The effect of same-sex marriage legalization on health insurance coverage for same-sex households in the United States.

Case studies using the synthetic control method

Abstract

This paper studies the effect of state level same-sex marriage legalization on health insurance coverage levels for same-sex households in the United States. Between 2000 and 2012 a number of US states legalized same-sex marriage. For same-sex couples this provided the possibility to enter into marriage and receive the same spousal benefits as are available to opposite sex couples. Utilizing the synthetic control method this paper analyzes four case studies. Positive but not significant effects of legalization are found for same-sex households as well as the general public in the cases of Iowa, New Hampshire and New York. For the state of Vermont, a positive and significant effect is found for the total population but not for same-sex households as a sub-group. This effect is not robust to backdating or a change in research design. These findings provide a critique of utilizing BRFSS survey data to make inferences for same-sex households as well as a basis for further analysis of the determinants of health insurance coverage.

Name student: Wouter van Dam
Student ID number: 453826

Supervisor: Dr. K.F.J. Spiritus
Second assessor: Dr. A.L. Boring

Date final version: 13-07-2021

ERASMUS UNIVERSITY ROTTERDAM Erasmus School of Economics

Bachelor Thesis Economics and Business Economics, major in Policy Economics.

Table of contents

1. Introduction

Advancement in LGBT rights has been one of the most salient societal changes in the United States in the 21st century (Smith, 2008). A policy question in this arena that has been frequently debated over the past 20 years is that of same-sex marriage. Marriage gives couples many legal rights as well as economic advantages through decreased taxes and more conferred benefits. One of the economic advantages is the conferral of employer sponsored health-insurance to a married partner. This paper seeks to evaluate this economic effect of marriage legalization.

Between 2000 and 2010 number of US states implemented some form of same-sex marriage or domestic partnership whilst in other states no such implementation took place (Sherkat et. al, 2011). This paper will address one of the effects of same-sex marriage legalization and answer the following research question: **What is the effect of same-sex marriage legalization on health insurance coverage for same-sex households in the United States?**

This effect is driven by individuals who are first uninsured but after marrying, once same-sex marriage is legalized, are insured on their spouse's employer sponsored health insurance. In 2019 55% of the US population was covered through an employer sponsored health insurance plan, making it by far the most common form of coverage (Keisler-Starkey & Bunch, 2020). As employer sponsored health insurance plans only apply to an employee and their legally recognized partner, same-sex marriage legalization is needed to make this coverage available to partners in a same-sex relationship. This research will analyze if the legal possibility of this type of coverage actually leads to a significant increase in health insurance coverage for same-sex households.

Earlier research at the national and state levels using a difference-in-difference approach, and at an individual level using a two-way fixed effects analysis finds a positive impact of same-sex marriage legalization on health insurance coverage rates (Tumin & Kroeger, 2020; Downing & Cha, 2020; Carpenter et. al., 2021). This research builds upon these findings and applies the synthetic control method to analyze 4 case studies of states where same-sex marriage was legalized between 2008 and 2010. Utilizing self-reported health insurance coverage rates aggregated to the state level this research analyzes the outcomes for Iowa, Vermont, New Hampshire and New York. Researching the effect for same-sex households as well as the population as a whole to find an estimated treatment effect.

This paper is organized as follows. Section 2 provides an overview of the literature on the relationship between same-sex marriage and health insurance coverage. Section 3 discusses the methodology and application of the synthetic control method. Section 4 describes the institutional background and Section 5 describes the data and data permutations. Section 6 presents the results and Section 7 concludes and presents a discussion of limitations and policy implications.

2. Literature review

This section provides an overview of the relevant existing literature on this subject as well as the relevance of this study in relation to existing studies. Previous research has shown that same-sex marriage legalization has an impact on variables ranging from employment and homeownership levels to reported family level wellbeing (Delhommer & Hammermesh, 2021; Badgett, 2020). In earlier research one of the most pronounced effects of same-sex marriage legalization is an increase in health insurance coverage: same-sex marriage legalization extends employer sponsored health insurance to the same-sex partners of employees for whom this was not available before legalization (Buffie, 2011). A limited number of private sector firms offer health-care coverage to unmarried same-sex domestic partners but research on this topic found that less than 15% of surveyed firms that offered health insurance coverage provided any form of this to same-sex domestic partners (Kaiser Family Foundation, 2004).

Earlier research has focused on analyzing the effects of marriage legalization on health insurance coverage at different levels of aggregation. Tumin and Kroeger (2020) use national health interview survey data and find that after nation-wide same-sex marriage legalization in the United States in 2015 the number of same-sex couples for which both individuals have private health insurance coverage increases from 79% to 88% (p = .003). In the same time-period, they find no such increase for opposite sex couples which they use as a comparison for the effect they find. This assumes that without same-sex marriage legalization both groups would have followed a similar trend over time.

Downing and Cha (2020) use data from the American Community Survey to compare state level differences in marriage recognition in a difference-in-difference approach exploiting the state level time-variance in same-sex marriage legalization. They find that marriage equality led to a 0.61 percentage point increase (P = .03) in employer-sponsored health insurance in an analysis of the total population. Their difference-in-difference approach leans heavily on a parallel trends assumption. Differences at the state level that violate this assumption are an issue here as marriage legalization is often implemented together with a number of different policies favoring LGBT individuals and they are unable to control for this.

Carpenter et. al. (2021) follow a research design that is the most similar to the method employed in this paper. Carpenter et. al. (2021) use data from the Behavioral Risk Factor Surveillance System health survey (BRFSS) as this paper does too. To analyze the effect, they employ a two-way fixed effects analysis for the period between 2008 and 2017. They find that men in same-sex households in states that have legalized same-sex marriage are 4,1 percentage points more likely to have any form of health-insurance (P = .016). They do not find the same effect for women in same-sex households. Their research includes a number of control variables but relies heavily on the assumption that there is exogenous variation in the timing of same-sex marriage legalization across states. Seeing that same-sex marriage legalization and its timing is a very salient topic of public debate and thus implementation is not likely to be exogenously timed, this is a shortcoming in their research. Even with the possible shortcomings in their methodology these papers all find a positive effect of same-sex marriage legalization on health-insurance coverage.

This research hopes to extend these insights. Using the synthetic control method proposed by Abadie & Gardeazabal (2003) and further elaborated by Abadie et. al (2010) this paper aims to overcome some of the shortcomings in earlier research and provide a better counterfactual for analysis. This method is further explained in Section 3.

3. Methodology

For analysis of the four case studies in this research the Synthetic Control Method will be utilized. Each state that undergoes treatment, the legalization of same-sex marriage, will be compared with a weighted combination of non-treated states: the donor pool. For this research this includes all other states in the BRFSS excluding: the District of Columbia, Connecticut and Massachusetts as they had a form of same-sex marriage or partnership legalization in or before this time period but do not provide a sufficiently long pre-treatment, or training, period; or in the case of D.C. an insufficiently large sample size. The state of Hawaii is also excluded as there was no BRFSS data available for 2004. For each case study the three other treated states are excluded to guard for bias in the counterfactual as they also undergo or have undergone treatment. The following three subsections provide a description of the synthetic control method, the limitations and advantages of this method and the relevant assumptions for its use.

