

NBER WORKING PAPER SERIES

CAN DIFFUSE DELIVERY SYSTEM REFORMS IMPROVE POPULATION HEALTH?
A STUDY OF THE STATE INNOVATION MODELS INITIATIVE

Partha Deb
Anjelica Gangaram
Hoda Khajavi

Working Paper 26360
<http://www.nber.org/papers/w26360>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2019

We thank participants at the 8th Annual Conference of American Society of Health Economists, especially Muzhe Yang and Rui Li, and seminar participants at the Office of the Assistant Secretary for Planning and Evaluation (ASPE) for their insights that helped us improve this paper. We thank Jun Li for providing a number of substantive and editorial comments. No funding or compensation was received for this particular study. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Partha Deb, Anjelica Gangaram, and Hoda Khajavi. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Can Diffuse Delivery System Reforms Improve Population Health? A Study of the State Innovation Models Initiative

Partha Deb, Anjelica Gangaram, and Hoda Khajavi

NBER Working Paper No. 26360

October 2019

JEL No. I1,I18

ABSTRACT

We examine the effects of the State Innovation Models (SIM) on population-level health status. The SIM initiative provided \$250 million to six states in 2013 for delivery system reforms. We use data from the Behavioral Risk Factor Surveillance System for the years 2010 -- 2016 to compare health of the populations in 6 SIM states to 15 states that were not involved in any aspects of SIM. We examine changes in health using an event study design. We develop a Latent Class Profile model that takes multiple measures of latent health into a common, latent health status to study the effect of the intervention. Such models can yield informative estimates where separate estimation of measures do not. We find that individuals in states that implemented SIM saw significant improvements in health across a number policy-relevant subpopulations.

Partha Deb
Hunter College
Department of Economics
695 Park Avenue
Room 1524 West
New York, NY 10065
and NBER
partha.deb@hunter.cuny.edu

Hoda Khajavi
Hunter College
hnourikhajavi@gradcenter.cuny.edu

Anjelica Gangaram
University of Michigan
1415 Washington Heights
Ann Arbor, MI 48109
agangara@umich.edu

1 Introduction

Much media and academic attention has been placed on the provisions of the Affordable Care Act (ACA) aimed at expanding health insurance coverage. However, a major component of the law is aimed at providing innovative healthcare delivery and financing mechanisms to improve patient care, population health, and reduce healthcare costs. Under these auspices, the Center for Medicare and Medicaid Innovation (CMMI), a division of the Centers for Medicare and Medicaid Services (CMS), was established to test interventions in payment and service delivery models.

One of these innovations is the State Innovation Models (SIM) initiative, which provides grants to states to design and test alternative healthcare and payment systems. SIM is a large, diffuse, delivery system reform intervention. It provided \$250 million to six states in 2013 for providers to shift toward value-based payment models and invest in infrastructure and training for providers to transform primary care and integrate behavioral health. No specific healthcare sectors were specified, nor were specific approaches to transformation outlined a priori. States were asked to submit plans for large-scale changes in healthcare provision and financing. States are ideally suited to design and test such alternative, large-scale mechanisms because they have regulatory authority to convene both public and private stakeholders.

In 2012, CMMI invited states to submit SIM proposals for the first round of model design or model testing awards. Model testing awards were intended for states with fully developed proposals for state-wide health system transformation. CMMI provided funding for states to implement their innovation plans based on proposals that described each state's vision for health system transformation and strategies the state intended to use to implement and test its model. On April 1, 2013, six states – Arkansas, Maine, Massachusetts, Minnesota, Oregon, and Vermont — were awarded a total of \$250 million to implement and test their innovation plans over a period of 3.5 years. These states were chosen because they built upon existing CMS approaches such as Accountable Care Organizations (ACOs), Medicaid health homes,

and/or bundled or episode-based payment models. All six states included Medicare, Medicaid, Children’s Health Insurance Program (CHIP), and commercial payers in their innovation plans. Most innovation plans emphasized improving care for high-risk, high-cost populations. See Center for Medicare and Medicaid Innovation (2019) for details of the initiative, for information on subsequent rounds of the initiative and a compendium of relevant materials.

In this paper, we examine the effects of the first round of the SIM initiative on health outcomes in the six states that implemented their innovation plans as compared to 15 states (Alabama, Alaska, Florida, Georgia, Indiana, Kansas, Louisiana, Mississippi, Missouri, Nebraska, North Carolina, North Dakota, South Carolina, South Dakota, and Wyoming). The 15 control states were not involved in the first round of model design initiatives, nor designated for the second round of model testing awards.

We take a population focus in our study (Hughes et al., 2015; Auerbach et al., 2013). If the numerous healthcare delivery infrastructure changes undertaken by states that participated in SIM were meaningful, the initiative should have resulted in population-level effects. We use data from the Behavioral Risk Factor Surveillance System (BRFSS) for the years 2010 - 2012 and 2014 - 2016 to examine the effects of SIM. We exclude data from 2013 because that was the year in which states learned of their awards and worked with CMMI to obtain the funding necessary to begin implementation. We use an event study approach (alternatively known as a generalized interrupted time series approach) to study the effect of the SIM initiative. This statistical design is very similar to a difference-in-difference model but the generalized interrupted time series approach is agnostic about trends in the pre-program or baseline period. We choose to be agnostic because we cannot confidently argue for common trends prior to 2013.

In our application of the generalized interrupted time series model, we treat 2012 as the base year and estimate differential trend coefficients for each of the other years. Given the numerous changes to the healthcare landscape induced by the ACA, we

prefer not to assume, a priori that the health trends in these six treated states were the same as those in the control states. The ACA, and the healthcare landscape more generally, induced many changes in healthcare delivery and financing, in the public and private sectors, across states. These changes may or may not have differentially affected the states in our study. It is unlikely that other state level innovations had the scope and scale of SIM, nevertheless we err on the side of caution. Our modeling strategy allows us to judge, ex post, whether there was a significant shift in the trend after implementation relative to the trend before it.

Health status is a notoriously difficult construct to measure in part because it is multidimensional and latent. Surveys such as the BRFSS typically have a number of measures that plausibly capture components of health status. Instead of estimating models for individual measures, we develop a Latent Class Profile (LCP) model that uses three measures of health status simultaneously, assuming that the regression model is for an underlying common latent measure of health status (Chung et al., 2011; Oberski, 2016). The measures of health include an ordinal measure of health status, which we assume follows an ordered logistic distribution, a count of days affected by physical health problems, and a count of days affected by mental health problems, which we assume follow Poisson distributions. These outcomes are jointly modeled as being driven by a common discrete latent factor, health. We then assume that the set of observed covariates, including the treatment design covariates and a set of demographic and socioeconomic characteristics, affect the latent factor directly, as in Chung et al. (2011), rather than individual measures of outcome.

The ability of the latent class profile model to isolate the common factor in the measures of health status yields substantial gains in statistical power. In situations where only small effects can be plausibly expected, modeling multiple measures of a common underlying outcome separately can yield inconsistent (not in a statistical sense) estimates across equations. The results of our empirical analysis are consistent with these features – we find consistent improvements in health due to SIM in the

LCP framework.

We are not aware of any peer-reviewed literature evaluating this large initiative. A contractor for CMMI has been evaluating the effect of the implementation of these models in these six states, drilling down to specific subpopulations and examining the effects in each state separately (RTI International, 2016, 2017, 2018). Findings in their reports suggest that the SIM initiatives in these states have produced considerable movement towards value-based purchasing models of payment in Medicaid and commercially insured populations. Clinicians in these states have credited SIM-related investments in health information exchanges and quality measurement as key determinants of changes in care delivery intended to improve care coordination. There is consistent evidence of better care coordination in primary care and physician services and some improvements in screening measures and medication adherence outcomes. These improvements did not lead to fewer ED visits or hospital admissions, except in Vermont and Arkansas among Medicaid beneficiaries.

