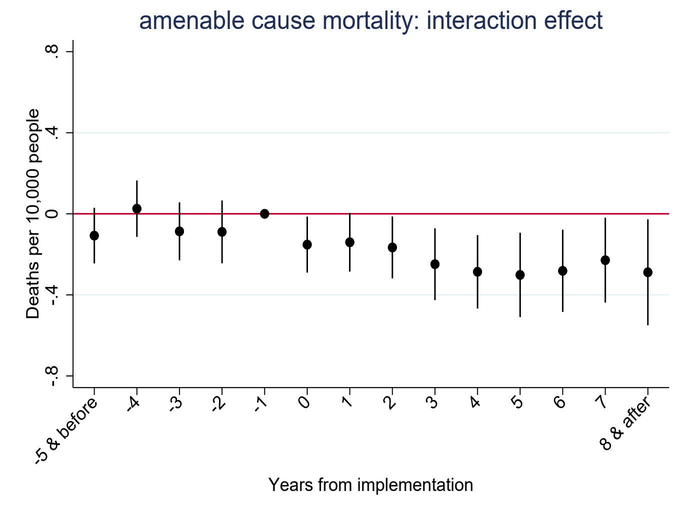



*Notes:* The above figures plot the event study coefficients $\phi_k$ from equation (8). Outcome variable is number of deaths of elderly over 65 years per 10,000 population, for cardiovascular mortality (panel a) respiratory mortality (panel b), suicides (panel c), drug and alcohol poisoning (panel d), opioid overdose mortality (panel e). Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017

Figure A36: Heterogenous effects of the policy on mortality of adult population by baseline physician availability



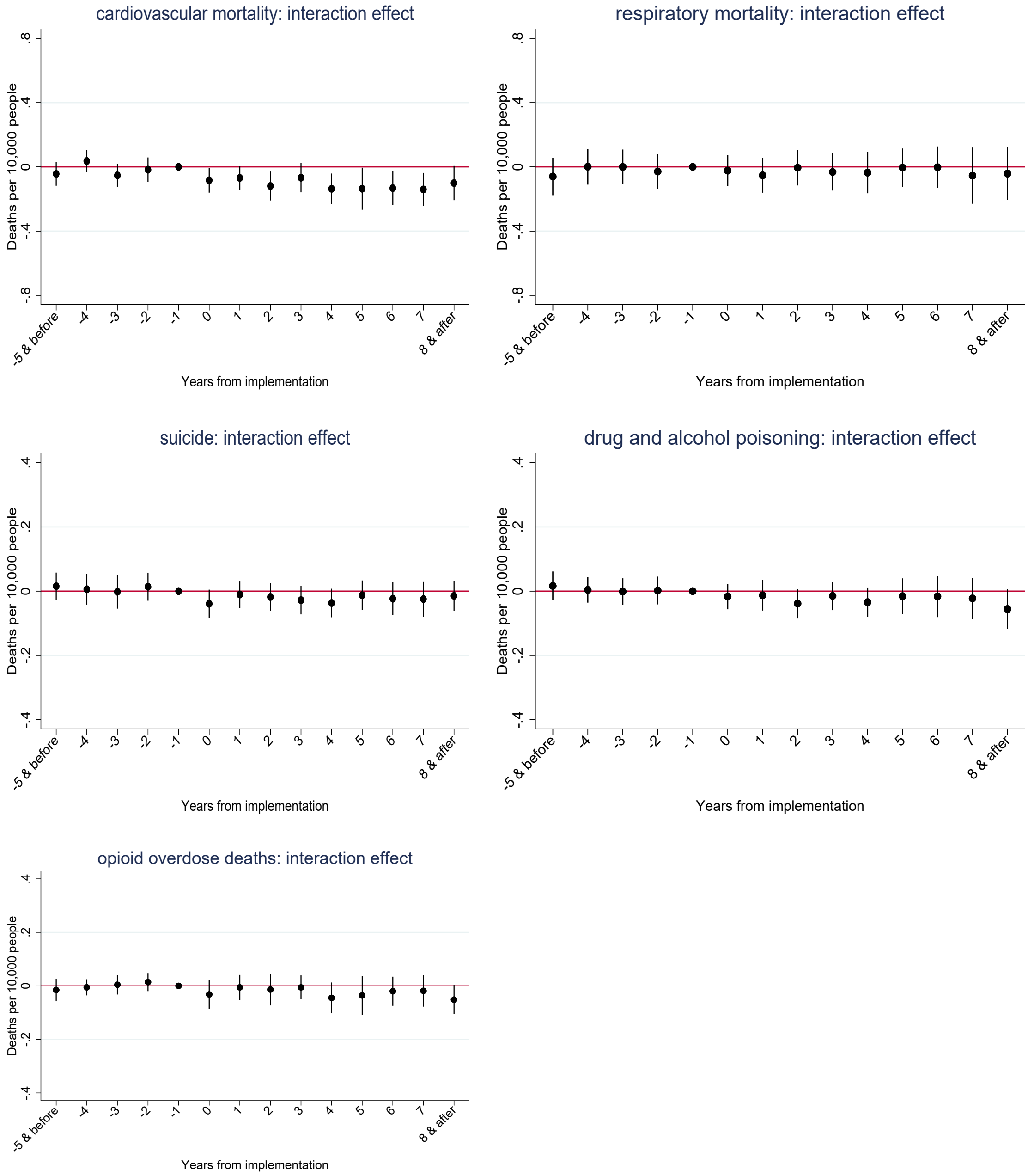

(a)




(b)

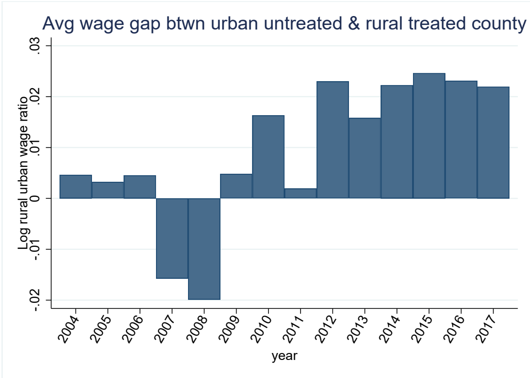
*Notes:* The above figures plot the event study coefficients $\phi_k$ from equation (8). Outcome variable is number of deaths of all adults over 20 years per 10,000 population, for all cause mortality (panel a) and amenable cause mortality (panel b). The list of amenable health conditions included in the paper is reported in Table A17. Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017.

Figure A37: Heterogenous effects of the policy on mortality of adult population by baseline physician availability: Breakdown by causes of death



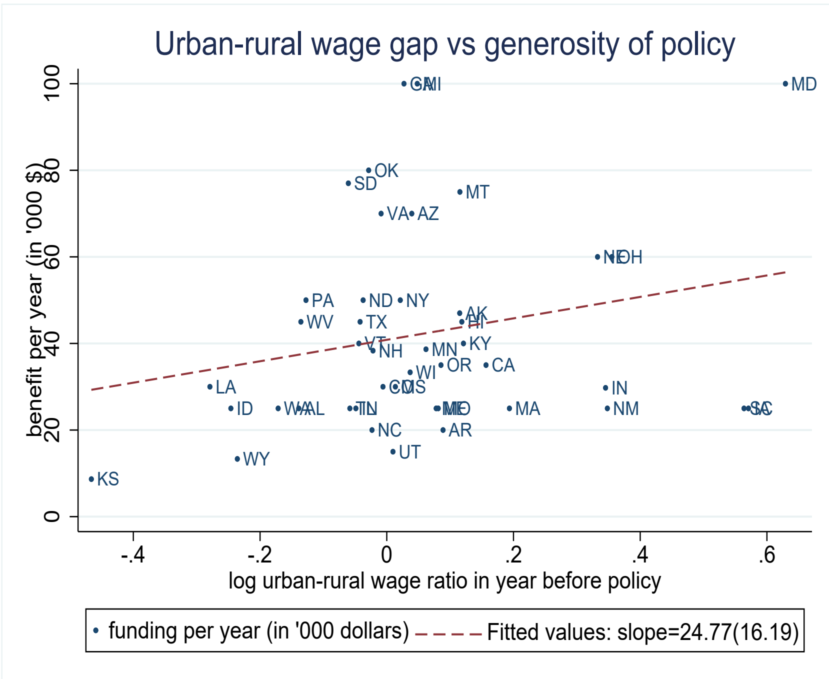

*Notes:* The above figures plot the event study coefficients $\phi_k$ from equation (8). Outcome variable is number of deaths of all adults over 20 years per 10,000 population, for cardiovascular mortality (panel a) respiratory mortality (panel b), suicides (panel c), drug and alcohol poisoning (panel d), opioid overdose mortality (panel e). Vertical lines represent 95% confidence intervals. All specifications include county level controls, county fixed effects and state-by-year fixed effects. Standard errors are clustered at the county level. Main data source: National Vital Statistics System CDC, 1999-2017.