3.1. Description of the synthetic control method

The synthetic control method enables the construction of a weighted combination of non-treated states, through a data driven approach, to represent the treated state had it not undergone same-sex marriage legalization. This approach is outlined below together with the relevant equations and interpretation. To construct this synthetic control this paper will utilize the source code as provided Abadie, Diamond & Hainmueller (2010). After construction of the synthetic control a comparison between the outcome variable, in this case the population percentage that is uninsured, of the treated state compared to the corresponding synthetic control gives the approximated treatment effect.

For each case studies with the synthetic control method $J+1$ states are observed where the first state $j=1$ is the treated state and all other $j = 2, \dots, J+1$ are untreated states in the so-called donor pool. This panel of states is observed over the same time span $t = 1, 2, \dots, T$ where $t = 1, \dots, T0 - 1$ are the time periods before legalization in state J= 1 and $t = T_0, \dots, T$ the year of legalization and the years after. With T_0 denoting the year that same-sex marriage was legalized in the treated state. The outcome variable, or variable of interest, $Y_{j,t}$ is defined as the uninsured population percentage at time t in state j. $Y_{j,t}^N$ is the counterfactual outcome (synthetic control) for treatment state j, which is the weighted sum of the outcomes of donor states $j = 2, \dots, J+1$ at time t. The treatment effect for state j at time t then is α_{jt} and is derived as:

$$\alpha_{it} = Y_{j,t} - Y_{j,t}^N.$$

This computes the difference between the observed uninsured population percentage for state j at time t and the counterfactual uninsured population percentage without treatment for state j at time t. After T_0 we only observe $Y_{j,t}$ and therefore need to estimate $Y_{j,t}^N$. Abadie et. al. (2010;2015) approximate $Y_{j,t}^N$ via a multifactor model on the basis of observed variables and comparison units from the donor pool. $Y_{j,t}^N$ is computed as a weighted sum of donor states $j = 2, \dots, J+1$. With weights w_j the counterfactual is approached by:

$$Y_{j,t}^N = \sum_{j=2}^{J+1} w_j \, Y_{j,t}$$

Where all weights w_j sum to one: $\sum_{j=2}^{J+1} w_j = 1$ and $w_j \geq 0$. These conditions avoid extrapolation and ensure the synthetic control be a weighted average of the donor pool units. With $W = (w_2, \dots, w_{J+1})$, W is defined as a generic (J x 1) vector of weights. With $\mathrm{V_z}$ reflecting the regression coefficient of the Z'th control variable in a regression on the outcome variable, the predictive power or importance of the Z'th variable can be taken into account. $X_1$is a $(K \times 1)$ vector of pre-treatment control variables K for the treatment state and X_0 is a $(K \times J)$ matrix of the same pre-treatment control variables from the donor pool. Optimal weights w_j^* are computed through a squared difference minimization process selecting w_j so that it minimizes the difference between the synthetic and true value of each control variable X (Abadie et al., 2015):

$$min \sum_{Z=1}^{Z} V_z (X_{1Z} - X_{0Z} W)^2$$

Control variables K are defined following earlier research on the determinants of state level health-insurance coverage. Chernew and Cutler (2005) use the current population survey and find a significant effect of the median household income, the unemployment rate and the percentage of non-white residents and the percentage of working woman in a state on the coverage rate in that state. Cebula (2006) finds the same effect for the median family income, the state unemployment rate and the state's population that classifies as Hispanic and further elaborates on the impact of education levels and female labor force participation. For this analysis I follow Carpenter et. al (2021) and use the state year unemployment rate, per capita personal income and percentage of the population that identifies as Hispanic as well as the population percentage with any college education and with a college degree. I exclude individuals aged 60 or over to prevent an interference effect of Medicaid coverage.

In line with the empirical literature, this paper follows the hypothesis that same-sex marriage legalization has a positive and significant effect on health-insurance coverage as it becomes available to same-sex households. For the practical application of the synthetic control method this research will utilize Stata with the *Synth* and *Synth_runner* extension packages (Abadie et. al., 2015, Galiani & Quistorff, 2017).

3.2. Limitations and advantages of synthetic control method

As a method of policy evaluation there are several advantages of but also limitations to the use the synthetic control method. This section will present these limitations and advantages together with their relevance to the specific setting of this research. A main drawback of the synthetic controls use for policy evaluation is that to evaluate the significance of results common statistical inference techniques are not applicable. In order to still statistically infer the treatment effects Abadie et al. (2003; 2010) use placebo tests. This test applies the synthetic control method to each control unit with the same application as the treated unit, analyzing it as if it were the treated state. A treatment effect found for an untreated state is a measure for how likely treatment effects are to show up by chance. By comparing the placebo effect for each control state to that of the treatment state the P-value is ascertained as the fraction of control units with a treatment effect larger than that of the treated unit. The standardized version of this p-value contains a vector of the proportion of placebo t-statistics relative to the pre-treatment Root Mean Square Prediction Error (RMSPE). This standardized p-value retains the same meaning as a classical p-value if the treatment is assigned randomly (Boutell et. al., 2018). In this case treatment cannot be said to be assigned randomly as the political process that leads to legalization is brought about by both political considerations as well as the public perception in that state. Even so, a treatment effect that is much larger than any of the placebo effects, and thus has a low P-value, will be a sign that the estimated effect is a true treatment effect.

A second limitation of the synthetic control method relates to the external validity of its policy outcome. Same-sex marriage legalization happens at a state level but is dependent on the policy landscape of that state. In states where same-sex marriage legalization is enacted we can expect a stronger focus on equal rights legislation and a more social interpretation of existing law and policy. For the area of health insurance coverage, where we find big differences between states, some states will have more inclusive health insurance policy frameworks. This would mean that the effect that is found for treated states is not generalizable to same-sex marriage legalization in all states and is biased downward: in more right leaning states legalization would be expected to have a higher effect.

Finally, the synthetic control method is vulnerable to a reverse causality of the treatment if the outcome variable has a direct effect on treatment application. A reverse effect of health insurance coverage on same-sex marriage legalization is however unlikely. Same-sex marriage legalization is not presented as a policy specifically aimed at changing the health insurance coverage rate so will not be implemented as a reaction to a decrease or expected decrease of the coverage rate in a certain state.

In addition to the aforementioned limitations, there are a few important advantages the synthetic control method provides in comparison to other forms of policy evaluation or economic inference as presented in Section 2. A main driver of the formulation of the synthetic control method is that it formalizes the choice of control units and their weights. Other studies utilize a difference-in-difference model where the researcher has to select a potential control state themselves. The synthetic control method uses available data to match on the basis of observable factors and thus creates a synthetic control state that is not biased by a researcher's selection. Including all applicable states in the donor pool provides an unbiased donor pool as no pre-selection is made.

Secondly, the synthetic control method alleviates some concerns on endogeneity (Fremeth et al., 2013). The synthetic control method selects its weights on the basis of observable time-varying confounders. As these control variables are selected on the basis of their explanatory power for the outcome variable it is assumed that the constructed synthetic control matches the treated unit on both the observed and unobserved factors that affect the outcome variable. This assumption is held when the synthetic control maintains a good fit in the periods before the treatment (Bouttell et al., 2018). Section 3.3. further elaborates on this assumption.