Although the effects of SIM on health are unknown, research on other CMMI initiatives is suggestive of the virtues of such designs. For instance, in a group of CMMI initiatives known as Episode-based Payment Initiatives, CMS holds health-care providers accountable for the services they provide to Medicare beneficiaries during an episode of care. Dummit et al. (2016) show that when hospitals were held accountable for total payments and quality of care during an episode of total joint replacement, average Medicare payments decreased by \$1166, while the quality of care was maintained. In an other approach, known as Accountable Care Organizations (ACOs), CMMI provided incentives for healthcare providers, including physicians and hospitals, to create networks to provide a coordinated care to Medicare beneficiaries. McWilliams (2016) shows that ACOs that joined the program between 2012 and 2014 generated \$287 million net savings for Medicare.

The remainder of the paper is organized as follows. In the next section, we describe the State Innovation Models in more detail. In section 3, we describe the latent class

profile model formally. The data and samples are described in section 4 and results are described in section 5. We discuss our findings and conclude in section 6.

2 Background

The State Innovation Models (SIM) initiative, promulgated by the Center for Medicare and Medicaid Innovation (CMMI), a division of the Centers for Medicare and Medicaid Services (CMS), provides states with funding, learning tools and expert technical assistance to modernize their multi-payer healthcare payment systems and to improve the delivery of healthcare. “The distinctive purpose of the SIM initiative is to test whether new models with potential to improve care and lower costs in Medicare, Medicaid, and CHIP will produce better results when implemented in the context of a state-sponsored plan that involves multiple payers, broader state innovation, and larger health system transformation to improve population health” (Van Vleet and Paradise, 2014).

The CMMI sought out the first round of state innovation proposals in the summer of 2012, and on the basis of these proposals presented states with one of three awards: a model design award, a model pre-test award, or a model test award. Model design and pre-test awards provided support for states to develop and design their own state innovation plans. Model design awards were given to sixteen states: California, Connecticut, Delaware, Hawaii, Idaho, Illinois, Iowa, Maryland, Michigan, New Hampshire, Ohio, Pennsylvania, Rhode Island, Tennessee, Texas, and Utah. Three states, Colorado, New York, and Washington, received pre-test awards. Arizona, California, Colorado, Connecticut, Delaware, Hawaii, Idaho, Illinois, Iowa, Kentucky, Maryland, Michigan, Montana, Nevada, New Hampshire, New Jersey, New Mexico, New York, Ohio, Oklahoma, Pennsylvania, Rhode Island, Tennessee, Utah, Virginia, Washington, West Virginia, and Wisconsin took part in round two of the SIM initiative as either model test or model design states. No pre-test awards were awarded during round two of the initiative.

The model test awards allowed states to “implement and test strategies” outlined in their proposals. The six states that received model test awards in Round 1 were Arkansas, Maine, Massachusetts, Minnesota, Oregon, and Vermont. The awards were announced on February 21, 2013 and the performance period for these awardees began on April 1, 2013. Our paper focuses on the impact that these state innovation plans had on the health of their populations.

As previously mentioned, SIM interventions designed to implement delivery system reforms varied across each of the six test states. The states focused their innovation efforts on Medicaid, CHIP, and commercial payers. They also implemented or had plans to implement interventions that targeted Medicare. Oregon also included its state employee healthcare system in its innovation plan. However, since each model test awardee developed its innovation plan independently, differences as well as commonalities can be found in the healthcare delivery system and payment reform strategies they employed. Each of the test states included and built upon state specific strategies that predated SIM, in addition to implementing new strategies in their innovation plans. There is no easy way to classify the test states based upon their innovation plans because the six test states did not implement sufficiently similar strategies, and the pre-SIM landscape was different for each of the test states.

The model test awardees used a two-pronged approach with regards to healthcare delivery reform. All the states, following SIM guidelines, emphasized improving primary care and integrating primary care with additional healthcare services. Patient-centered medical homes (PCMH) are primary care facilities that focus on preventative care and chronic condition management. PCMH were central to the states’ aim to improve primary care and all six test states implemented strategies to improve existing PCMH or create new ones. Additionally, Arkansas and Maine used SIM resources to fund new patient-centered health homes. These health homes focused on coordinating and improving physical and behavioral health for medically complex, high-risk, high cost Medicaid populations. Maine also instituted behavioral health homes that pro-

vide mental healthcare in addition to primary care. With the exception of Arkansas, the test states focused reforms within ACOs; these organizations are focused on providing coordinated care to patients. Arkansas, Maine, Minnesota, and Oregon also devised new work force models that incorporated team based care and nontraditional healthcare workers. Patients may be treated by a team of healthcare professionals or obtain information and resources from community health workers or peer wellness specialists.

The six SIM states did not limit their healthcare delivery system reform efforts to primary care. They also worked to connect primary care services to other health services with the goal of allowing patients to experience better overall care. These states varied in their scope and scale of integration of one or more of the following types of services with primary care: speciality care, oral healthcare, behavioral healthcare, long term care, and social services.

In addition to improving healthcare delivery, the states used a variety of strategies to reduce healthcare costs. As a result, the test states implemented a variety of alternative payment strategies. The states used a combination of three or more of the following payment models: per-member-per-month payments, bundled payments, bonus payments, shared savings, shared savings and risk, prospective payment, and partial/global capitation. These payment models were meant to link the payment for services to the efficacy and quality of care patients receive, not the number of procedures received.

3 Econometric model

We analyze three measures of health status jointly using a Latent Class Profile (LCP) model (Lazarsfeld and Henry, 1968; Chung et al., 2011). The measures are described in detail in section 4 below. The LCP model assumes that the measures are co-determined by a common discrete latent factor. We assume that the set of observed covariates, including the treatment design covariates and a set of demographic and

socioeconomic characteristics affect the latent factor directly.

The conceptual structure of the LCP model was originally developed in Lazarsfeld and Henry (1968). It is closely related to the latent class model, which is the nomenclature used when the measurements are binary (Goodman, 1974). It is also closely related to the Grade of Membership Model (Manton et al., 1992; Portrait et al., 1999) for multiple binary measurements. Erosheva (2005) and Oberski (2016) provide overviews of these models. Erosheva (2005) derives a unifying finite mixture model framework for these models.

To be more specific, we use three measurements: a count of days affected by physical health problems, p_i , a count of days affected by mental health problems, m_i , and self-reported health status on a five-point multinomial (ordinal) scale, s_i . Suppose that the population can be divided into C classes or subpopulations on the basis of their underlying, latent health status. For an individual in class $c \in 1, 2, \dots, C$, let the counts of days affected by physical and mental health problems each follow a Poisson distribution so that

$$f(p_i) = \frac{\psi_c^{p_i} \exp(-\psi_c)}{p_i!} \quad (1)$$

and

$$f(m_i) = \frac{\mu_c^{m_i} \exp(-\mu_c)}{m_i!}, \quad (2)$$

where ψ_c and μ_c are class-specific unknown parameters that determine the Poisson distributions. We assume that the ordinal measure of health status, that takes values $j = 1, 2, \dots, 5$, follows an ordered logistic distribution so, for an individual in class c ,

$$g(s_i(j)) = \frac{1}{1 + \exp(-\xi_{j,c})} - \frac{1}{1 + \exp(-\xi_{j-1,c})} \quad (3)$$

where $\xi_{j,c}$ are class-specific threshold parameters with $\xi_{0,c}$ defined as $-\infty$ and $\xi_{5,c}$ as $+\infty$. It is important to note that these class-specific distributions are not parameterized with covariates. By keeping these vectors of values fixed (in expectation) within each class, the underlying health class has an unconditional interpretation. Also see Chung et al. (2011) for further discussion.