Figure A38: Comparison of wages between rural treated counties and urban untreated counties




*Notes:* The above figure shows the mean wages of physicians (Family and general practitioners and general pediatricians). The wage data is taken from Occupational Employment and Wage statistics (OEWS) metropolitan and non-metropolitan area estimates. These estimates are not adjusted for age and hence do not represent accurate wages of new physicians entering the labor market.

Figure A39: Relationship between urban-rural wage gap in year before policy and benefits offered




*Notes:* The wage data is from IPUMS USA 1970-2014 and is for the category "Physicians and surgeons". The policy spans from 1978 to 2015.

Table A1: Effects of the policy on entry of eligible practicing physicians by speciality

| | (General FM+IM) | (General Ob/Gyn) | (General Pediatrics) | (FM+IM speciality) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| policy dummy | 0.0234** | 0.0625*** | 0.0670*** | 0.0387** |
| | (0.0113) | (0.0221) | (0.0229) | (0.0152) |
| Mean dependent variable | 41.7 | 4.5 | 5.3 | 4.5 |
| Controls | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes |
| Observations | 48,964 | 48,964 | 48,964 | 48,964 |

Notes: Outcome variable is log physicians per 100,000 population for the four broad eligible specialities at a county-year level. Physicians consist of MDs and DOs. All specifications include county and state-by-year fixed effects. County level controls include unemployment rate, per capita income, percentage of population in different age groups (20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65+), percentage of black, white, male and Hispanic population. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A2: Effects of the policy on entry of eligible physicians by distance to largest metro county in the state

Dependent variable: log(physicians per 100,000 population)

| | (1) |
|---|---|
| $\alpha_1$: policy dummy*I (Distance<= 50 miles) | 0.1067*** |
| | (0.0290) |
| $\alpha_2$:policy dummy*I (50 < Distance<= 90 miles) | 0.0942*** |
| | (0.0294) |
| $\alpha_3$:policy dummy*I (90 < Distance<= 130 miles) | 0.0621** |
| | (0.0273) |
| $\alpha_4$:policy dummy*I (130 < Distance<= 170 miles) | -0.0319 |
| | (0.0285) |
| $\alpha_5$: policy dummy*I (Distance > 170 miles) | -0.0328 |
| | (0.0311) |
| Controls | Yes |
| County FE | Yes |
| State X year FE | Yes |
| P value for test of equality | |
| $\alpha_1$ & $\alpha_2$ & $\alpha_3$ & $\alpha_4$ & $\alpha_5$ | 0.0011 |
| $\alpha_1$ & $\alpha_4$ | 0.0049 |
| $\alpha_1$ & $\alpha_5$ | 0.0012 |
| $\alpha_2$ & $\alpha_4$ | 0.0022 |
| $\alpha_2$ & $\alpha_5$ | 0.0046 |
| $\alpha_3$ & $\alpha_4$ | 0.0118 |
| $\alpha_3$ & $\alpha_5$ | 0.0044 |
| Observations | 48,964 |

Notes: Outcome variable is log physicians per 100,000 population at a county-year level. Physicians include both MDs and DOs. Distance variable represents the distance between a county and nearest largest metro county within the state. The specification includes county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The p values correspond to the test of equality of coefficients in various distance bins. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

** p<0.01, ** p<0.05, * p<0.1

Table A3: Effects of the policy on entry of eligible practicing physicians by size of benefits

| | (Baseline) (1) | (Controls) (2) |
|---|---|---|
| policy dummy | 0.0052 | 0.0067 |
| | (0.0215) | (0.0213) |
| Mean dependent variable | 55.0 | 55.0 |
| Controls | No | Yes |
| County FE | Yes | Yes |
| State X year FE | Yes | Yes |
| Observations | 48,964 | 48,964 |

Notes: Outcome variable is log physicians per 100,000 population at a county-year level. Physicians include both MDs and DOs. The generosity of the policy in county c year t is defined as log of total benefit amount in (c,t) cell divided by minimum service period in that cell. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

** p<0.01, ** p<0.05, * p<0.1

Table A4: Effects of the policy on entry of eligible practicing physicians: Heterogeneity by natural and non-natural amenities

| | (1) | (2) |
|---|---|---|
| policy dummy*Dummy for high amenity counties | 0.0683** | |
| | (0.0298) | |
| policy dummy*Standardised amenity index | | 0.0407*** |
| | | (0.0147) |
| Mean dependent variable | 48.5 | 54.8 |
| Controls | Yes | Yes |
| County FE | Yes | Yes |
| State X year FE | Yes | Yes |
| Observations | 48,964 | 48,964 |

Notes: Outcome variable is log physicians per 100,000 population at a county-year level. Physicians consist of MDs and DOs. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated counties. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A5: Who is screened in by the policy?

| | (ranked) | (foreign) | (unranked) | (state of training) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| policy dummy | 0.0339** | 0.0274** | −0.0494** | 0.0834*** |
| | (0.0133) | (0.0114) | (0.0196) | (0.0145) |
| Mean dependent variable | 0.162 | 0.076 | 0.759 | 0.794 |
| Controls | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes |
| Observations | 46,835 | 46,835 | 46,835 | 46,835 |

Notes: Outcome variable is proportion of incoming physicians from ranked US medical school (col 1), proportion of incoming physicians from foreign medical school (col 2) , proportion of incoming physicians from unranked US medical school (col 3) and proportion of incoming physicians who start their job in their state of training (col 4). All outcome variables are aggregated to county-year level using individual longitudinal data. Physicians consist of MDs and DOs. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of dependent variable in treated counties. Main data source: AMA physician masterfile, 1996-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A6: Effects of the policy on choice of specialisation of training physicians: first year residents

| | (Primary care/treated residents) | (Speciality/untreated residents) |
|---|---|---|
| | (1) | (2) |
| policy dummy | 0.0217*** | −0.0243** |
| | (0.0057) | (0.0110) |
| Mean dependent variable | 396 | 595 |
| Controls | Yes | Yes |
| State FE | Yes | Yes |
| Year FE | Yes | Yes |
| Observations | 1,275 | 1,275 |

Notes: Outcome variable is log number of first year residents in treated specialities (col 1) and untreated specialities(col 2) at a state-year level. See Appendix B for definition of treated and untreated specialities. All specifications include state and year fixed effects. Standard errors are clustered at county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated states. Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

*** p<0.01, ** p<0.05, * p<0.1

Table A7: Effects of the policy on hospital reimbursement and physician reimbursement for Medicare enrollees

| | (Hospital reimbursement) | (Physician reimbursement) |
|---|---|---|
| | (1) | (2) |
| policy dummy | $-0.0144^{**}$ | -0.0033 |
| | (0.0062) | (0.0095) |
| Mean dependent variable | 4153.35 | 1846.24 |
| Controls | Yes | Yes |
| County FE | Yes | Yes |
| State X year FE | Yes | Yes |
| Observations | 38,320 | 38,320 |