Lastly, in comparison to the more widely utilized difference-in-difference model the synthetic control method is able to relax the parallel trends assumption of this method. As the synthetic control method doesn't require one single control unit but the weighted application of different control units to follow a parallel trend it provides a relaxation on this assumption. Both the pre-treatment fit, and the previously discussed placebo tests are a relevant control for this.

3.3. Assumptions of synthetic control method

Despite the relaxation in respect to other methods of policy evaluation the synthetic control method does still require a number of assumptions to be met to provide a correct application. For these assumptions there are no statistical tests to determine if they are satisfied. This is why there is an increased importance of the examination of these assumptions at a theoretical level. This section presents the relevant assumptions of the synthetic control method and their application.

One main assumption is that the treatment effect only affects the treated unit and only does this after implementation of the treatment. There should be no spillover effects and no anticipation effect. The legal framework surrounding marriage legalization in the Unites States before *United States v. Windsor* (570 U.S. 744 (2013)) entailed that same-sex unions would not be recognized in federal law and would not require direct recognition in other states than the state where the union was entered into. This means that direct spillover effects are absent. Any indirect spillover effect of increased health insurance coverage through out of state employment in donor states is seen as negligible. An anticipation effect at a personal level is unlikely and we do not expect firms to alter their benefit policies on the basis of anticipating same-sex marriage legalization (Hansen et. al. 2020). To ensure the validity of the donor pool, states with same-sex marriage legalization in period of analysis are excluded from the donor pool, ensuring a valid reference group.

One of the assumptions that is significantly relaxed compared to other methods is the assumption that there is no endogeneity of the outcome. For the synthetic control method unobserved confounders do not provide an issue because it assumes that matching on observable variables will mean that there is also a match on any unobservable confounders, as also mentioned in Section 3.2. By matching closely on the observable control variables, it is assumed that the synthetic control is also similar for any unobserved variables that influence the outcome and there is no endogeneity of the outcome. As long as there is a close pre-intervention synthetic control fit, when the synthetic control is able to match closely in the period before same-sex marriage legalization, it is assumed that unobserved confounding variables will also be balanced (Abadie, Diamond, Hainmuller; 2010).

To achieve this correct fit for the synthetic control, the assumption of no dormant factors must also be met. This relates to the factors affecting the relationship between the control variables and the outcome variable. This assumption requires that these factors must apply to the post-treatment period in the same way as they do to the pre-treatment period (Hollingsworth & Wing, 2020). To ensure that there are no significant changes in the overall legal framework surrounding both health insurance (the application of the Affordable Care Act) or national marriage legalization data is restricted to 2002-1012. This will be further elaborated upon in Section 4.

Seeing that weights are restricted to be non-negative to avoid extrapolation, the outcome variable for the treatment state has to fall within the convex hull of the outcomes for the donor states. As can be seen in Appendix B Figures B.1-B.10, this is the case for each control state. A treatment state with outliers falling outside of the convex hull of donor state outcomes would lead to synthetic control that is unable to match. In addition to this, as the synthetic control weights are selected by matching on the control variables these control variables must also fall within the convex hull of the donor states values for that control variable.

Lastly, the synthetic control method assumes conditional independence of the outcome. This entails that treatment application must not be dependent on potential outcomes after matching on the underlying factor structure (Hollingsworth & Wing, 2020). After conditioning on the selected control variables, treatment assignment should not depend on the outcome variable. Same-sex marriage legalization has been highly politicized in the Unites States (Willetts, 2011; Badgett, 2009) and it is not likely that the decision to implement it is dependent on the insurance coverage rates in a certain state.

4. Institutional background

This section will provide the institutional background and an explanation of the setting for each case-study included in this research. This research analyzes data for the period between 2002 and 2012 and includes case studies for four states that legalized same-sex marriage between 2009 and 2011. With the states of Iowa (same-sex marriage legalized 2009), Vermont (same-sex marriage legalized 2009), New Hampshire (same-sex marriage legalized 2010) and New York (same-sex marriage legalized 2011) as treated states. Massachusetts and Connecticut, two states that legalized some form of same-sex marriage before 2019 are excluded as their treatment years do not provide a long enough training period for the synthetic control method (see Section 3).

In the analysis the BRFSS survey years from 2002 to 2012 will be included. Earlier years are excluded as the changes in the policy environment over time changes the relationship that the synthetic control method finds between control variables and the outcome variables, which would prove an issue for the assumption of no dormant factors as explained in Section 3.3. Due to changes in the legal framework at a national level in 2013 (United States v. Windsor, 570 U.S. 744 (2013)) and 2014 (Obergefell v. Hodges, 576 U.S. 644 (2015)) the years after 2012 have not been included. Additionally, the timeline described in Carpenter et. al. (2021)

shows the rapid acceleration of same-sex marriage legalization in all states in the years after 2013.[1] If states implement same-sex marriage legalization in the same time-period they cannot be included in the donor pool. For any of the states legalizing same-sex marriage after 2012 the donor pool would be severely limited. This insufficiently large donor pool would make the synthetic control method unable to analyze these years.

Same-Sex Marriage legalization can take place through a number of different methods, by way of Judicial rulings, referenda or legislative and executive action (Willetts, 2011). This means that the 'treatment application' for same-sex marriage legalization can differ across states. From the four cases studied in this research Iowa is the only state where same-sex marriage was legalized through a judicial ruling. On appeal to a lower court ruling Iowa's State Supreme Court held that limiting marriage to opposite-sex couples violated the equal protection clause of the Iowa Constitution and marriage thus had to be opened to same-sex couples (Varnum v. Brien, 763 N.W.2d 862 (Iowa 2009)). Vermont, New Hampshire and New York enacted legislation at the state level that legalized same-sex marriage. Across different methods of same-sex marriage legalization, differences in the potential effect are not expected (Sherkat et. al., 2011).

5. Data

This section will provide the sources and codification of the variables used as well as an explanation of the classification of same-sex households in Section 5.1. For the outcome variable Health insurance coverage data from the Behavioral Risk Factor Surveillance System (BRFSS) is used (CDC, 2002-2012). The BRFSS is a large, annual, telephone-based survey coordinated by the Center for Disease Control in the United States (Pierannunzi, 2013). The annual sample size includes more than 400,000 individuals. The BRFSS includes data on whether an individual has health insurance coverage as a yes or no question. It does not specify the type of insurance. For this reason, the analysis of this paper will relate to the population percentage that is entirely uninsured, those that report having no form of health insurance. For the total population BRFSS sampling weights were utilized to compute the population percentage. For the same-sex household sub-population BRFSS sampling weights were reweighted over the sum of weights in their survey wave as shown in Simon, Soni, and Cawley (2017). This entails reconstructing each individual's sample weight as the fraction of that individuals assigned sample weight over the sum of all sample weights in the same-sex household subset for that year. As the synthetic control method does not employ standard

[1] For a further overview and timeline of same-sex marriage legalization please see Hansen, Martell, & Roncolato (2020), Table 1.

forms of statistical inference standard errors did not need to be clustered as they have no further role in the results or analysis.

The control variables as presented in Section 3.1 and Appendix Table A.1 derive from the following sources. The monthly state level unemployment rate is provided by the US Bureau of labor Statistics and yearly State level personal income per capita is derived from data of the US Bureau of Economic Analysis. Finally, Population percentages pertaining to Hispanic heritage, college education and degree attainment are retrieved from the U.S Census Bureau's 2002-2012 waves of the Current Population Survey.