Let π_c denote the probabilities of membership in each class of health status. The classes are non-trivial so that $0 < \pi_c < 1$. In addition, they are mutually exclusive and exhaustive so that $\sum_{c=1}^C \pi_c = 1$. Then, the probability density function for a C -component finite mixture (Deb and Trivedi, 1997; McLachlan and Peel, 2004), is denoted by

$$f(p_i, m_i, s_i) = \sum_{c=1}^C \pi_c f(p_i; \psi_c) f(m_i; \mu_c) g(s_i; \xi_{0c}, \xi_{1c}, \dots, \xi_{5c}). \quad (4)$$

Let \mathbf{z}_i denote a vector of observed covariates including the design vector. We parameterize the class probabilities with covariates using a multinomial logit functional form, i.e.,

$$\pi_c(\mathbf{z}_i; \boldsymbol{\theta}_c) = \frac{1}{1 + \exp(-\mathbf{z}_i \boldsymbol{\theta}_c)} \text{ for } c = 1, 2, \dots, C \quad (5)$$

so that

$$f(p_i, m_i, s_i) = \sum_{c=1}^C \pi_c(\mathbf{z}_i; \boldsymbol{\theta}_c) f(p_i; \psi_c) f(m_i; \mu_c) g(s_i; \xi_{0c}, \xi_{1c}, \dots, \xi_{5c}). \quad (6)$$

Denote the time periods of data in the study as $t = T_0, \dots, 0, \dots, T$ where T_0 denotes the first year of the study, 0 denotes the period before the treatment begins and T denotes the final period of data available for analysis. Let $I(t)$ denote the indicator for time t . In addition, let A_{it} denote the treatment status of observation i at time t ; $A_{it} = 1$ if the observation belongs to the treatment group and $A_{it} = 0$ if it belongs to the control group. Note that treatment assignment is designated by the state of residence (S_{it}) of the individual. We specify $\mathbf{z}_i \boldsymbol{\theta}_c$ using an interrupted time series or event study design:

$$\mathbf{z}_i \boldsymbol{\theta}_c = \alpha_c + \sum_{t=T_0}^T \beta_{tc} A_{it} I(t) + \mathbf{x}_i \boldsymbol{\gamma}_c + \boldsymbol{\delta}_{sc} + \boldsymbol{\zeta}_{tc} + \boldsymbol{\kappa}_c S_{it} \Omega(t) \quad (7)$$

where $\boldsymbol{\delta}_{sc}$ and $\boldsymbol{\zeta}_{tc}$ denote state and time fixed effects in class c respectively. State-specific time trends are introduced via $S_{it} \Omega(t)$, where Ω denotes a polynomial in t . We also allow for linear state-specific time trends. The vector \mathbf{x}_i includes a number of covariates described in section 4 and month (of survey) fixed effects.

For our empirical analysis, we assume that health status can be categorized into two classes. We estimate the parameters of the 2-class LCP model specified by equations 6 and 7 using maximum likelihood. We conduct inference based on standard errors adjusted for clustering at the state level. Post estimation, we estimate the sample average time effects, τ_{tc} on the class probability, at time t in class c , as

$$\tau_{tc} = \frac{1}{1 + \exp(-[\alpha_c + \sum_{t=T_0}^T \beta_{tc} I(t) + \mathbf{x}_i \boldsymbol{\gamma}_c + \boldsymbol{\delta}_{sc} + \boldsymbol{\zeta}_{tc} + \boldsymbol{\kappa}_c S_{it} \Omega(t)])} - \frac{1}{1 + \exp(-[\alpha_c + \mathbf{x}_i \boldsymbol{\gamma}_c + \boldsymbol{\delta}_{sc} + \boldsymbol{\zeta}_{tc} + \boldsymbol{\kappa}_c S_{it} \Omega(t)])}. \quad (8)$$

This measure follows the logic of treatment effects in nonlinear models described in Athey and Imbens (2006) and Puhani (2012).

4 Data

As previously mentioned, we use data from the Behavioral Risk Factor Surveillance System (BRFSS) to examine the effects of the SIM initiative on population health outcomes in the six, round one model test states. The BRFSS is a nationwide telephone survey that collects data about US residents' self-perceived health status, risky health behaviors, chronic health conditions, and use of preventative healthcare services. Sponsored by the Centers for Disease Control and Prevention (CDC) and supported by a number of federal and state government agencies, the BRFSS currently collects data in all 50 states, the District of Columbia, and three US territories (Centers for Disease Control and Prevention, 2019).

We use data from the 2010 – 2012 and 2014 – 2016 BRFSS surveys to examine the effects of the SIM initiative. Data from 2013 is excluded from our analysis because that was the year in which the states learned of their awards and worked with CMMI to obtain the funding necessary to begin implementation. We conduct our analysis using the sample of individuals ages 45 and older.

Our sample consists of 832,081 individuals ages 45 or older with 45.8% of the individuals being elderly (ages 65 or older). The population is primarily non-Hispanic

white (89.1%), female (60.6%), married (55.4%), and 60% have more education than a high school diploma. We define a low socioeconomic status group as individuals who have a household income less than \$35,000 per year and education level of no more than a high school diploma. Such individuals comprise 28% of the sample.

We use three measures of health status, elicited in the survey using the following language:

1. Would you say that in general your health is (1) excellent; (2) very good; (3) good; (4) fair; (5) poor
2. Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?
3. Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?

Overall, 78.4% of the study population feel good or better about their general health status, with on average 4.9 and 3.1 days per month with “not good” physical and mental health, respectively.

Table 1 shows summary statistics for the outcomes. The number of days in “not good” physical health is somewhat lower in treated states both before and after the State Innovation Models initiative. For both groups, the mean number of days also declines slightly after SIM as compared to the period prior. The differences in mean “not good” mental health days, across treatment and control states and before and after SIM, are small. Self-rated health status is slightly better in the treated states relative to control states before and after SIM with a negligible mean decline in the rating after SIM.

The distribution of age, gender, race and ethnicity, and marital status are very similar across treated and control states and stable across the pre- and post-SIM

years for our main sample (table 2). Individuals in treated states have higher levels of education, are less likely to be black and more likely to live in rural areas. The percentage of survey respondents contacted by cell phone is similar across treated and control groups and stable over time.

We undertake a number of additional analyses using the following subsamples of our data:

1. Individuals ages 65 and older;
2. Individuals 45 and older of low socioeconomic status (earning less than \$35,000 per year and with a high school diploma or less);
3. Individuals 45 - 64 of low socioeconomic status;
4. Individuals 45 and older of a minority race or ethnicity;
5. Individuals 45 and older living in rural areas;
6. Individuals 65 and older living in rural areas.

The first subsample targets Medicare eligible individuals. They typically have higher healthcare needs but are well insured by Medicare. The second subsample attempts to target the Medicaid eligible population. Unfortunately, BRFSS does not ask about health insurance coverage, except for a couple of years. We used the information on Medicaid coverage in those years to validate our proxy variables – low income and low education – as predictors of Medicaid coverage. The third sample uses the same low socioeconomic status proxies for Medicaid coverage for the under-65 population. As table 3 shows, this sample is substantially smaller than the main sample and the smallest of all subsamples.