Notes: Outcome variable is log hospital reimbursement per Medicare enrollee in column 1 and log physician reimbursement per Medicare enrollee in column 2 at a county-year level. Both hospital reimbursements and physician reimbursements are price, age, race and sex adjusted. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. The mean of the dependent variable is the baseline average of level of dependent variable in treated counties. Main data source: Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A8: Effects of the policy on quality of primary care of Medicare beneficiaries

| | (Hemoglobin a1c test) | (Blood lipids test) | (Eye examination) | (Mammogram) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| policy dummy | 0.0152 | -0.0009 | 0.0074 | -0.0020 |
| | (0.0089) | (0.0076) | (0.0115) | (0.0081) |
| Mean dependent variable | 0.0058 | 0.0052 | 0.0048 | 0.0041 |
| Controls | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes |
| Observations | 38,320 | 38,320 | 38,320 | 38,320 |

Notes: Outcome variable is log per capita diabetic medicare enrollees having a1c test (col 1), log per capita diabetic medicare enrollees having blood lipids test (col 2), log per capita medicare enrollees having eye examination (col 3), log per capita female medicare enrollees having at least one mammogram (col 4) at a county-year level. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: Area Health Resource File, Dartmouth Atlas of Healthcare, 2000-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A9: Effects of the policy on entry of specialists and older physicians: Falsification tests

| | (physician MDs :eligible) | (physician DOs: eligible ) | (older physicians: ineligible) | (specialists: ineligible) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| policy dummy | 0.0491*** | 0.0783** | 0.0064 | 0.0104 |
| | (0.0140) | (0.0332) | (0.0105) | (0.008) |
| Mean dependent variable | 47.3 | 7.58 | 30.05 | 24.2 |
| Controls | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes |
| Observations | 48,964 | 48,964 | 48,964 | 48,964 |

Notes: Outcome variable is log physicians per 100,000 population or log specialists per 100,000 population at a county-year level. Columns 1 and 2 report the results of main specification from Table 2. Older physicians comprise those above the age of 45 years. Columns 3 and 4 include both MDs and DOs. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017.

** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A10: Difference-in-differences Estimator Decomposition

| | (MD entry) | (DO entry) | (Access to physician) | (ER visit) | (Hospital admission) | (Hospital reimbursement) |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Timing comparisons | 0.0149 | 0.0199 | 0.0172 | 0.0036 | 0.0139 | 0.00013 |
| | [0.0023] | [0.0023] | [0.0012] | [0.0012] | [0.0012] | [0.0012] |
| Always vs Timing | 0.0049 | 0.0433 | 0.0209 | -0.0265 | -0.0933 | -0.0112 |
| | [0.0209] | [0.0209] | [0.0178] | [0.0178] | [0.0178] | [0.0178] |
| Never vs Timing | 0.0563 | 0.0840 | 0.0275 | -0.0564 | -0.1082 | -0.0177 |
| | [0.9498] | [0.9498] | [0.9423] | [0.9423] | [0.9423] | [0.9423] |
| Always vs Never | 1.3286 | -0.9219 | 0.3542 | -1.1291 | -1.1410 | -1.0242 |
| | [0.0003] | [0.0003] | [0.0002] | [0.0002] | [0.0002] | [0.0002] |
| Within comparisons | -0.1825 | -0.0812 | -0.1497 | 0.1039 | 0.0597 | 0.0683 |
| | [0.0267] | [0.0267] | [0.0385] | [0.0385] | [0.0385] | [0.0385] |
| Mean Dependent variable | 47.3 | 7.58 | 0.081 | 0.096 | 80.19 | 4153.35 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes | Yes |
| StatexYear FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 48,964 | 48,964 | 38,320 | 38,320 | 38,320 | 38,320 |

Notes:Outcome variable is log physicians per 100,000 population in columns 1 and 2, log per capita Medicare enrollees having at least one ambulatory visit to physician in column 3, log per capita ER visits of Medicare beneficiaries in column 4, log preventable hospital stays rate of Medicare beneficiaries in column 5 and log hospital reimbursement per Medicare enrollee in column 6 at a county year level. Each row corresponds to a 2X2 estimator with corresponding weights reported in brackets below the estimates. All specifications include county and state-by-year fixed effects.

Table A11: Difference-in-differences Estimator Decomposition- Training physicians

| | (Primary care/treated residents) | (Speciality/untreated residents) |
|---|---|---|
| | (1) | (2) |
| Timing comparisons | 0.0231 | 0.0302 |
| | [0.0952] | [0.0952] |
| Always vs Timing | 0.0154 | -0.0249 |
| | [0.8306] | [0.8306] |
| Never vs Timing | 0.1697 | -0.1806 |
| | [0.0575] | [0.0575] |
| Always vs Never | 1.2873 | -1.6952 |
| | [0.00025] | [0.00025] |
| Within comparisons | -0.2040 | 0.2611 |
| | [0.0165] | [0.0165] |
| Mean Dependent variable | 396 | 595 |
| Controls | Yes | Yes |
| State FE | Yes | Yes |
| Year FE | Yes | Yes |
| Observations | 1,275 | 1,275 |

Notes: Outcome variable is log number of first year residents in treated specialities (col 1) and untreated specialities(col 2) at a state-year level. Each row corresponds to a 2X2 estimator with corresponding weights reported in brackets below the estimates. All specifications include state and year fixed effects.

Table A12: Effects of the policy on entry of eligible practicing physicians and health outcomes: Population weighted

| | (MDs) | (DOs) | (Access to physicians) | (ER visits) | (Hospital admissions) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| policy dummy | $0.0505^{***}$ | $0.0812^{**}$ | $0.0213^{**}$ | $-0.0512^{***}$ | $-0.1021^{***}$ |
| | (0.0153) | (0.0344) | (0.0092) | (0.0173) | (0.0361) |
| Mean dependent variable | 56.8 | 7.87 | 0.085 | 0.099 | 83.9 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 48,964 | 48,964 | 38,320 | 38,320 | 38,320 |

Notes: Outcome variable is log physicians per 100,000 population in columns 1 and 2, log per capita Medicare enrollees having at least one ambulatory visit to physician in column 3, log per capita ER visits of Medicare beneficiaries in column 4 and log preventable hospital stays rate of Medicare beneficiaries in column 5 at a county year level. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Regressions and mean of dependent variables are weighted by county population. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017, Area Health Resource File, Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A13: Effects of the policy on entry of eligible practicing physicians and health outcomes: Propensity score reweighted sample

| | (MDs) | (DOs) | (Access to physicians) | (ER visits) | (Hospital admissions) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| policy dummy | 0.0597*** | 0.1043*** | 0.0214*** | −0.0520*** | −0.1016*** |
| | (0.0144) | (0.0329) | (0.0081) | (0.0172) | (0.0353) |
| Mean dependent variable | 47.41 | 7.92 | 0.085 | 0.093 | 85.6 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 49,033 | 49,033 | 38,374 | 38,374 | 38,374 |

Notes: Outcome variable is log physicians per 100,000 population in columns 1 and 2, log per capita Medicare enrollees having at least one ambulatory visit to physician in column 3, log per capita ER visits of Medicare beneficiaries in column 4 and log preventable hospital stays rate of Medicare beneficiaries in column 5 at a county year level. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$, where T is the indicator for treatment and p is the estimated propensity score. Main data source: Area Health Resource File and AMA physician masterfile, 1995-2017, Area Health Resource File, Dartmouth Atlas of Healthcare, county level Medicare claims, 2000-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A14: Effects of the policy on speciality choice of first year training physicians: Heterogenous effects

| | (General FM+IM) | (General Ob/Gyn) | (General Pediatrics) | (FM+IM speciality) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| policy dummy | -0.0011 | 0.0786** | 0.0494*** | 0.0563*** |
| | (0.0332) | (0.0334) | (0.0187) | (0.0211) |
| Mean dependent variable | 231 | 25 | 61 | 79 |
| Controls | Yes | Yes | Yes | Yes |
| State FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 1,275 | 1,275 | 1,275 | 1,275 |

Notes: Outcome variable is log number of first year residents for the four broad eligible specialities at a state-year level (See Appendix B). All specifications include state and year fixed effects. Standard errors are clustered at county level and reported in parenthesis. Main data source: GME Track Association of American Medical Colleges and National Resident Matching Program, AY 1994-95 to 2018-19.