5.1. Classification of same-sex households

The 2002 to 2012 waves of the BRFSS don't include direct data on respondent's sexual orientation. Unfortunately, there are no sufficiently large data sets available for this time period that include a direct question on this. Following Buchmueller and Carpenter (2010) and Carpenter et. al (2021) this analysis relies on respondents' answers on the composition of their household and their relationship status to proxy for same-sex households. Specifically, individuals that report living in a household with 2 same-sex individuals above the age of 18 are coded as a same-sex household. Carpenter et. al (2021) estimate that 11 to 29 percent of these same-sex households are likely to be sexual minorities (compared to 0.1 to 1.4 percent of other households).

6. Results

This research covers four separate states as case studies using the same methodology. For clarity each individual case study is presented in a separate section. Section 6.1 provides the most extensive discussion of general interpretation as this is applicable to all cases. Sections 6.1-6.4 all include graphical representation of the specific case study and reflection on their outcome. State weights and predictor variable means are provided in Appendix B and discussed in the following sections. The results in this section will provide the answer to the main research question and form the basis for the conclusion.

6.1. Iowa

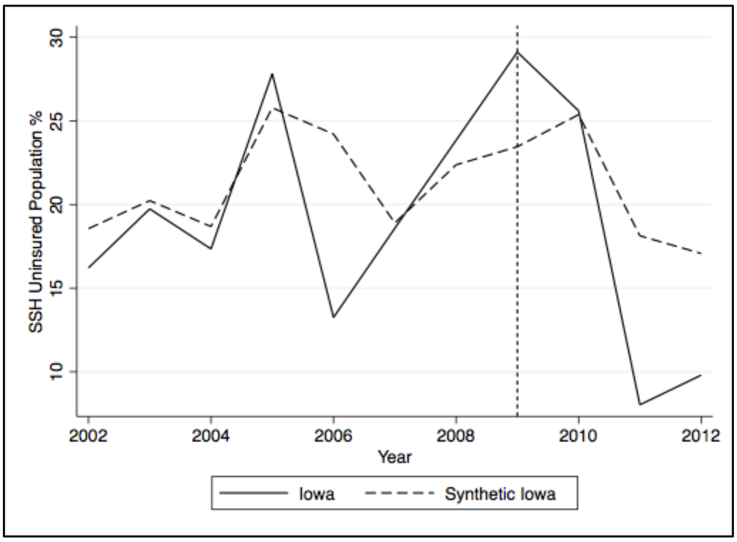


Figure 1 Trends in uninsured population percentage: Iowa vs. Synthetic Iowa for same-sex households, with treatment year 2009.

Figure 1 shows the development of the population percentage that is uninsured for the state of Iowa and the synthetic control created by the synthetic control method for the same-sex household (SSH) population. The applicable weights assigned to each donor state are provided in Table B.1 of Appendix B. The synthetic control consists of three states. Mississippi, Nebraska and Wisconsin are given a weight of 0.023, 0.576 and 0.401 respectively. This combination is found to best replicate the uninsured population percentage in Iowa for the pre-treatment period, minimizing the RMSPE. Figure 1 shows that the synthetic control approximates the true values for Iowa between 2002-2005 but there is clear deviation in 2006.

Table B.2 of Appendix B displays to what extent the synthetic Iowa matches the real prediction variable values of Iowa. The synthetic control generally approximates the true values for Iowa in these predictor means well, with the largest deviation for the Hispanic population percentage at 33% of the original value. Overall this is a positive measure for how well the synthetic Iowa is able to mimic the path that the health insurance rate for Iowa would follow without implementation of same-sex marriage legalization.

The estimated treatment effect is the gap in the uninsured population percentage between the synthetic Iowa and Iowa. We see a mixed trend with a small positive effect on the uninsured rate in 2009 and 2010 and a (larger) negative effect in 2011 and 2012. These results would suggest a mixed effect on the uninsured population percentage: with same-sex marriage legalization leading to health insurance coverage rates for individuals in same-sex households first decreasing and increase after two years.

To evaluate the significance of these effects, placebo tests where run on each donor state in the donor pool. By comparing the treatment effect of the treated state to that of the untreated donor states the significance of the effects that are found can be determined. The placebo tests are visualized in Figure B.1 of Appendix B. Computation of these placebo tests show that the effects that are found are not statistically significant. Statistical significance for each treatment year varies between 0.20 and 1. All year effects are well above a traditional significance level of 5% or 10%. As treatment distribution is not randomly assigned this P-value should not be regarded as a classical measure of significance. Even so, P-values of this magnitude are evidence that the effect that is found is not significant compared to effects for the donor pool and thus cannot be attributed as treatment effects.

There seem to be issues with the synthetic control fit as can be ascertained visually and in evaluating the standardized P-values. The large year over year variance in the uninsured population percentage could provide an issue in this case as this makes it more difficult to provide a tight synthetic control fit. As seen in Figure 1 the progression of the uninsured population percentage for same-sex households is irregular and has large transitory shocks between years, ranging between 9% and 28%. This irregularity in the outcome variable could be caused by an unbalanced weighting of survey respondents or issues with approximating same-sex couples by coding for same-sex households. As the BRFSS was not designed to approximate or evaluate the behavior of same-sex households this could be a reason of concern for inference on the basis of same-sex household subsets. For this reason, it can be beneficial to also evaluate the outcome for the total population. An analysis of the total population will include all same-sex couples, with the risk that a possible effect is overshadowed by variations for other population groups. The possible increase in synthetic control fit could however outweigh the negative side of analysis at a total population level.

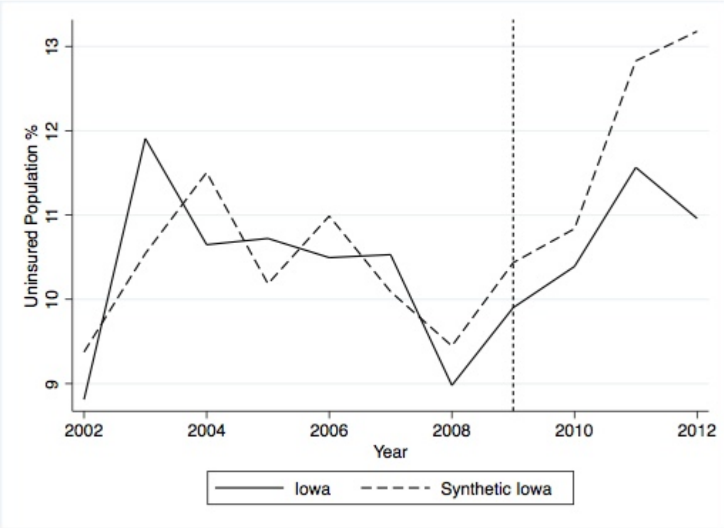


Figure 2 Trends in uninsured population percentage: Iowa vs. Synthetic Iowa, total population, with treatment year 2009.