The remaining subsamples focus on sub-populations of interest to policy-makers but are not explicitly part of the design elements in SIM. In subsample 4, we limit our attention to individuals of minority race or ethnicity. In the final two cases, we limit

our attention to rural individuals, with the final sample focusing on just the rural elderly. For each of these subsamples, we report the sample means of each outcome along with associated sample sizes for control and treated groups for the pre- and post-SIM periods in table 3. Features of the measures of health status observed in our main sample persist across subsamples.

5 Results

Our statistical design is agnostic about trends in the pre-program or baseline period. In other words, we conduct an event study for which we set the effect of the differential trend to zero for 2012 and estimate effects for 2010 and 2011 in the pre-program period and 2014, 2015, and 2016 during the implementation of SIM.

A selected set of parameter estimates of the latent class profile model from the full sample of individuals ages 45 and older are shown in table 4. To be precise, these are the parameters associated with equation 7. The coefficients on the interactions of treatment assignment and pre-treatment years 2010 and 2011 are not statistically significant. The coefficients on the interactions of treatment assignment and the post-treatment years 2014-2016 are all statistically significant and positive. Thus the probability of being in latent class 1 increases among the treated population as a result of the treatment. Below we show that latent class 1 refers to better health. As expected, married, those with higher education and higher income are more likely to be in class 1. Women are significantly less likely to be in class 1. Black and Hispanic individuals and individuals over 65 years old are more likely to be in class 1. These findings may appear counterintuitive. However, they are consistent with studies that find subjective health status being reference dependent. Groot (2000) shows that adaptation behaviour leads to the scale of reference of the subjective health measure changing with age. Lindeboom and van Doorslaer (2004) find clear evidence of shifting reference points by age and gender, but not by income, education or language. Among effects of covariates not shown, it is worth noting that some state

and year fixed effects are statistically significant. Coefficients on some state-specific time trends are also statistically significant. In addition, a few month fixed effects are significant.

In order to interpret the effect sizes of treatment and to interpret health status in the two classes, we calculate marginal effects from the model and display those in figure 1. The top figure of the panel shows the marginal effects and associated 95% confidence intervals of the year indicators on the probability of being in latent class 1. These effects are estimated using equation 8 which is derived from equations 5 - 7. Note that 2013 observations were not considered as that was the year in which SIM was implemented in the treated states. The estimated effects for 2010 and 2011 are small and statistically insignificant. The effects for 2014-2016 are positive and highly significant. The estimated effect sizes of about 2 percentage points can be interpreted as change from a base class probability of 75%.

The bottom-left panel of the figure shows the estimated mean numbers of days in “not good” physical and mental health days. These are estimated from equations 1 and 2 respectively. Individuals in class 1 are estimated to have 0.62 “not good” physical health days while individuals in class 2 have 17.59 “not good” physical health days. Individuals in class 1 have 0.43 “not good” mental health days while individuals in class 2 have 11.16 “not good” mental health days. The bottom-right figure in the panel shows the estimated proportion in each category of self-reported health by latent class. These are estimated from equation 3. The parameter estimates imply that 19% of individuals in class 1 report excellent health while only 4% in class 2 report excellent health. On the other end of the scale, 9% of individuals in class 1 report being in fair health while 32% in class 2 report being in fair health. Less than 1% of individuals in class 1 report being in poor health. This is stark contrast to 24% of individuals in class 2. Moving forward, we refer to individuals in class 1 as “healthy” individuals while those in class 2 are “unhealthy”.

Taking the results of the LCP model together, it shows that, for the sample

of individuals ages 45 and older, SIM increased the likelihood of being healthy by a statistically significant 2 percentage points. More precisely, the estimates suggest that treated and control states had similar levels of health status prior to SIM and that the innovations in healthcare delivery initiated via SIM led to significant improvement in population health in the states that implemented it.

Figure 2 shows the estimates of the latent class models for the other six samples. The first panel (2a) displays the estimated effects for the subsample of individuals 65 years and older. The figure on the left side of the panel shows the marginal year effects on the probability of being in class 1. It shows that SIM had a statistically significant effect on the health status of older (Medicare eligible) individuals in the treated states. The probability of being in class 1 (the better health status class) increased by 3 percentage points. The figure in the center of the panel shows that individuals in class 1 have substantially fewer days in “not good” physical and mental health. The figure at the right shows that individuals in class 1 are much more likely to be in fair or poor health and much less likely to be in excellent or very good health.

When the sample is restricted to low socioeconomic status individuals (as defined by low income and low education) (panel 2b), the positive effects of SIM on health continue to be present. The yearly effects are close to zero prior to SIM and significantly positive post SIM, although the confidence intervals are larger than in the prior samples in large part because the sample size is 28% of the full sample. In the estimates shown in panel 2c, the sample was restricted to low socioeconomic status individuals ages 45-64 in our attempt to better target individuals who might be Medicaid eligible. The latent classes estimated have qualitatively similar characteristics as for the other subsamples: individuals in class 1 are in better health than individuals in class 2. The baseline rate of membership in class 1 is 0.60, which is lower than the rates for the samples that include all income levels but similar to that for the sample of individuals of all ages but of low socioeconomic status. Individuals of low socioeconomic status are more likely to be unhealthy. The estimates show positive

effects in 2014 and 2016. The effect in 2014 is just shy of being statistically significant at conventional levels but the 2016 estimate is. The effect in 2015 is close to zero. Note that the confidence intervals are large corresponding with the relatively small sample (less than 12% of the original sample).

The effects shown in panel 2d are based on the sample of minority individuals. Minority individuals comprise 13% of the sample in control states and 6% of the sample in treated states. The estimates are, once again, less precise but suggestive of improved health post-SIM.

When we restrict the sample to individuals living in rural areas (panel 2e), we find that SIM has a statistically significant effects on improvement of health status, with the effects plausibly increasing over the period of treatment from 2014-2016. Finally, in panel 2f, we display the effects of SIM for the sample of older individuals who live in rural areas. We find statistically significant and positive effects of SIM. Compared to all rural individuals, the effects for the elderly sample are somewhat larger.

The evidence that there are no residual pre-trend differences between treated and control states in our specification is consistent across samples. Results of the formal tests of the joint hypothesis that corresponds to differential pre-trends are reported in table 5. The smallest p-value of the tests across the 7 samples is 0.178. These can also be observed informally in the figures.

In order to establish the virtues of the latent class model for the purposes of evaluating the effects of SIM relative to simpler, single equation models, we estimate a Poisson regression for the count of “not good” physical health days, another Poisson regression for the count of “not good” mental health days and an ordered logit regression for the ordinal measure of self-reported health status. Covariates are specified as in the LCP model using the interrupted time series design. The results of this analysis are reported in table 6. The first panel of results refers to the sample of individuals ages 45 and older. In the Poisson regression for “not good” physical health days, SIM has a significant health-improvement effect in 2014 and 2016. The effect of SIM is

negative but not statistically significant in 2015. The effects on mental health are all negative but never statistically significant at conventional levels. The effects of SIM on self-reported health status (parameter estimates from an ordered logit regression) show *positive* significant effects implying worse self-reported health. But these findings contradict those from the Poisson regressions. In the second panel, which focuses on the sample of individuals 65 and older, the estimates from the Poisson regression for physical health days and the estimates from the ordered logit regression for self-reported health status all point in the same direction and are statistically significant. In this case, SIM shows improving health for two of the three measures. But there is no evidence of improvement on the mental health measure.