*** p<0.01, ** p<0.05, * p<0.1

Table A15: Effects of the policy on hospital length of stay in California by payer category: patient level records

| | (Medicare) (1) | (Medicaid) (2) | (Private insurance) (3) | (Uninsured) (4) | (Unscheduled) (5) |
|---|---|---|---|---|---|
| policy dummy | -0.0085 | 0.0164 | -0.0012 | 0.0094 | 0.0142 |
| | (0.0053) | (0.0123) | (0.0091) | (0.0109) | (0.0194) |
| Mean dependent variable | 3.87 | 12.84 | 2.47 | 3.12 | 3.09 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Hospital FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 5548 | 5548 | 5548 | 5548 | 5548 |

Notes: Outcome variable is log length of stay per patient for Medicare patients (col 1), Medicaid patients (col 2), private insurance patients (col 3), uninsured patients (col 4) and log number of unscheduled admissions (col 5) at a hospital-year level. All specifications include hospital and year fixed effects. Standard errors are clustered at hospital level and reported in parenthesis. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017 .
*** p<0.01, ** p<0.05, * p<0.1

Table A16: Effects of the policy on hospital charges in California by payer category: patient level records

| | (Medicare) (1) | (Medicaid) (2) | (Private insurance) (3) | (Uninsured) (4) | (Unscheduled) (5) |
|---|---|---|---|---|---|
| policy dummy | 0.0376 | 0.0305 | 0.0222 | 0.0192 | 0.0385 |
| | (0.0398) | (0.0224) | (0.0199) | (0.0211) | (0.0285) |
| Mean dependent variable | 16852.33 | 13675.9 | 11009.84 | 12876.89 | 14074.81 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Hospital FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 5548 | 5548 | 5548 | 5548 | 5548 |

Notes: Outcome variable is log total charges per patient for Medicare patients (col 1), Medicaid patients (col 2), private insurance patients (col 3), uninsured patients (col 4) and log number of unscheduled admissions (col 5) at a hospital-year level. All specifications include hospital and year fixed effects. Standard errors are clustered at hospital level and reported in parenthesis. Regressions are weighted by $T + (1 - T) * \frac{p}{1-p}$ where T is the indicator for treatment and p is the estimated propensity score. Main data source: OSHPD patient discharge data at the day level, 1999-2017 .
*** p<0.01, ** p<0.05, * p<0.1

Table A17: List of amenable health conditions

| (Conditions)<br>(ICD-10 codes) | |
|---|---|
| Infectious & Parasitic Diseases | A00-B99 |
| Neoplasms (ALL) | C00-D48 |
| Disorders of thyroid gland | E00-E07 |
| Diabetes Mellitus | E10-E14 |
| Epilepsy | G40-G41 |
| Chronic rheumatic heart diseases | I05-I09 |
| Hypertensive diseases | I10-I13, I15 |
| Ischemic heart diseases | I20-I25 |
| Cardiomyopathy | I42 |
| Atrial fibrillation and flutter | I48 |
| Other cardiac arrhythmias | I49 |
| Heart failure | I50 |
| Cerebrovascular diseases | I60-I69 |
| All respiratory diseases | J00-J98 |
| Gastric and duodenal ulcers | K25-K27 |
| Gastrojejunal ulcers | K28 |
| Diseases of appendix | K35-K38 |
| Hernia | K40-K46 |
| Diseases of gallbladder and biliary tract | K80-K83 |
| Acute pancreatitis | K85 |
| Infections of the skin and subcutaneous tissue | L00-L08 |
| Infectious arthropathies | M00-M02 |
| Glomerular diseases | N00-N07 |
| Renal tubulo-interstitial diseases | N10-N15 |
| Renal failure | N17-N19 |
| Unspecified contracted kidney, small kidney unknown cause | N26-N27 |
| Hyperplasia of prostate | N40 |
| Pregnancy, childbirth and the puerperium | O00-O99 |
| Conditions originating in the perinatal period | P00-P96 |
| Misadventures to patients during surgical and medical care | Y60-Y69, Y83-84 |

Notes: The above table reports the list of amenable health conditions used in the paper and their corresponding ICD-10 codes, following Sommers, Long and Baicker(2014).

Table A18: Effects of the policy on mortality of elderly population

| | All cause mortality | | Amenable cause mortality | |
|---|---|---|---|---|
| | (Baseline) | (Controls) | (Baseline) | (Controls) |
| | (1) | (2) | (3) | (4) |
| policy dummy | $-2.6271^{**}$ | $-2.3841^{**}$ | $-2.0872^{**}$ | $-2.0623^{**}$ |
| | (1.1107) | (1.0292) | (0.9579) | (0.9368) |
| Controls | No | Yes | No | Yes |
| County FE | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes |
| Observations | 40,449 | 40,449 | 40,449 | 40,449 |

Notes: Outcome variable is number of deaths of elderly over 65 years per 10,000 population , for all cause mortality (cols 1 & 2) and amenable cause mortality (cols 3 & 4) at a county-year level. The list of amenable health conditions included in the paper are reported in Table A17. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: National Vital Statistics System CDC, 1999-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A19: Effects of the policy on mortality of elderly population by cause of death

| | (Cardiovascular) | (Respiratory) | (Suicides) | (Drug poisoning) | (Opioid overdose) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | |
| policy dummy | $-0.7392^{**}$ | -0.0538 | -0.0192 | $-0.0452^{*}$ | 0.00292 |
| | (0.3166) | (0.3897) | (0.0394) | (0.0229) | (0.0236) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 |

Notes: Outcome variable is number of deaths of elderly over 65 years per 10,000 population, for cardiovascular mortality (col 1), respiratory mortality (col 2), suicides (col 3), drug and alcohol poisoning (col 4), opioid overdose mortality (col 5) at a county-year level. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: National Vital Statistics System CDC, 1999-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A20: Effects of the policy on mortality of adult population

| | All cause mortality | | Amenable cause mortality | |
|---|---|---|---|---|
| | (Baseline) | (Controls) | (Baseline) | (Controls) |
| | (1) | (2) | (3) | (4) |
| policy dummy | $-2.9573^{**}$ | $-2.6816^{**}$ | $-2.4561^{**}$ | $-2.3650^{**}$ |
| | (1.2576) | (1.1970) | (1.1086) | (1.0337) |
| Controls | No | Yes | No | Yes |
| County FE | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes |
| Observations | 40,449 | 40,449 | 40,449 | 40,449 |

Notes: Outcome variable is number of deaths of all adults over 20 years per 10,000 population , for all cause mortality (cols 1 & 2) and amenable cause mortality (cols 3 & 4) at a county-year level. The list of amenable health conditions included in the paper are reported in Table A17. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: National Vital Statistics System CDC, 1999-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A21: Effects of the policy on mortality of adult population by cause of death

| | (Cardiovascular) | (Respiratory) | (Suicides) | (Drug poisoning) | (Opioid overdose) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | |
| policy dummy | $-0.9126^{**}$ | 0.0576 | -0.0547 | 0.0248 | 0.0255 |
| | (0.3594) | (0.2689) | (0.1006) | (0.1334) | (0.1216) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 |

Notes: Outcome variable is number of deaths of all adults over 20 years per 10,000 population, for cardiovascular mortality (col 1), respiratory mortality (col 2), suicides (col 3), drug and alcohol poisoning (col 4), opioid overdose mortality (col 5) at a county-year level. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: National Vital Statistics System CDC, 1999-2017.