Figure 2 shows the development of the population percentage that is uninsured for the state of Iowa and the synthetic control created by the synthetic control method for the full population. We see that this utilizes a different set of donor state weights in Table B.3 of Appendix B. The overall balance for control variables seems to improve slightly. When evaluating the gap in the uninsured population percentage between the synthetic and real Iowa we find negative effects. In the years after treatment the population percentage that is uninsured drops compared to a setting without same-sex marriage legalization. Using the placebo analysis as outlined above standardized P-values ranging from 0.11 (for 2012) to 0.67 (for 2009) are found. Visual outputs for the Placebo analysis can be found in Figure B.3 of Appendix B. The effect for 2012 of minus 2.2 percentage points approaches a 10% significance level. Even so this is once again not significant at a traditional level, with recognition to the interpretation as provided earlier in this section.

The progression of the uninsured population percentage for the total population is smoother and contains smaller transitory shocks between years than that of same-sex households. This would support the argument that restricting BRFSS data to same-sex households is problematic as the survey weights are not constructed with this population group in mind or this sub-group captures a subpopulation that is highly irregular (O'Connell & Lofquist, 2009). A second potential explanation for the poor fit of the earlier analysis could be that there are other variables that have a strong effect on health insurance coverage for same-sex households that are not accounted for in this analysis. This implies that for same-sex households the underlying relationship between the control variables and the outcome variable is significantly different than for the general population. Meaning that the predictive power of the included controls and their correlation with the outcome variable are not able to provide a close match for the synthetic control of same-sex households as there are other, unobserved, factors that affect health insurance coverage.

To conclude, for the state of Iowa the synthetic control method is unable to capture a significant effect of same-sex marriage legalization on the health insurance coverage at any population level. Evaluating the total population provides a better matching synthetic control as also indicated by a lower RMSPE (Table B.1 and B.3 Appendix B) but still does not match tightly. The differences between total population and same-sex households population analysis that are observed will be further discussed in Section 7.

6.2. Vermont

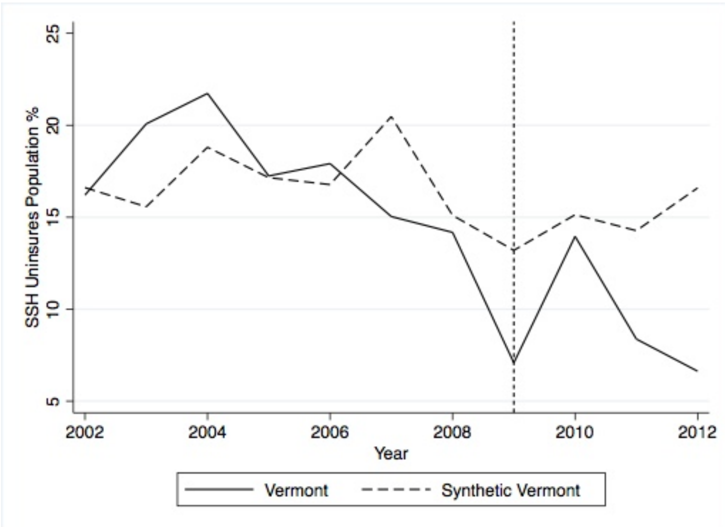


Figure 3 Trends in uninsured population percentage: Vermont vs. Synthetic Vermont for same-sex households, with treatment year 2009.

Figure 3 plots the development of the population percentage that is uninsured for the state of Vermont and the synthetic control created by the synthetic control method for the same-sex household population. For Vermont the applicable weights assigned to each donor state are provided in Table B.5 of Appendix B. The synthetic control consists of 6 states: California, Delaware, Maine, North Dakota, Oregon and Virginia. We see the synthetic control is only able to approximates the true values for Vermont in 2002, 2005, 2006 and 2008 in other years there is clear deviation.

Table B.6 of Appendix B shows the mean values for the control variables used as predictors. We see that for the majority of control variables there is a fair match. For the population percentage that is Hispanic the synthetic control is unable to match correctly, with a factor 8 increase. This can partially be explained by the extremely small Hispanic population in Vermont: it has the second lowest Hispanic population percentage of all U.S. states (U.S. Census bureau, CSP). Matching on this variable is a challenge for this specific case study. This provides a potential violation of the assumption that all control variables must fall within a convex hull of those variables for the donor states as discussed in Section 3.3.

The visual deviation, unbalanced matching on control variables and high RMSPE relative to the mean outcome variable are indications of a bad synthetic control fit. The synthetic Vermont is unable to adequately mimic the path that the health insurance rate for Vermont would follow without implementation of same-sex marriage legalization.

The outcome variable gaps suggest a negative effect on the uninsured population percentage: with same-sex marriage legalization leading to an increase in health insurance coverage for the same-sex household population as visualized in Figure B.3 of Appendix B. Placebo tests find high standardized P-values for each year except 2012. There is a negative 9.97 percentage point decrease in the uninsured rate in 2012. We can expect such an effect to be very large compared to a majority of the placebo tests and thus gain a low P-value. For 2012 the placebo tests compute a standardized P-value of 0.116, slightly above the 10% level. A standardized P-value at or near the 10% level cannot be interpreted as a traditionally significant result. To evaluate the design and methodological set up the same analysis was computed for the total population, following the same reasoning as presented in Section 6.1.

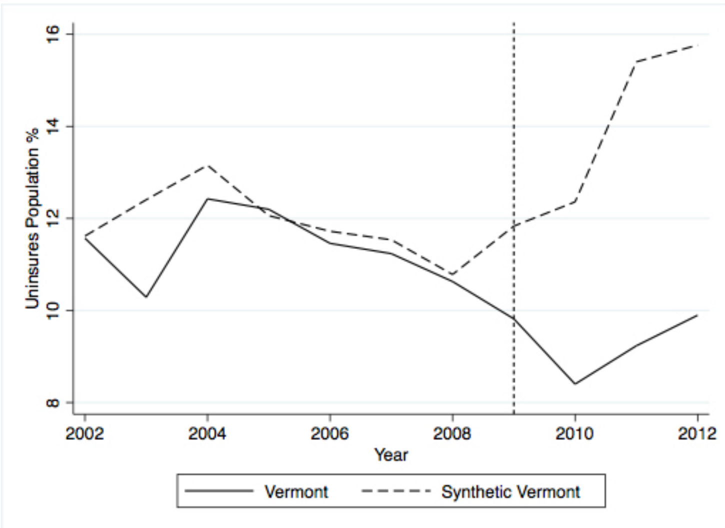


Figure 4 Trends in uninsured population percentage: Vermont vs. Synthetic Vermont, total population, with treatment year 2009.

The total population analysis provides a smoother progression of the outcome variable and provides an improved synthetic control fit. Figure 4 shows the development of the population percentage that is uninsured for the full population. Table B.7 of Appendix B shows the state weights for this analysis. Here we see Delaware and Virginia receiving different weights than in the same-sex household analysis and a change entirely in the make-up of the other weighted states.

Control variable balance, shown in Table B.8 of Appendix B, is comparable to that of the same-sex household analysis, with similar problems for the Hispanic population percentage. The gap between the synthetic and real Vermont shows a clear negative treatment effect. Placebo analysis as provided in Figure B.4 of Appendix B lead to computation of standardized P-values for this effect. These P-values correspond with the increase in fit that is visible.

Table 1. Estimated treatment effects and P-values, Vermont.