The results continue to be inconsistent, within and across samples, in the remaining panels. We view these as evidence of the lack of reliability of the regressions taken one measure at a time. Complementarily, we view these as evidence of the value of a latent class model that treats the three measures as noisy indicators of an underlying health status construct. Our econometric model is able to improve the signal-to-noise ratio and thus produces more reliable estimates of the effect of SIM on health status.

6 Conclusion

We have examined the effects of the State Innovation Models initiative implemented by the Center for Medicare and Medicaid Innovation on health outcomes in the six states that implemented their innovation plans as compared to 15 states that were not involved at all. We used data from the Behavioral Risk Factor Surveillance System for the years 2010 – 2012 and 2014 – 2016 to examine the effects of treatment. We used an event study design using a Latent Class Profile model for multiple measures of health status to study the effect of SIM. We chose to be agnostic about the pre-program trends in treatment and control states because of the changing landscape of healthcare financing and delivery induced, in part, by the Affordable Care Act. Although it is unlikely that other state level innovations had the scope and scale of

SIM, we chose not to pre-judge the effects of other innovations.

The results of our analysis are remarkably consistent. Individuals in states that implemented SIM saw significant improvements in health status in most of the eight subpopulations we considered. Importantly, we found no evidence of differential pre-intervention trends. The effects of SIM appear to be stronger for the sample of elderly, Medicare eligible individuals as compared to the sample of individuals aged 45 and older. The effects of SIM, for the elderly, are of similar magnitudes for the rural population as it is for the overall elderly population suggesting that the investments are impacting rural healthcare providers as well as urban and suburban ones. The effects of SIM on low socioeconomic status, 45-64 year old individuals are somewhat inconsistent in terms of statistical significance. Nevertheless, taking the substantially smaller sample size into account, it appears that SIM initiatives via Medicaid channels also improved the health of low socioeconomic status individuals.

We find that our use of the latent class profile model is justified. When each of the three health outcomes is modeled separately, the results are sporadically significant, inconsistent across measures and across samples. The ability of the latent class profile model to isolate the common factor in the measures of health status yields substantial gains in statistical power for two reasons. First, in the single equation models, covariates affect each outcome independently. In models where small effects are expected, such models can yield inconsistent estimates across equations. The LCP model forces consistency of covariate effects across measures. Second, there are gains in statistical power because all outcomes are modeled jointly.

In many analyses using finite mixture models, researchers search for the best fitting mixture models before interpreting results. We explored results in a 3-class LCP model but did not search for the best fitting LCP model more generally. We chose to report results from the 2-class LCP model because it created a clean and intuitive distinction between healthy and unhealthy states and because covariates could be interpreted monotonically and easily as either making better health more or

less likely. The 3-class model displayed monotonic health states but the interpretation of covariate effects (now from a multinomial logit specification with 2 coefficients for each covariate) became cumbersome and did not appear to produce any additional insight.

In conclusion, we find robust and substantial improvements in health status among the populations in states that implemented their SIM initiatives. The effects appear to be working through Medicare and Medicaid channels, although we cannot rule out other channels. In future work, we plan to examine the effects of SIM in the states that implemented programs in the second round.

This study has a number of limitations, nevertheless. First, although we have deliberately taken a population perspective, it is important to understand healthcare channels through which these innovations might have affected the population. The BRFSS, unfortunately, has no good measures of healthcare utilization. Second, the states participating in SIM were explicit about the involvement of Medicaid. We were unable to stratify by Medicaid insurance status (choosing to stratify by a proxy defined by low socioeconomic status instead) because the question of insurance type is asked only in a couple of BRFSS years. Finally, although it would be of some importance to be identify specific features of the states' innovations that made the most difference, with only 6 treated states and no clear low-dimensional distinction between the treatment states, we chose not to explore along those dimensions.

Additional caveats are also in order. First, other CMMI interventions existed in the treated states (and some control states) prior to SIM. Our evidence suggests that there were no population level differential effects prior to 2014. Second, the fact that our "Medicaid" results are weak can be attributed to the fact that we cannot directly identify Medicaid enrollees. The sample of low SES individuals likely contains a fraction of healthy individuals not enrolled in Medicaid (even if some of them may be eligible). Such individuals would decrease the effect size on changes in health status. Finally, although SIM initiatives were not targeted towards Medicare beneficiaries,

we find a robust effect for the elderly sample. These effects are plausibly driven by broad infrastructure and training changes funded via SIM combined with the fact that Medicare beneficiaries are frequent users of healthcare services, thus most exposed to systemic changes.

References

- Athey, S. and Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497.
- Auerbach, J., Chang, D. I., Hester, J. A., and Magnan, S. (2013). Opportunity knocks: Population health in State Innovation Models. Discussion Paper, Institute of Medicine, Washington, DC.
- Center for Medicare and Medicaid Innovation (2019). State Innovation Models initiative: Model test awards round one — Center for Medicare & Medicaid Innovation. <https://innovation.cms.gov/initiatives/State-Innovations-Model-Testing/>. (Accessed on May 26, 2019).
- Centers for Disease Control and Prevention (2019). Behavioral Risk Factor Surveillance System. <https://www.cdc.gov/brfss/index.html>. (Accessed on May 26, 2019).
- Chung, H., Anthony, J. C., and Schafer, J. L. (2011). Latent class profile analysis: An application to stage sequential processes in early onset drinking behaviours. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(3):689–712.
- Deb, P. and Trivedi, P. K. (1997). Demand for medical care by the elderly: A finite mixture approach. *Journal of Applied Econometrics*, 12(3):313–336.
- Dummit, L. A., Kahvecioglu, D., Marrufo, G., Rajkumar, R., Marshall, J., Tan, E., Press, M. J., Flood, S., Muldoon, L. D., Gu, Q., et al. (2016). Association between hospital participation in a medicare bundled payment initiative and payments and quality outcomes for lower extremity joint replacement episodes. *Journal of the American Medical Association*, 316(12):1267–1278.
- Erosheva, E. A. (2005). Comparing latent structures of the grade of membership, rasch, and latent class models. *Psychometrika*, 70(4):619–628.
- Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2):215–231.
- Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics*, 19(3):403 – 420.
- Hughes, L. S., Peltz, A., and Conway, P. H. (2015). State Innovation Model Initiative: A state-led approach to accelerating health care system transformation. *Journal of the American Medical Association*, 313(13):1317–1318.
- Lazarsfeld, P. F. and Henry, N. W. (1968). *Latent Structure Analysis*. New York, Houghton, Mifflin.

- Lindeboom, M. and van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, 23(6):1083 – 1099.
- Manton, K. G., Woodbury, M. A., Stallard, E., and Corder, L. S. (1992). The use of grade-of-membership techniques to estimate regression relationships. *Sociological Methodology*, pages 321–381.
- McLachlan, G. and Peel, D. (2004). *Finite Mixture Models*. John Wiley & Sons.
- McWilliams, J. M. (2016). Changes in medicare shared savings program savings from 2013 to 2014. *Journal of the American Medical Association*, 316(16):1711–1713.
- Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In Robertson, J. and Kaptein, M., editors, *Modern Statistical Methods for HCI*, pages 275–287. Springer International Publishing.
- Portrait, F., Lindeboom, M., and Deeg, D. (1999). Health and mortality of the elderly: the grade of membership method, classification and determination. *Health Economics*, 8(5):441–458.
- Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear “difference-in-differences” models. *Economics Letters*, 115(1):85–87.
- RTI International (2016). State Innovation Models (SIM) initiative evaluation: Model test year two annual report. <https://downloads.cms.gov/files/cmmi/sim-round1-secondannualrpt.pdf>. (Accessed on May 26, 2019).
- RTI International (2017). State Innovation Models (SIM) initiative evaluation: Model test year three annual report. <https://downloads.cms.gov/files/cmmi/sim-rd1mt-thirdannrpt.pdf>. (Accessed on May 26, 2019).
- RTI International (2018). State Innovation Models (SIM) initiative evaluation: Model test year five annual report. <https://downloads.cms.gov/files/cmmi/sim-rd1mt-fifthannrpt.pdf>. (Accessed on May 26, 2019).
- Van Vleet, A. and Paradise, J. (2014). The State Innovation Models (SIM) program: An overview — The Henry J. Kaiser Family Foundation. <https://www.kff.org/medicaid/fact-sheet/the-state-innovation-models-sim-program-an-overview/>. (Accessed on May 26, 2019).