*** p<0.01, ** p<0.05, * p<0.1

Table A22: Heterogeneous effects of the policy on mortality of elderly population by baseline physician availability

| | (All)<br>(1) | (Amenable)<br>(2) | (Cardiovascular)<br>(3) | (Respiratory)<br>(4) | (Suicides)<br>(5) | (Drugs)<br>(6) | (Opioid)<br>(7) |
|---|---|---|---|---|---|---|---|
| policy dummy | −0.1251** | −0.1045*** | −0.0557** | -0.0148 | -0.00154 | -0.0018 | -0.0017 |
| * Baseline physicians/10,000 pop | (0.0494) | (0.0396) | (0.0242) | (0.0491) | (0.0107) | (0.0044) | (0.0049) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 |

Notes: Outcome variable is number of deaths of elderly over 65 years per 10,000 population , for all cause mortality (col 1), amenable cause mortality (col 2) and by main causes of death (cols 3-7) at a county-year level. Baseline physicians refer to number of physicians in the year before the policy. The list of amenable health conditions included in the paper are reported in Table A17. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: National Vital Statistics System CDC, 1999-2017.
*** p<0.01, ** p<0.05, * p<0.1

Table A23: Heterogeneous effects of the policy on mortality of adult population by baseline physician availability

| | (All)<br>(1) | (Amenable)<br>(2) | (Cardiovascular)<br>(3) | (Respiratory)<br>(4) | (Suicides)<br>(5) | (Drugs)<br>(6) | (Opioid)<br>(7) |
|---|---|---|---|---|---|---|---|
| policy dummy | −0.2392** | −0.2285** | −0.1092** | -0.0277 | -0.0233 | -0.0252 | -0.0220 |
| * Baseline physicians/10,000 pop | (0.0983) | (0.0904) | (0.0460) | (0.0508) | (0.0259) | (0.0229) | (0.0270) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| County FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State X year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 | 40,449 |

Notes: Outcome variable is number of deaths of all adults over 20 years per 10,000 population, for all cause mortality (col 1), amenable cause mortality (col 2) and by main causes of death (cols 3-7) at a county-year level. Baseline physicians refer to number of physicians in the year before the policy. The list of amenable health conditions included in the paper are reported in Table A17. All specifications include county and state-by-year fixed effects. Standard errors are clustered at the county level and reported in parenthesis. Main data source: National Vital Statistics System CDC, 1999-2017.
*** p<0.01, ** p<0.05, * p<0.1

# A Description of eligible sites for the program

The eligible sites/facilities/counties have been obtained from the state health departments/websites. The selection criteria of these sites differs by state. In general they are:

1.The population to primary care physician ratio should be greater than two thousand to one.

2.A free clinic.

3.An area where the population groups face special health problems.

4.An area where physician practice patterns limit access to primary care or that have an unmet provider need $> 0.25$ Full time equivalent(FTE).

States like South Dakota have additional criteria that the area must not be located within a 20-mile radius extending from the city center of a city having more than 50,000 people. Several states like Colorado, Louisiana, Kentucky, Nebraska, North Carolina, Wisconsin etc. have a list of approved sites specified by their address,city and zip code or their census tract. In such cases, I consider the county of the address/census tract as the treated county.

The eligible facilities for this program generally consist of Rural health clinic, rural health clinic in conjunction with critical access hospitals,Indian Health Service outpatient centers, Tribal 638 Outpatient centers, rural private practice sites, hospital affiliated outpatient primary care clinic, private non-profit primary care and mental health clinics, long term care facilities, state correctional facilities, solo or group practices, state and county health department clinics, mobile units.

## B Treated and untreated specialities during the training stage

| Broad category | Subcategories |
|---|---|
| Family medicine | Family medicine, Geriatric medicine (FM), Sports medicine (FM), Clinical Informatics (FM), Psychiatry/Family Practice, Internal Medicine/Family Practice,.. |
| Internal medicine | Clinical Informatics (IM), Internal medicine, Critical care medicine (IM), Endocrinology, Diabetes, and Metabolism (IM), Infectious Disease (IM), Geriatric Medicine (IM), Pulmonary Disease & Critical Care Medicine (IM), Sports Medicine (IM), Nephrology (IM),.... |
| Obstetrics and Gynecology | Obstetrics and Gynecology |
| Pediatrics | Pediatrics, Child Abuse Pediatrics, Pediatrics/Emergency Medicine, Pediatrics/Anesthesiology, Pediatrics/Psychiatry/Child and Adolescent Psychiatry,.... |
| Surgical specialities | Otolaryngology, Ophthalmology, Orthopaedic Surgery, Orthopaedic Sports Medicine (OS), Surgery-General, Pediatric Surgery (GS), Surgical Critical Care (GS), Vascular Surgery (GS), Complex General Surgical Oncology (GS),... |
| Medical specialities | Hematology & oncology (IM), Interventional Cardiology (IM), Clinical Cardiac Electrophysiology (IM), Transplant Hepatology (IM),Internal Medicine/Neurology,Pediatric Cardiology (Pediatrics), ... |
| Other specialities | Pathology-Anatomic & Clinical, Cytopathology (P), Forensic Pathology (P), Neuropathology (P), Blood Banking/Transfusion Medicine (P), Neuroradiology (R-D), Vascular & Interventional Radiology (R-D), Nuclear Medicine, Neurodevelopmental Disabilities,.... |

Notes:Above are examples of treated and untreated specialities of trainees, which are considered in the paper. 'FM': Family medicine, 'IM': Internal medicine, 'OS': Orthopaedic surgery, 'GS': General surgery, 'P':Pathology, 'R-D': Radiology-Diagnostic.The first four broad categories comprise the treated specialities, while the last three are included in the set of untreated specialities. The exhaustive list is available from the author on request.

# C Robustness and alternative specification details

## C.1 Decomposing the difference-in-differences estimator

In this section, I discuss the results of decomposing the difference-in-differences estimate using Goodman-Bacon methodology(2018). In a standard difference-in-differences model, the average treatment effect is the difference in change in outcomes between treated and control groups, before and after the treatment. However, when treated units are exposed to the treatment at different times, as in the case of state+local loan repayment assistance policy , the difference-in-differences estimate is a weighted average of all possible 2 group-2 period (2X2) estimators. Some of these estimators compare between units treated at a particular time (treatment) to untreated units (pure controls). Other comparisons are made between units treated at two different times, using later treated group as control before its treatment starts and earlier treated group as control after its treatment starts. The estimates derived from comparisons between treatment groups treated at different points in time, bias the single coefficient estimator away from the true treatment effect. To identify the size of the bias in my estimates, I decompose the difference-in-differences model into five groups of estimators- Timing comparisons, always vs timing, never vs timing, always vs never, within comparisons. In Table A10, I present the decomposition results for the main outcomes. Each column, pertaining to a particular outcome, shows the five groups of 2X2 estimators with corresponding weights in brackets below the estimates. The following conclusions emerge: *First*, around 94-95% of the main estimate is explained by comparisons between treated counties and never treated/untreated counties. This finding is consistent with a significant number of untreated counties present in the sample. *Second*, only 0.12-0.23% of the main estimate is explained by timing variation among treatment groups i.e. counties treated later serving as controls for counties receiving treatment earlier and counties treated earlier acting as controls for later treated counties. This is in line with my empirical strategy whereby I compare treated and untreated counties within a treated state and a given year. Additionally, all treated counties in the state are treated at the same time the state is treated and these treated counties do not switch their treatment status during my sample period. Thus, possibilities of comparing earlier set of treated counties with later set of treated counties and vice versa are limited, reducing the scope of bias in the overall difference-in-difference estimate. *Third*, approximately 2% of the main estimate is accounted for by always vs timing comparison. This component arises because some counties(states) are treated before the start of my analysis period. *Fourth*, the remaining 2.7-3.8% of the main estimate is explained by within comparisons or "residual component". This residual component takes into account the variation in controls across always treated and never treated counties. However, most of the weight on this residual component arises from the variation due to inclusion of state-by-year fixed effects in the regression. The remaining weights are allocated to time varying controls which change the baseline estimates moderately as seen in Tables 2 and 3.

I also carry out a similar decomposition exercise for training physicians in Table A11. The results in Table A11 imply the following conclusions. *First*, majority of the variation (83%) in the main estimate is explained by "always vs timing comparison", while "never vs timing" comparison accounts for only 5.75% of the main estimate. This is consistent with the fact that there are only two untreated/never treated states in my sample. Recall that, for training physicians, I compare treated and untreated states, before and after the policy. *Second*, even though there are possibilities of comparing earlier treated states with later treated states, the weight attributed to timing comparisons is only 9.52% and thus does not bias the estimates significantly. *Third*, the residual component explains about 1.65% of the main estimate. Thus, the extent to which time varying controls may bias the true treatment effect is not significantly large.