Year	Estimated treatment effect in percentage points	P-value	Standardized P-value
2009	-2.01	0.093	0.093
2010	-3.96	0.0235	0.000
2011	-6.17	0.0465	0.023
2012	-5.87	0.0233	0.047

Table 1 shows the estimated treatment effects, P-value and standardized P-value for the synthetic control analysis for Vermont's total population. For 2010-2012 the standardized P-values correspond with traditional significance at the 5% level or lower. For 2009 significance of just under 10% is found. In combination with the relatively low RMSPE in Table B.10 of Appendix B this suggests a significant negative effect of same-sex marriage legalization on the uninsured population percentage.

To control the robustness of this effect further analysis was done where the treatment period was backdated. Backdating the treatment means instructing the synthetic control method to test for a treatment effect from 2006 onwards, three years before the actual treatment effect. The backdated synthetic control computation is presented in Figure B.5 of Appendix B. This backdated analysis shows an upward bias of the synthetic control even before same-sex marriage legalization (2009), with an effect already visible in 2007. Because the timing of the treatment effect is not robust to backdating the effect that is found is seen as less credible (Abadie, 2021).

An additional point of concern in interpreting this effect is that the effect that is found for the total population does not correspond to that of the same-sex household population as theorized. If this effect was driven by the mechanism as outline in Section 2 the same-sex household population would be expected to show this effect at a higher rate than the total population. For this reason, caution is warranted in the interpretation of these results. This will be further elaborated upon in Section 7.

As described in Section 2, earlier research finds a stronger effect specifically for the male same-sex households as opposed to female same-sex households. To examine if this is the case here, an additional analysis for the male same-sex households is executed. Figure B.6 of Appendix B shows the development of the population percentage that is uninsured if data is limited to only male same-sex households. There is a negative effect, but this is not significant because the pre-treatment fit is unable to match the highly irregular progression of the outcome variable. Analysis for only male same-sex households gives similar results to that of all same-sex households and shows the same limitations. This outcome further supports the points raised at the close of Section 6.1 that either the survey weights or make-up of the same-sex household subpopulation is insufficient or the underlying determinants of health insurance coverage for same-sex households differ to that of the general population.

6.3. New Hampshire

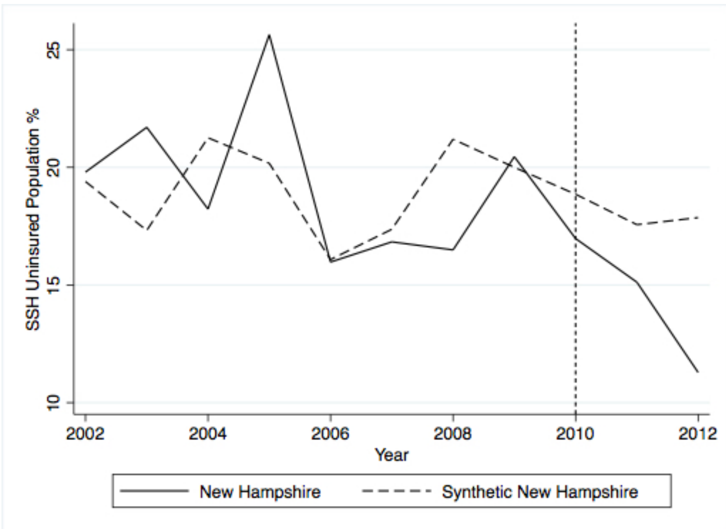


Figure 5 Trends in uninsured population percentage: New Hampshire vs. Synthetic New Hampshire for same-sex households, with treatment year 2010.

Figure 5 plots the development of the population percentage that is uninsured for New Hampshire as well as its synthetic counterpart. Table B.8 of Appendix B shows the 6 states that receive a weight in the construction of the synthetic control New Hampshire. We see the synthetic control is unable to match the outcome variable for all pre-treatment years other than 2006, 2007 and 2009. Table B.9 of Appendix B shows the variable balance of the control variables. This balance is fair but does not match closely for variables other than the Unemployment Rate and the lagged uninsured population percentage. The estimates treatment effects for the state of New Hampshire are negative. Similar to in Section 6.1 placebo tests to the donor pool as provided in Figure B.7 of Appendix B, and the standardized P-value computation show that none of these effects are significant. Issues and inference are comparable to Section 6.1.

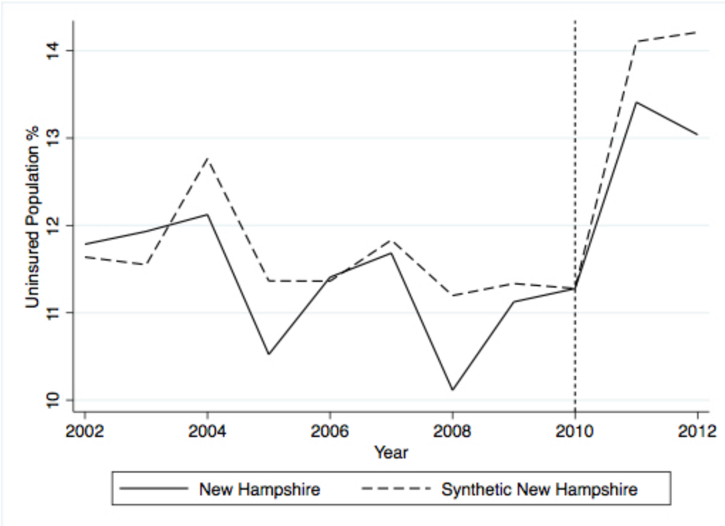


Figure 6 Trends in uninsured population percentage: New Hampshire vs. Synthetic New Hampshire, total population, with treatment year 2010.

Figure 6 shows the trend for New Hampshire when not restricted to the same-sex household population. Table B.10 of Appendix B shows that of the states receiving a weight in the synthetic control for the same-sex household population only New Jersey remains. The predictor mean balance as provided in Table B.11 is comparable to that of Table B.9 for the same-sex household population. Synthetic control fit improves both visually as well as with a reduced relative RMSPE in Table B.10 compared to Table B.8. Even with these improvements to the fit the placebo tests visualized in Figure B.8 produce no P-values, standardized or otherwise, that are significant at a level below 30%. Comparable to Section 6.1, this analysis is unable to provide a conclusion to the effect of same-sex marriage legalization for the state of New-Hampshire.

6.4. New York

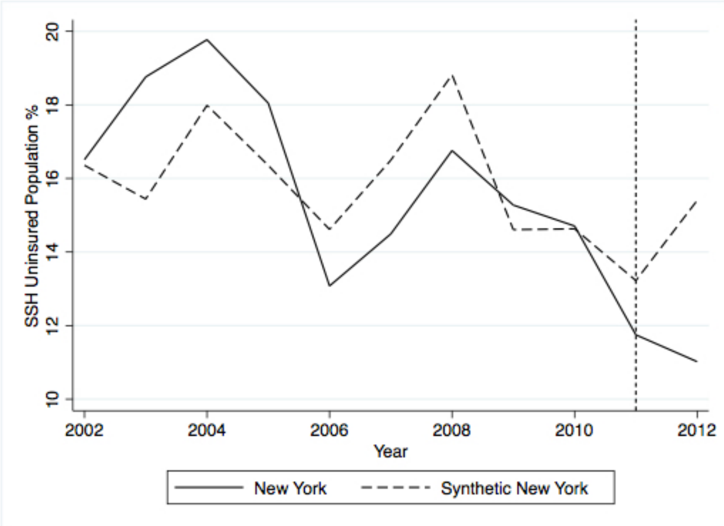


Figure 7 Trends in uninsured population percentage: New York vs. Synthetic New York for same-sex households, with treatment year 2011.