Table 1: Summary statistics of outcome measures

	Control before SIM	Treated before SIM	Control after SIM	Treated after SIM
0 not good physical health days	0.62	0.63	0.63	0.63
1 – 4 not good physical health days	0.13	0.14	0.13	0.14
5 – 9 not good physical health days	0.05	0.05	0.05	0.05
10 – 19 not good physical health days	0.06	0.06	0.06	0.06
≥ 20 not good physical health days	0.13	0.12	0.13	0.11
mean not good physical health days	5.11	4.64	4.94	4.49
0 not good mental health days	0.73	0.71	0.74	0.72
1 – 4 not good mental health days	0.10	0.12	0.10	0.12
5 – 9 not good mental health days	0.04	0.05	0.04	0.05
10 – 19 not good mental health days	0.05	0.05	0.05	0.05
≥ 20 not good mental health days	0.08	0.07	0.07	0.07
mean not good mental health days	3.25	3.23	3.03	2.97
excellent health	0.14	0.18	0.14	0.18
very good health	0.30	0.34	0.31	0.35
good health	0.32	0.29	0.32	0.30
fair health	0.16	0.13	0.15	0.13
poor health	0.08	0.06	0.07	0.05
mean health status	2.74	2.55	2.69	2.53
N	308,391	124,172	285,157	114,361

Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS.

Table 2: Summary statistics of covariates

	Control before SIM	Treated before SIM	Control after SIM	Treated after SIM
Female	0.62	0.61	0.59	0.58
45-54 years old	0.24	0.27	0.21	0.23
55-64 years old	0.30	0.32	0.30	0.31
65 years and older	0.45	0.41	0.48	0.45
Married	0.55	0.53	0.56	0.56
< HS education	0.10	0.07	0.08	0.05
HS graduate	0.34	0.29	0.31	0.27
Some college	0.27	0.26	0.28	0.27
College graduate	0.29	0.38	0.32	0.40
Black	0.11	0.03	0.10	0.03
Hispanic	0.02	0.03	0.03	0.02
Income < 15,000	0.11	0.10	0.09	0.08
Income 15,000– < 25,000	0.17	0.15	0.15	0.13
Income 25,000– < 35,000	0.11	0.10	0.10	0.09
Income 35,000– < 50,000	0.13	0.13	0.13	0.13
Income \geq 50,000	0.32	0.38	0.36	0.42
Income missing	0.15	0.14	0.17	0.16
Rural	0.44	0.28	0.27	0.23
Cell phone	0.06	0.07	0.34	0.35
N	308,391	124,172	285,157	114,361

Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS.

Table 3: Summary statistics of outcome measures for alternative samples

Sample		Control before SIM	Treated before SIM	Control after SIM	Treated after SIM
65 & older	Not good physical health days	5.42	4.97	5.05	4.63
	Not good mental health days	2.24	2.15	2.22	2.15
	Health 5 point scale	2.85	2.69	2.75	2.61
	N	139,831	51,384	137,979	51,779
45 & older of low SES	Not good physical health days	7.45	7.31	7.37	7.27
	Not good mental health days	4.36	4.59	4.31	4.44
	Health 5 point scale	3.21	3.10	3.16	3.07
	N	98,528	31,548	78,324	24,817
45 - 64 of low SES	Not good physical health days	8.38	8.52	8.56	8.68
	Not good mental health days	6.60	7.14	6.38	6.72
	Health 5 point scale	3.28	3.16	3.26	3.16
	N	40,554	13,211	32,422	10,453
45 & older of minority race or ethnicity	Not good physical health days	6.38	6.62	6.07	6.07
	Not good mental health days	4.17	4.64	3.88	4.31
	Health 5 point scale	3.09	2.96	2.99	2.86
	N	55,407	13,277	50,697	11,216
45 & older in rural areas	Not good physical health days	5.23	4.83	5.17	4.89
	Not good mental health days	3.09	3.16	2.81	2.87
	Health 5 point scale	2.80	2.61	2.77	2.63
	N	135,825	35,055	78,390	26,831
65 & older in rural areas	Not good physical health days	5.55	4.99	5.31	4.90
	Not good mental health days	2.18	2.19	2.17	2.20
	Health 5 point scale	2.90	2.72	2.85	2.69
	N	65,479	16,058	43,674	14,495

Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS. Low SES is defined as individuals who have income<35,000 and education level no more than a high school diploma.

Table 4: Selected Parameter Estimates

	coefficient	standard error
SIM state \times 2010	0.024	0.018
SIM state \times 2011	0.014	0.021
SIM state \times 2014	0.102	0.022***
SIM state \times 2015	0.076	0.021***
SIM state \times 2016	0.102	0.019***
2010	0.040	0.012**
2011	0.027	0.011*
2014	-0.004	0.013
2015	-0.063	0.011***
2016	-0.071	0.011***
SIM state	0.235	0.018***
Female	-0.165	0.009***
55 - 64 years old	-0.002	0.010
65 years and older	0.357	0.022***
Married	0.145	0.013***
HS graduate	0.360	0.020***
Some college	0.336	0.019***
College graduate	0.649	0.028***
Black	0.190	0.029***
Hispanic	0.125	0.052*
Income 15,000– < 25,000	0.525	0.015***
Income 25,000– < 35,000	0.913	0.019***
Income 35,000– < 50,000	1.145	0.017***
Income \geq 50,000	1.561	0.017***
Income missing	0.995	0.024***
Rural	0.035	0.018
Cell phone	0.012	0.011

Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS. A selection of parameter estimates from the main sample are shown. Parameter estimates on the full set of state and month indicators and on state-specific linear trends are not shown. Standard errors are adjusted for clustering at the state-level. *, ** and *** denote statistical significance at 0.1, 0.05 and 0.01 levels respectively.

Table 5: Wald χ^2 test of differential pre-treatment trends

Sample	$\chi^2(2)$ test	p value
45 & older	1.77	0.414
65 & older	0.54	0.765
45 & older of low SES	2.54	0.281
45 - 64 of low SES	1.17	0.557
45 & older of minority race	0.47	0.791
45 & older in rural areas	3.07	0.215
65 & older in rural areas	3.45	0.178

Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS. Low SES is defined as individuals who have income<35,000 and education level no more than a high school diploma. Wald $\chi^2(2)$ test of the joint hypothesis that the coefficients of 2010 and 2011 year interacted with treated states indicators are equal to zero.