### C.2 Balanced panel and population weights

I estimate balanced panel fixed effect models using event studies as shown in Figure A21-Figure A26. A concern with the baseline specification in my main results on entry of physicians is that the increase in point estimates over time in Figure 4 may represent changing composition of treated states rather than actual entry of physicians. To address this concern, I consider a balanced panel over the event window -4 to +6, dropping the unbalanced states and years. Thus, states with less than 4 years of pre-periods or 6 years of post-periods are dropped from the sample. The results are largely similar to the unbalanced panel results, suggesting that changing composition of states do not drive the dynamic treatment effects.

As an additional robustness check, I also estimate specifications that include county population weights. Inclusion of these weights helps to determine whether the policy has heterogeneous effects based on local population. For example, if the policy has larger effects in higher population counties, then WLS estimates that put greater weights on these more populous counties will be larger than the corresponding OLS estimates.The results are presented in tab:TableA12 and are highly similar to the baseline results. This seems to suggest that treatment effects of the policy are not significantly stronger in more populous counties.

### C.3 Propensity score matching for urban and rural treated counties

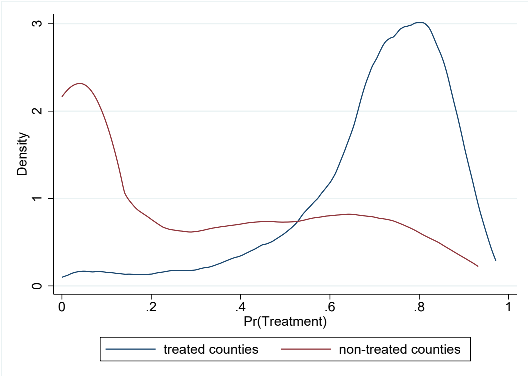
When I include the treated states of New Jersey, Rhode Island, Washington DC and Delaware with their corresponding treated counties being present in urban areas, I use propensity score matching to choose comparable untreated counties from the pool of both rural and urban counties. I first estimate the propensity score for being selected to receive the treatment. For this purpose, I use an iterative procedure (Imbens and Rubin, 2015) to select covariates from a rich set of pre-treatment county variables, measured in the year before the policy, to include in the propensity score model. [78] After selecting covariates using the above method, I estimate the following logit specification for calculating the propensity scores:

$$Logit(Pr(Treated_c)) = \beta_0 + \beta_1 X_c \quad (9)$$

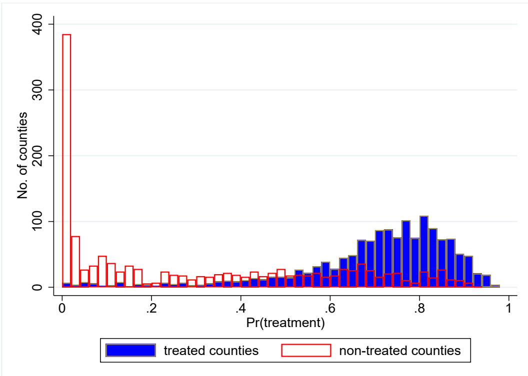
The dependent variable is whether county c in state s has implemented the policy during my sample period. $X_c$ consists of the vector of county level covariates chosen by Imbens and Rubin method. These include % of total population in different age groups, population density, % of population in poverty, unemployment rate, median household income, total employment, % of male population in different age groups, % of white population of age groups 20-39, 40-64, % hispanic population of age groups 40-64, 65 and above, % black population of age groups 40-64, 65 and above, quadratic of % population in poverty, quadratic of % of total population in age groups 20-39, 40-64 and above 65 years. Figure A40 shows the distribution of estimated propensity scores using kernel density and histogram. I drop observations with propensity scores outside the common support $(8.1 * 10^{-8}, 0.932)$. After trimming the sample, I estimate equations 1 and 4 using $T + (1 - T) * \frac{p}{1-p}$ as weights, where T is the indicator for treatment and p is the estimated propensity score. This weighting method produces a consistent estimator of average treatment effect of treated counties(Imbens, 2004).The results using the reweighted sample are presented in tab:TableA13.

[78]The list of original covariates inputted into the model is available from the author on request.

Figure A40: Propensity score distribution : Urban and rural treated counties




(a) Kernel density




(b) Histogram

*Notes:* Main data source: Area Health Resource File, Bureau of Labor Statistics, Surveillance, Epidemiology and End Results (SEER) program, National Cancer Institute for the pre-treatment i.e. 'year before policy' period.

# D Data Appendix

Below, I briefly describe the process of linking the data scraped from various websites with AMA data. This linkage is done for the current year based on publicly available information like medical and training details, current office location, current speciality, gender, date of birth. I employ several rounds of fuzzy matching to link the datasets, and manually checking each match for false positives. After all the rounds, the few unmatched records are matched manually, again using a sufficient set of the above attributes. Once the datasets are linked and re-identified for the current year, they can be linked for all the previous years because of masked identifiers provided in AMA data linking physicians across the years.

# E Propensity score matching for California health outcomes

I fit a propensity score model to select control counties in California. To select variables for inclusion in the logit model for propensity scores, I use the double lasso procedure of Belloni et al(2014a) and Urminsky et al(2016). [79] The double lasso procedure prevents overfitting of the propensity score model and allows for imperfect selection of controls in either of the above two steps. It provides an efficient data-driven method to select a subset of influential confounders from a broader set of potentially confounding variables.

After following the above data dependent covariate selection method, I estimate the following logit model:

$$Logit(Pr(Treated_c)) = \beta_0 + \beta_1 X_c \tag{7}$$

The dependent variable is whether the hospital in county c has implemented the policy in 2003. $X_c$ consists of the vector of variables (both county level and hospital level) chosen by the data driven procedure.[80] These variables include patient demographics like age, share female, share white at a hospital level. It also includes county level covariates like per capita personal income, median household income, log population, log population density, percentage of population in age groups ranging between 20-85 plus, percentage of black population, percentage of female population, percentage of hispanic population, percentage of white population, percentage of population in poverty, unemployment rate. The distribution

[79] I estimate each of the steps using square root lasso described in Belloni et al(2014a) and Belloni et al(2014b). The original set of pre-policy variables inputted into the model include county level variables like: unemployment rate, poverty rate, real median household income, per capita personal income, population, population density, uninsured rate for non-elderly adults, percentage of population in 5-year age groups ranging between 20-65 plus, percentage of male population, percentage of black population, percentage of hispanic population, percentage of white population, percentage of female population, percentage of non hispanic population and hospital level variables like: share white, share female, average age of admitted patients, number of diagnoses, case mix index.

[80] Increasing the lasso penalty parameter from the optimal level results in fewer variables selected in $X_c$. In that case, the estimated propensity scores do not accurately adjust for pre-existing differences between treated and control units, reducing the power to detect effects on admissions.

of propensity scores is depicted in Figure A41. The treated and control hospitals (counties) display significant overlap for almost the whole distribution, except some non-treated hospitals being present at the very lower tail of the distribution. I drop those observations in the tails where there is no overlap (propensity scores outside the range of 0.048 and 0.998). This removes 47 hospitals from the pre-treatment sample, 4 in the treated counties and 43 in the control counties. The final sample consists of 12 hospitals in treated counties and 280 hospitals in control counties. [81] Despite the fact that trimming causes the estimand to be applied to a subsample of matched counties, it has advantages. There is improvement in the asymptotic variance and robustness properties of the estimator by omitting unmatched units (Imbens and Rubin (2015)). Additionally, the excluded units make the estimators sensitive to outliers (Young, 2018).