Figure 7 shows the development of the population percentage that is uninsured for New York as the final case study in this paper. The five states that form this synthetic control are seen in Table B.12 in Appendix B. The synthetic control fit and predictor balance as presented in Table B.13 of Appendix B is comparable to that found in Section 6.2. As New York's same-sex marriage legalization occurred in the before last year of the timeline in this paper there are only two years where a treatment effect can be ascertained. For both years the treatment effect is negative indicating increased health insurance coverage after same-sex marriage legalization. Pursuant to the placebo tests provided in Figure B.9 of Appendix B the P-value of these effects are not significant and lead to the same conclusion as same-sex household population analysis for Sections 6.1-6.3.

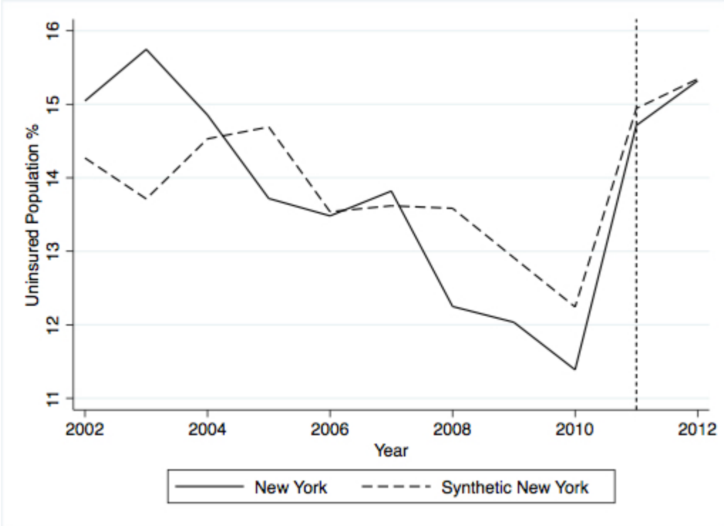


Figure 8 Trends in uninsured population percentage: New York vs. Synthetic New York, total population, with treatment year 2011.

Figure 8 provides the visualization of New York's total uninsured population percentage compared to that of its synthetic control. Table B.14 and B.15 of Appendix B provide state weights and control variable means for this iteration. The increase in synthetic control fit as compared to analysis for the same-sex household population is less pronounced than in sections 6.1 to 6.3. For the two post-treatment years there is a very small but negligible negative effect of same-sex marriage legalization. In accordance with the placebo tests as presented in Figure B.10 of Appendix B the standardize P-values suggest significance levels of this result of higher than 80%. For the state of New York this analysis finds no significant effects of same-sex marriage legalization on the health insurance coverage rate.

7. Discussion and conclusion

This research analyzes the effect of same-sex marriage legalization on the health insurance coverage for same-sex households in the United States. Four case studies were completed to evaluate the effect of legalization in Iowa, Vermont, New Hampshire and New York through application of the synthetic control method. For Iowa, New Hampshire and New York no significant effects where found. Across both the same-sex household population and the total population the case studies found negative effects but not at a significant level. Vermont, in an analysis of its total population is the exception to this. The treatment effect was found by comparing the outcome variables to a synthetic control state made up of a weighted combination of donor states selected through the synthetic control method. Compared to this synthetic control Vermont saw a 2.01, 6.17 and 5.87 percentage point decrease in the reported uninsured population percentage in the first, third and fourth year of same-sex marriage legalization in that state. The decreases in these years where significant at a 5% level. The decrease in uninsured population percentages suggests that same-sex marriage legalization leads to increased health insurance coverage for the general population in Vermont. However, this effect was not robust to backdating and a specification to only male same-sex households did not find any significant effect either. Furthermore, the mechanism behind this treatment effect would be expected to show at a higher rate for the same-sex household population than the general population, but the treatment effect for this sub-population is not significant.

In the analysis of the different case studies both the same-sex household population, as coded through the BRFSS, as well as the general population was analyzed. This research is unable to draw a conclusion to the effect of same-sex marriage legalization on health insurance coverage for the same-sex households. The state hypothesis of a positive effect is not supported, and this research finds no conclusive evidence that such an effect does or does not exist.

A further examination of the results and the synthetic control fit provides additional implications of this research. The difference between same-sex household population analysis and total population analysis show that the health insurance rates for same-sex households in the BRFSS follow a trajectory that has larger transitory shocks and is less smooth. For each case study the synthetic control fit for the same-sex household population is relatively poor, even though it is matched on control variables that are generally accepted to be good determinants of the health insurance coverage rate. This could be driven by issues in approaching the same-sex population by coding for same-sex households. Coding for same-sex households could inadvertently include a highly irregular subpopulation, or BRFSS weights for these respondents could provide unbalanced survey weights. Here the limited availability of data on the topic of sexual orientation proves to be a limiting factor of this paper. Another possible reason for this issue is more profound: the underlying relationship between the control variables and the health insurance coverage rate for those living in same-sex households could significantly differ from that of the general population.

As is clear in Section 6 synthetic control analysis for same-sex households has difficulty providing a close fit. The imperfect fit suggests a violation of the assumption that the observed factors are able to select a close match on relevant unobserved factors as this does not seem to hold for the same-sex household population. This means that the included control variables are not sufficient and there are relevant unobserved factors not captured by selecting a counterfactual on the basis of observable control variables. This would entail that other papers that don't include a broader set of control variables do not fulfil the necessary, and often heightened, assumptions for the relevant model. The violation of this assumption provides an important limitation to the interpretation of earlier research as has been presented in section 2 and an important factor to be taken into account for further research on this topic.

An additional limitation of this research is the issue of external validity. The application of this research rests strongly on the specific policy landscape of the United States where health-insurance is not mandated and is strongly driven by employer sponsoring. As such it mostly provides insight for the ongoing debate on the effect of nationwide same-sex marriage legalization in the United States. However, these insights can still be relevant to further elaborate analysis methods for other benefits of marriage, also in other jurisdictions.

Finally, one of the main recommendations for future research stems from the problematic nature of same-sex household codification as a means to approach analysis relating to LGBT individuals and same-sex couples. This conclusion provides a critique of earlier research such as Carpenter et. al. (2021) but also drives home one of the suggestions for future research given in that research. The limitation provided by not having sufficient data on same-sex relationships severely limits policy analysis in this field. For future research, also in other countries, data collection on sexual orientation as well as a broadening of existing relationship questions in large scale surveys is suggested. Additionally, further research into the determinants of health insurance coverage specifically for same-sex households could confirm or dismiss the questions raised about differences in the underlying determinants of health insurance coverage for this subpopulation. Both these extensions can lead to a better understanding and analysis of the effects of policies such as same-sex marriage legalization. This could also be used in analysis of many other socio-legal constructs and their economic impact on the LGBT population.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association 105* (490): 493–505.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country*. American economic review*, 93(1), 113-132.

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2), 391-425.Badgett, M. L. (2009). When gay people get married: What happens when societies legalize same-sex marriage. NYU Press.