Table 6: Year effects in interrupted time series models for each health measure

Sample		Physical health not good days	Mental health not good days	Self-reported health status
45 & older	Treated in 2010	-0.013 (0.016)	-0.023 (0.016)	-0.034 (0.023)
	Treated in 2011	-0.003 (0.011)	0.004 (0.018)	-0.002 (0.015)
	Treated in 2014	-0.037** (0.015)	-0.026 (0.019)	0.025** (0.010)
	Treated in 2015	-0.012 (0.020)	-0.015 (0.021)	0.037** (0.019)
	Treated in 2016	-0.035*** (0.010)	-0.002 (0.021)	0.024 (0.020)
65 & older	Treated in 2010	0.015 (0.023)	-0.008 (0.031)	-0.023 (0.034)
	Treated in 2011	0.003 (0.022)	0.048* (0.028)	-0.002 (0.018)
	Treated in 2014	-0.089*** (0.023)	-0.037 (0.025)	-0.043** (0.019)
	Treated in 2015	-0.061** (0.024)	0.002 (0.031)	-0.047* (0.025)
	Treated in 2016	-0.106*** (0.020)	0.018 (0.026)	-0.045* (0.027)
45 & older of low SES	Treated in 2010	0.022 (0.021)	-0.003 (0.025)	-0.045 (0.040)
	Treated in 2011	0.011 (0.026)	0.006 (0.033)	-0.029 (0.024)
	Treated in 2014	-0.039* (0.022)	-0.037 (0.041)	0.032 (0.025)
	Treated in 2015	-0.030 (0.024)	-0.030 (0.027)	0.057* (0.032)
	Treated in 2016	-0.066*** (0.025)	-0.027 (0.027)	0.023 (0.040)
45 - 64 of low SES	Treated in 2010	0.010 (0.022)	0.011 (0.033)	-0.087* (0.046)
	Treated in 2011	0.001 (0.029)	-0.016 (0.042)	-0.085** (0.041)
	Treated in 2014	0.044 (0.033)	-0.047 (0.054)	0.095*** (0.036)
	Treated in 2015	0.042 (0.036)	-0.048 (0.042)	0.133*** (0.047)
	Treated in 2016	0.006 (0.020)	-0.070** (0.029)	0.089** (0.039)

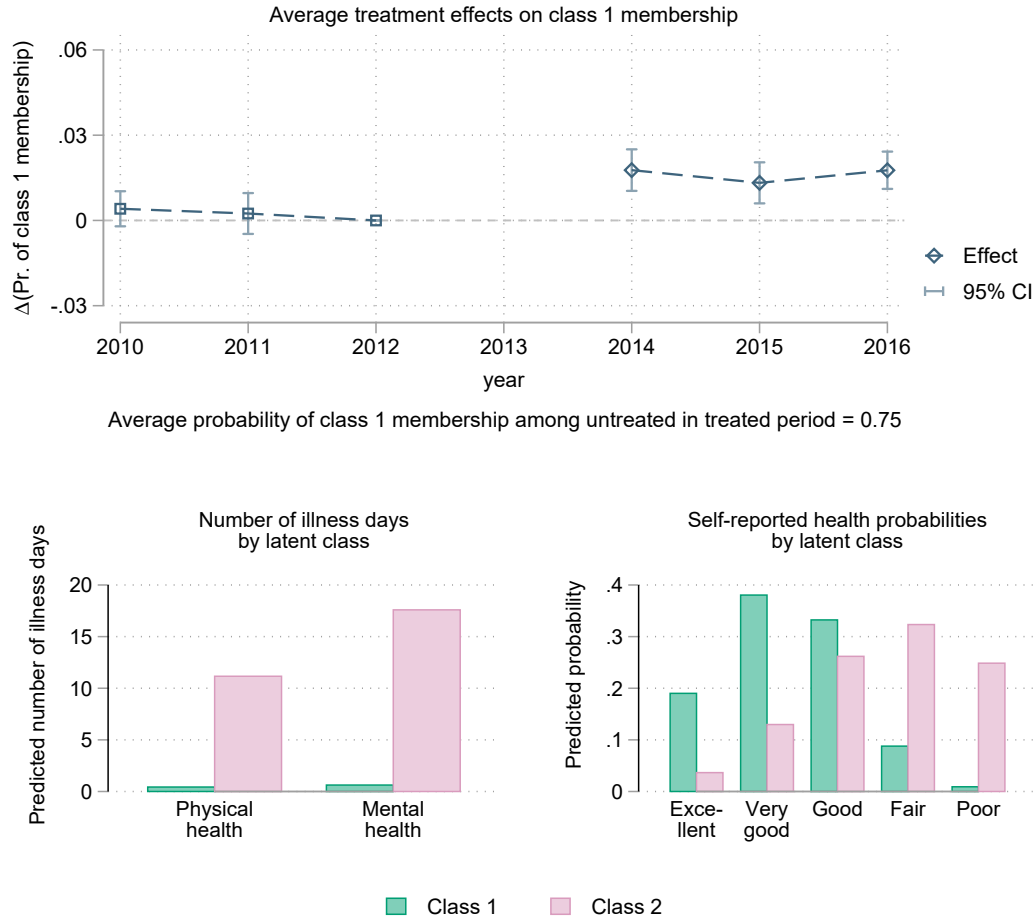
Continued on next page

Table 6: Year effects in interrupted time series models for each health measure

Sample		Physical health not good days	Mental health not good days	Self-reported health status
45 & older of minority race or ethnicity	Treated in 2010	-0.005 (0.025)	0.104** (0.044)	-0.196*** (0.040)
	Treated in 2011	0.013 (0.030)	0.028 (0.030)	-0.016 (0.033)
	Treated in 2014	-0.096* (0.049)	0.028 (0.049)	0.006 (0.039)
	Treated in 2015	-0.018 (0.032)	0.042 (0.039)	0.033 (0.027)
	Treated in 2016	-0.036 (0.025)	0.073* (0.040)	0.053* (0.032)
45 & older in rural areas	Treated in 2010	-0.042** (0.018)	-0.041 (0.028)	0.030 (0.019)
	Treated in 2011	0.005 (0.018)	0.011 (0.025)	0.041 (0.038)
	Treated in 2014	-0.028 (0.023)	-0.014 (0.020)	0.031** (0.015)
	Treated in 2015	-0.043 (0.030)	-0.127*** (0.037)	0.021 (0.022)
	Treated in 2016	-0.074*** (0.020)	-0.006 (0.034)	-0.029 (0.024)
65 & older in rural areas	Treated in 2010	-0.006 (0.024)	-0.151* (0.081)	0.037 (0.034)
	Treated in 2011	0.033 (0.024)	0.022 (0.038)	0.037 (0.034)
	Treated in 2014	-0.067** (0.033)	0.042 (0.035)	0.023 (0.022)
	Treated in 2015	-0.061* (0.034)	-0.003 (0.083)	-0.036 (0.048)
	Treated in 2016	-0.096*** (0.021)	0.148*** (0.057)	-0.042 (0.027)

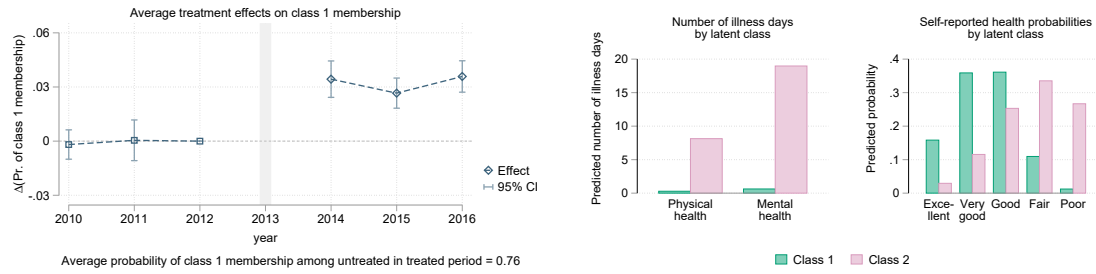
Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS. Low SES is defined as individuals who have income<35,000 and education level no more than a high school diploma. Models for the counts of “not good” physical health days and “not good” mental health days are Poisson regressions. The model for the ordinal measure of self-reported health status is an ordered logit regression. Demographic and socioeconomic status characteristics of the individual included as covariates include gender, age, education, race and ethnicity, income group, rural/urban status, and whether the survey was conducted by cell phone. All models included state and year indicators, a state-specific linear year trend and month of survey indicators. Estimates of the differential time trend coefficients are reported. Cluster-robust standard errors are in parentheses. ***, ** and * denote statistical significance at 1%, 5% and 10% respectively.