# F Heterogeneity by baseline physicians per capita

I consider the following equation:

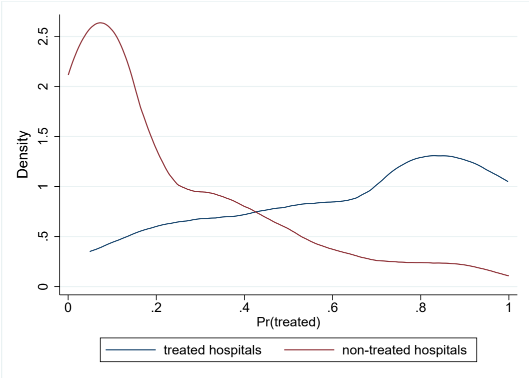
$$Y_{ct} = \sum_{k=-6,k\neq-1}^{+9} \alpha_k \mathbb{I}(t = D_c{+}k) + \sum_{k=-6,k\neq-1}^{+9} \phi_k \mathbb{I}(t = D_c{+}k)(Baseline\ physicians/10,000\ population)_c$$
$$+ \beta X_{ct} + \gamma_c + \delta_{st} + \epsilon_{ct} \quad (8)$$

In the above equation, the outcome variable is number of deaths per 10,000 population. Baseline physicians is number of physicians in the year before the policy. The coefficient of interest is $\phi_k$ on the interaction between policy dummy $D_c$, the indicator function $\mathbb{I}(t = D_c{+}k)$ and the baseline number of physicians per 10,000 population in county c. The coefficient $\phi_k$ gives the relative treatment effect of more current physicians in places having more baseline physicians. A negative estimate suggests that adding a physician reduces mortality more in a place with many physicians at baseline than it does in a place with fewer at baseline. All the other terms in the above equation are similar to equation (1). The results of equation (8) hold up if I consider the alternative specification where I divide the baseline physicians per capita into above median (high density areas) and below median (low density areas) and examine the number of deaths separately in high density and low density areas.

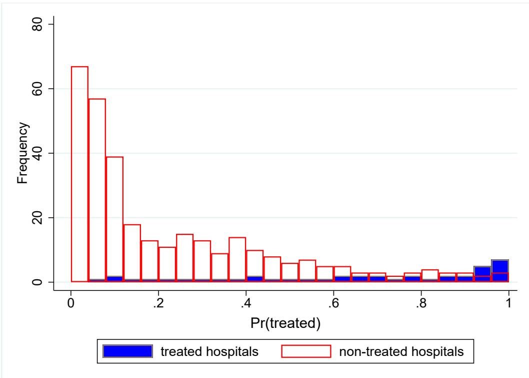
The pooled estimates reported in tab:TableA22 echo the event study findings. Column

[81] 12 and 280 hospitals denote the unique number of hospitals each year.

Figure A41: Propensity score distribution : California patient level records




(a) Kernel density




(b) Histogram

*Notes:* Main data source: OSHPD patient discharge data at the day level, Area Health Resource File, Bureau of Labor Statistics, Surveillance, Epidemiology and End Results (SEER) program, National Cancer Institute for the pre-treatment period 1999-2002.

1 implies that among treated counties, adding a physician to a place with more physicians at baseline reduces overall mortality by an additional 0.1251 deaths per 10,000 population relative to a place with fewer physicians at baseline. This effect is largely accounted for by amenable cause mortality , with an additional drop of 0.1045 deaths per 10,000 population in a treated place with more baseline physicians compared to a place with fewer baseline physicians. The detailed cause of death analysis in columns 3-7 suggest that cardiovascular deaths of elderly decline by 0.056 per 10,000 population in a treated place with higher baseline physicians. In contrast, I do not find evidence of significant decrease in respiratory mortality,

suicides, drug poisoning and opioid overdose deaths of elderly in treated places having more baseline physicians. To ascertain the precision of the null effects, I consider the range of effects that can be eliminated by the point estimates. In case of respiratory disease, I can reject mortality declines outside (-0.1110,0.0814) per 10,000 population while in case of suicides, drug poisoning and opioid overdoses, effects outside (-0.0225,0.0194); (-0.0104,0.0068) and (-0.0113, 0.0079) per 10,000 population respectively can be ruled out (see columns 4-7).

# G Calculation of geographic distance

I calculate the distance between the centroids of each county in my sample and the nearest largest metro county of that state. As per USDA Rural-Urban continuum codes, these metro counties can be counties in metro areas of 1 million population or more, counties in metro areas of 250,000 to 1 million population or counties in metro areas of fewer than 250,000 population(in small states like Montana), depending on the state of practice of the physician. I calculate the great circle distance between the centroids using the haversine formula and then identify the minimum distance between a county in my sample and the largest metro county within the state. This distance is then converted to miles. As a robustness check, I also calculate the distance in statute miles between the centroids by computing the length of the great circle arc, located on the surface of a sphere, in Matlab. I obtain largely similar results by doing so.

I obtain the centroids of the counties from NHGIS shapefiles and convert them to WGS1984 projection in ArcGIS. Recall that, a consistent set of counties is followed over the sample period, with county boundary changes taken into account. Any missing latitudes and longitudes during the merging of the coordinates file to the main data file is filled in manually.[82]

# H Construction of artificial amenity index

In this section, I first describe the components used to construct the amenity index. Retail environment includes clothing stores, eating and drinking places, amusement and recreation facilities per capita. Environmental quality includes expenditure on parks per capita, median

[82]see, for example, https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.2000.html.

air quality index (EPA), number of good, moderate and unhealthy days (as per EPA). Transportation amenities comprises transport infrastructure (highways, airports, parking spots) and transportation utilities(passenger transit, bus charter service, school bus, taxicabs, bus terminal and service, rural bus transportation etc.) per capita. Education amenities consist of county level expenditure per pupil, government expenditure on elementary, secondary education and libraries per capita. Health amenities include government expenditure on hospitals and other health facilities per capita. Crime includes number of police officers, number of violent crimes and number of property crimes per capita. I use principal component analysis (PCA) to combine these different categories of amenities with different scales into a single dimension.

In the below table A24, I report the weights on each amenity. Overall, this index appears to accurately capture an area's attractiveness in terms of non-natural amenities. Additionally, I also find that counties closer to the large cities within the state (eg: Fountain, IN; Jefferson, KS; Allegany, NY; Cayuga, NY; Montgomery, NC; Clarion, PA) have higher values of the index than counties in the interior. This supports an area's desirability as seen through the index's lens.

# I Conceptual Framework

This section provides a conceptual framework for understanding when provision of information in the form of introduction of a loan repayment program, increases the hiring rate of those medical trainee applicants who have the highest debt. I build on the simplified framework of Bartik and Nelson(2021), adapted to my setting where an information source is added, rather than removed, to ease the debt of medical trainees. [83]

Suppose an employer's pool of applicants i is drawn from groups $g_1$ (trainees with higher debt) and $g_2$ (trainees with lower debt). For both groups, the potential match quality $\mu$ follows a uniform distribution on the unit interval.

[83] I will abstract from applicants' costs to acquire information about the policy, as it is not central to my analysis. Moreover, these costs are likely to be low for in-state medical trainees relative to out-of-state trainees, because of widespread availability of the program information on their medical school or training institution websites. Recall that, I show a large increase in the proportion of medical trainees who choose to remain in their training state after the implementation of the program.

Table A24: Principal Component Analysis for non-natural amenity index

| | (Loading) | (Unexplained variance) |
|---|---|---|
| | (1) | (2) |
| *Panel A: Retail index* | | |
| Eating places per 1000 residents | 0.6167 | 0.3451 |
| Apparel stores per 1000 residents | 0.4920 | 0.5831 |
| Recreation & amusement facilities per 1000 residents | 0.6145 | 0.3499 |
| *Panel B: Environment index* | | |
| EPA median air quality index | -0.6199 | 0.2075 |
| Good days (EPA) | 0.6008 | 0.3057 |
| Moderate days (EPA) | -0.4298 | 0.619 |
| Unhealthy days (EPA) | -0.2540 | 0.867 |
| Expenditure on parks | 0.0743 | 0.9386 |
| *Panel C: Transportation index* | | |
| Transport services per 1000 residents | 0.7071 | 0.3227 |
| Transport infrastructure per 1000 residents | 0.7071 | 0.3227 |
| *Panel D: Education index* | | |
| Expenditure per pupil | 0.1957 | 0.9237 |
| Government expenditure on elementary & secondary education | 0.6935 | 0.0421 |
| Government expenditure on libraries | 0.6934 | 0.0424 |
| *Panel E: Health index* | | |
| Government expenditure on hospitals | 0.7071 | 0.0519 |
| Government expenditure on other health facilities | 0.7071 | 0.0519 |
| *Panel G: Crime index* | | |
| Property crimes per 1000 residents | 0.7071 | 0.2641 |
| Violent crimes per 1000 residents | 0.7071 | 0.2641 |
| *Panel H: Aggregate index* | | |
| Retail index | -0.0722 | 0.9384 |
| Environment index | 0.1300 | 0.9125 |
| Transportation index | 0.4079 | 0.6306 |
| Education index | 0.6272 | 0.1766 |
| Health index | 0.6205 | 0.1952 |
| Crime index | -0.1818 | 0.9267 |