Figure 1: Marginal year effects on latent health class: Individuals 45 & older

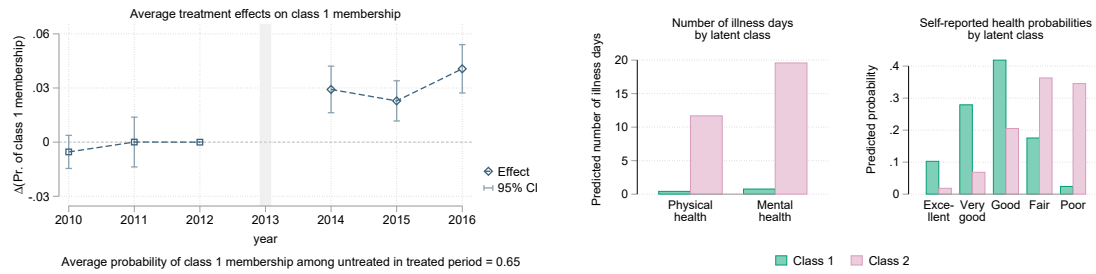


Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS. Latent class profile model is estimated where the measures of latent health status are a count of “not good” physical health days, a count of “not good” mental health days are Poisson regressions and an ordinal measure of self-reported health status. Demographic and socioeconomic status characteristics of the individual included as covariates include gender, age, education, race and ethnicity, income group, rural/urban status, and whether the survey was conducted by cell phone. All models included state and year indicators, a state-specific linear year trend and month of survey indicators. Estimates of the differential time trend marginal effects are shown in the figure. Cluster-robust standard errors are used to display 95% confidence intervals.

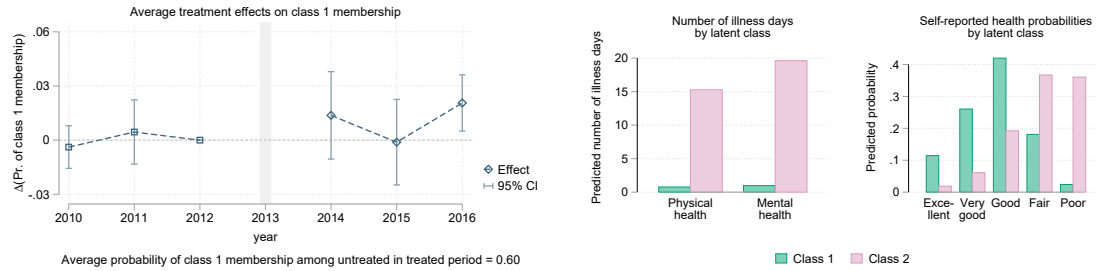
Figure 2: Marginal year effects on latent health class



(a) Individuals 65 & older

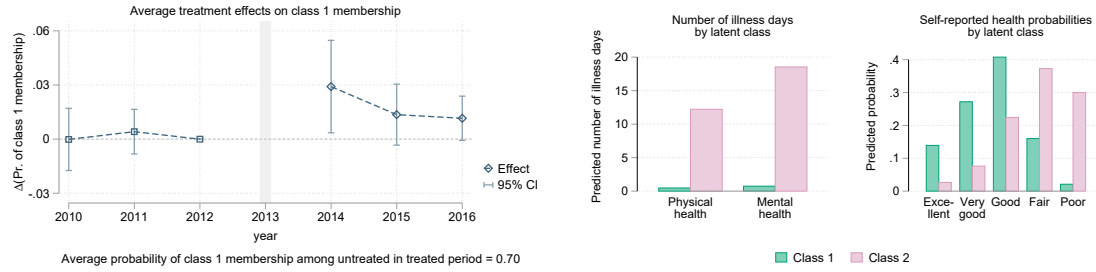


(b) Individuals 45 & older of low SES

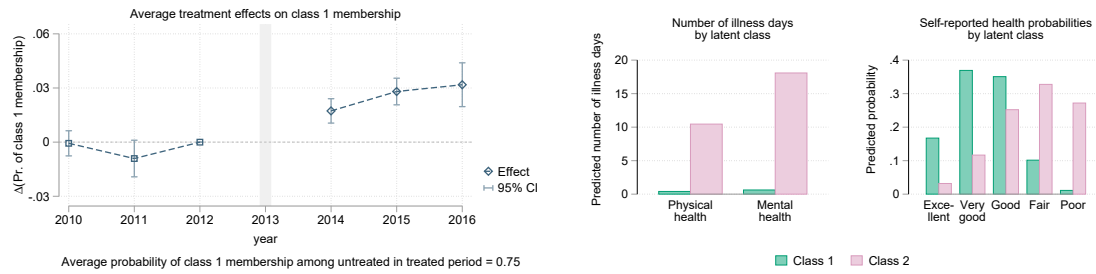


(c) Individuals 45 - 64 of low SES

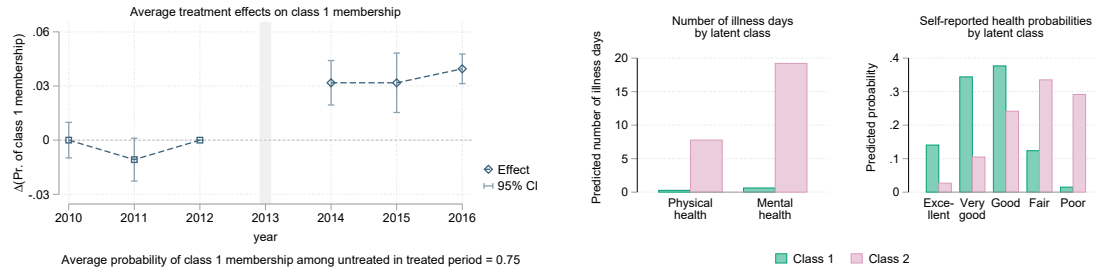
Figure 2: Marginal year effects on latent health class



(d) Individuals 45 & older of minority race or ethnicity



(e) Individuals 45 & older living in rural areas



(f) Individuals 65 & older living in rural areas

Note: Sample of observations from 2010-2012 (pre-SIM) and 2014-2016 (post-SIM) BRFSS. Low SES is defined as individuals who have income < \$35,000 and education level no more than a high school diploma. Latent class profile model is estimated where the measures of latent health status are a count of “not good” physical health days, a count of “not good” mental health days are Poisson regressions and an ordinal measure of self-reported health status. Demographic and socioeconomic status characteristics of the individual included as covariates include gender, age, education, race and ethnicity, income group, rural/urban status, and whether the survey was conducted by cell phone. All models included state and year indicators, a state-specific linear year trend and month of survey indicators. Estimates of the differential time trend marginal effects are shown in the figure. Cluster-robust standard errors are used to display 95% confidence intervals.