Notes:All amenity variables are in logs. In panels A-G, I report loadings used in construction of each subindex. In Panel H, I report loadings on each subindex to construct the aggregate non-natural amenity index.

$$\mu_{i,g_1} \sim U(0,1) \tag{9}$$

$$\mu_{i,g_2} \sim U(0,1) \tag{10}$$

An employer wants to hire a share $\gamma$ of applicants, and I assume that it hires those with the highest expected match quality.[84] I assume wages of newly hired physicians are fixed, so

[84]This may be a simplified assumption in a few states, with the employer selecting among the eligible candidates based on the priority of receipt of application. Otherwise, I believe this is a reasonable assumption

that firms do not prefer to hire low quality matches at lower wages.[85] The applicants have two signals, denoted by s, to reveal their match quality. [86] For group $g_1$, signal s has a probability $p_{g_1,s}$ of perfectly revealing the trainee's type and a probability 1-$p_{g_1,s}$ of providing zero information about the trainee's type.[87]

Below, I show how the two groups' hiring rates change, when employers have an additional signal available due to the roll-out of the loan repayment program. This signal is in the form of applicants disclosing the debt amount to the employer to be eligible for the policy. As mentioned in the institutional details in the main text, the design of this program is such that low debt trainees receive the full amount of their debt, while higher debt trainees receive only partial funding, bounded above by the benefits offered by the state of the practicing physician. Hence, it seems that this temporary program is more likely to provide higher incentives to relatively lower debt medical trainees to apply for it. Suppose the additional signal is denoted by s=2. The change in both groups' hiring rates depends on how the other available signals differs across the groups. The hiring rate of group $g_1$, before the policy, when only signal 1 is available, is: $P(hired = 1|s = 1) = P(hired = 1) * P(s = 1|hired = 1)$. Therefore, [88]

$$\lambda_{g1,1} = (\frac{2\gamma}{p_{g_1,1} + p_{g_2,1}}) * p_{g_1,1} \tag{11}$$

The first term on the right hand side shows the share of the applicant types above the hiring threshold. The second term, on the other hand, shows the probability of an applicant, who is above the hiring threshold, to have his type revealed.

The hiring rate of group $g_1$, after the policy, when both signals 1 and 2 are available, is: $P(hired = 1|s = 1 \ or \ s = 2) = P(hired = 1) * P(s = 1 \ or \ s = 2|hired = 1)$.

$$\lambda_{g1,2} = (\frac{2\gamma}{p_{g_1,1} + p_{g_2,1} + (1 - p_{g_1,1})p_{g_1,2} + (1 - p_{g_2,1})p_{g_2,2}}) * (p_{g_1,1} + (1 - p_{g_1,1})p_{g_1,2}) \tag{12}$$

to explain the empirical findings.

[85] I present suggestive evidence in section 5.6 that the loan repayment program does not have a significant effect on a physician's wages.

[86] I do not talk about the bias of the signal here because my empirical estimates do not indicate that the signal disproportionately favors any particular gender or race/ethnicity.

[87] This extreme "all or nothing approach" may not be realistic, but serves my purpose of illustrating the importance of relative precision of a new signal in hiring.

[88] This expression depends on the assumption that employers do not hire applicants who provide no signal or whose signal has no information content.

The first term denotes the share of applicants above the hiring threshold when both signals 1 and 2 are available. The second term reflects the probability that signal 1 or 2 reveals the applicant type.

Relative to equation 14, the change in denominator in the first term on the right hand side of equation 15, represents the effect of an endogenously higher hiring threshold when signal 2 becomes available. The change in the second term shows the effect of more applicants from group $g_1$ indicating their type to be above that higher threshold. Subtracting equation 15 from equation 14 and after simplification, group $g_1$'s hiring rate can increase after the policy iff the following condition holds:

$$\frac{p_{g_1,2}}{p_{g_2,2}} > \frac{\left(\frac{p_{g_1,1}}{1-p_{g_1,1}}\right)}{\left(\frac{p_{g_2,1}}{1-p_{g_2,1}}\right)} \tag{13}$$

Equation 16 shows that the rollout of the loan repayment program will increase group $g_1$'s hiring rate if and only if the ratio of signal 2's probabilities for group $g_1$ relative to group $g_2$ is greater than the corresponding odds ratio for signal 1. This condition reflects the relative advantage of group $g_1$ in signal 2. If $p_{g_1,2} > p_{g_2,2}$ i.e. signal 2 is more informative for higher debt applicants than lower debt applicants, then hiring rate of higher debt applicants will increase only if the baseline signals are either equally informative for higher debt and lower debt applicants ($p_{g_1,1} = p_{g_2,1}$) or these signals are less informative for higher debt applicants relative to lower debt ones ($p_{g_1,1} < p_{g_2,1}$). This is the standard intuitive case.

Condition 16 can also hold even if $p_{g_1,2} < p_{g_2,2}$, as long as $p_{g_1,1} << p_{g_2,1}$. Intuitively, this implies that if baseline signals for higher debt applicants are more noisy than for lower debt applicants, then addition of signal 2 can increase hiring of higher debt applicants even if signal 2, by itself, provides lower information about them. Signal 2 can fail to deliver accurate information about the debt level of applicants if very high debt ones under-report their loans, to be eligible for the policy. The employer is then unable to precisely distinguish between low-debt and high-debt applicants. I provide empirical evidence in support of more hiring of higher debt applicants like foreign physicians and physicians from the bottom ranks of US medical schools, after the introduction of signal 2. This can possibly be explained by the relative advantage of these higher debt trainees in signal 2– the higher debt trainees are more disadvantaged under the baseline screening tools relative to lower debt trainees. In contrast, some progressively higher debt applicants from unranked US medical schools experience lower hiring rates after the program, presumably because these applicants have a higher advantage under the baseline signals compared to the lower debt ones. Either, these highest debt trainees do not apply for the program, or they are not hired by the employer, as their signals are uninformative and employers have limited budget. Thus, the introduction of the policy does not help in their hiring process. The fact that not all higher debt applicants' hiring rates go up after signal 2, indicates that precision of signal varying by individuals, matters more for their successful hiring. [89] The baseline signals i.e. signal 1 can consist of screening tools for reimbursement based programs or other productivity shocks which do not mandate disclosure of loan amounts. These signals are then more likely to convey muddled information about the debt level of trainees.

In the extreme case of baseline signals either conveying no information for the higher debt applicants or complete information for the lower debt ones i.e. $p_{g_1,1} = 0$ or $p_{g_2,1} = 1$, hiring rates of higher debt applicants will increase after the policy as long as the policy reveals any information about those applicants i.e. $p_{g_1,2} > 0$. The above framework can be extended to consider changes in the hiring rates of treated speciality trainees compared with untreated speciality ones as they navigate through their program years. As per the predictions of this framework, treated speciality residents will see an increase in hiring rates if and only if signal 2 conveys relatively more precise information about them as compared to signal 1. This claim in borne out in the data too. Intuitively, if signal 1 cannot differentiate between treated primary care specialities and untreated non-primary care specialities (with both of them being equally eligible), the treated group can possibly benefit from the introduction of signal 2, even if signal 2 fails to provide precise information about the debt type of the applicants.

The main takeaway of the framework is that, the hiring rates of a particular group can increase, if that group has a relative and not necessarily an absolute advantage, in the information content of the new signal vis-a-vis that of baseline signals.

[89] In other words, the group mean of the posterior distribution matters less relative to the variance of the posterior.