Badgett, M. L. (2020). The Economic Case for LGBT Equality: Why Fair and Equal Treatment Benefits Us All. Beacon Press.

Bouttell, J., Craig, P., Lewsey, J., Robinson, M., & Popham, F. (2018). Synthetic control methodology as a tool for evaluating population-level health interventions. *Journal of Epidemiology & Community Health*, 72(8), 673-678.

Buchmueller, T., & Carpenter, C. S. (2010). Disparities in health insurance coverage, access, and outcomes for individuals in same-sex versus different-sex relationships, 2000–2007. *American journal of public health*, 100(3), 489-495.

Buffie, W. C. (2011). Public health implications of same-sex marriage. *American Journal of Public Health*, 101(6), 986-990.

Carpenter, C. S., Eppink, S. T., Gonzales, G., & McKay, T. (2021). Effects of Access to Legal Same-Sex Marriage on Marriage and Health. *Journal of Policy Analysis and Management.*

CDC (2002-2012). Behavioral Risk Factor Surveillance System Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2002-2012.

Cebula, R. J. (2006). A further analysis of determinants of health insurance coverage. *International Advances in Economic Research*, 12(3), 382-389.

Chernew, M., Cutler, D. M., & Keenan, P. S. (2005). Increasing health insurance costs and the decline in insurance coverage. *Health services research*, 40(4), 1021-1039.

Delhommer, S. M., & Hamermesh, D. S. (2021). Same-Sex Couples and the Gains to Marriage: The Importance of the Legal Environment. *Journal of Policy Analysis and Management*.

Downing, J., & Cha, P. (2020). Same-Sex Marriage and Gains in Employer-Sponsored Insurance for US Adults, 2008–2017. *American journal of public health*, 110(4), 537-539.

Fremeth, A., Holburn, G. L., & Richter, B. K. (2013). Making Causal Inferences in Small Samples using Synthetic Control Methodology: Did Chrysler Benefit from Government Assistance? *Available at SSRN 2135294*.

Galiani, S., & Quistorff, B. (2017). The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal*, 17(4), 834-849.

Hansen, M. E., Martell, M. E., & Roncolato, L. (2020). A labor of love: The impact of same-sex marriage on labor supply. *Review of Economics of the Household*, 18(2), 265-283.

Hollingsworth, A., & Wing, C. (2020). Tactics for design and inference in synthetic control studies: An applied example using high-dimensional data. *Available at SSRN 3592088*.

Kaiser Family Foundation and Health Research and Educational Trust. (2004). *Employer Health Benefits, 2004 Annual Survey*. 2004 http://www.kff.org/insurance/7148/upload/2004-Employer-Health-Benefits-Survey-Full-Report.pdf. (accessed June 3, 2021).

Keisler-Starkey, K., & Bunch, L. N. (2020*).* Health insurance coverage in the United States: 2019. *United States Census Bureau.*

Sherkat, D. E., Powell-Williams, M., Maddox, G., & De Vries, K. M. (2011). Religion, politics, and support for same-sex marriage in the United States, 1988–2008. *Social Science Research,* 40(1), 167-180.

Simon, K., Soni, A., & Cawley, J. (2017). The impact of health insurance on preventive care and health behaviors: Evidence from the first two years of the ACA Medicaid expansions. *Journal of Policy Analysis and Management*, 36, 390–417.

Smith, M. (2008). Political institutions and lesbian and gay rights in the United States and Canada. Routledge.

U.S. Bureau of Economic Analysis (2002-2012), "SAINC1 Personal Income Summary" https://apps.bea.gov/iTable/iTable.cfm?reqid=70&step=30&isuri=1&year_end=-1&acrdn=1&classification=non-industry&state=0&yearbegin=-1&unit_of_measure=levels&major_area=0&area=xx&year=2012,2011,2010,2009,2008,2007,2006,2005,2004,2003,2002&tableid=21&category=421&area_type=0&statistic=3 (accessed May 20, 2021).

U.S. Census Bureau (2002-2012), Current Population Survey, generated by Wouter van Dam; using https://www.census.gov/cps/data/cpstablecreator.html, (accessed May 27, 2021)

O'Connell, M., & Lofquist, D. (2009, April). Counting same-sex couples: Official estimates and unofficial guesses. *Annual Meetings of the Population Association of America*, Detroit, MI.

Pierannunzi, C., Hu, S. S., & Balluz, L. (2013). A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004–2011. *BMC medical research methodology*, 13(1), 1-14.

Tumin, D., & Kroeger, R. A. (2020). Private Health Insurance Coverage Among Same-Sex Couples After Obergefell v. Hodges. *LGBT health*, 7(5), 264-270.

Willetts, M. C. (2011). Registered domestic partnerships, same-sex marriage, and the pursuit of equality in California. *Family Relations* 60(2), 135-149.

Appendix A

Descriptive statistics data

Table A.1. Data description.

Variables	Source	Years	Description
Uninsured Population percentage	*BRFSS*	2002-2012	Weighted percentage of BRFSS respondents indicating they do not have 'any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare?'
Unemployment rate	*BLS*	2002-2012	Monthly state level unemployment rate in %
Personal income per capita	*BEA*	2002-2012	State level total personal income in dollars divided by total midyear population
Hispanic population %	*CPS*	2002-2012	Percentage of the population that identifies as Hispanic
Population % with college education	*CPS*	2002-2012	Percentage of the population with education exceeding High-School education
Population % with a college degree	*CPS*	2002-2012	Percentage of the population with a college degree

Note. BRFSS is the Behavioral Risk Factor Surveillance System provided by the CDC (2002-2012), *BLS* is the US Bureau of labor Statistics accessed through the Public Data Application Programming Interface (API)[2], *Bea* is the US Bureau of Economic Analysis, *CPS* is the Current Population Survey retrieved through the US Census Bureau (2002-2012).

[2] BLS.gov cannot vouch for the data or analyses derived from these data after the data have been retrieved from BLS.gov.

Appendix B

Iowa, same-sex households

Table B.1. State weights in the synthetic Iowa for same-sex households.

State	Weight	State	Weight
Alabama	0	Montana	0
Alaska	0	Nebraska	0.576
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	-
California	0	New Jersey	0
Colorado	0	New Mexico	0
Connecticut	-	New York	-
Delaware	0	North Carolina	0
District of Columbia	-	North Dakota	0
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	*	South Dakota	0
Kansas	0	Tennessee	0
Kentucky	0	Texas	0
Louisiana	0	Utah	0
Maine	0	Vermont	-
Maryland	0	Virginia	0
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0.401
Mississippi	0.023	Wyoming	0
Missouri	0		
RMSPE:		4.392	

Table B.2. Predictor means for Iowa, same-sex households.

	Iowa	
Variables	Real	Synthetic
Unemployment rate	4.129	4.197
Personal income per capita	33,148	34,804
Hispanic population %	4.867	6.485
Population % with college education	23.642	23.258
Population % with a college degree	17.254	17.792
% uninsured 2008	23.842	22.360
% uninsured 2005	27.823	25.856
% uninsured 2002	16.213	18.593

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2008, variables. See Table A.1. for the description and units for each variable.

a b

Figure B.1. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for Iowa and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents Iowa.

Iowa, total population

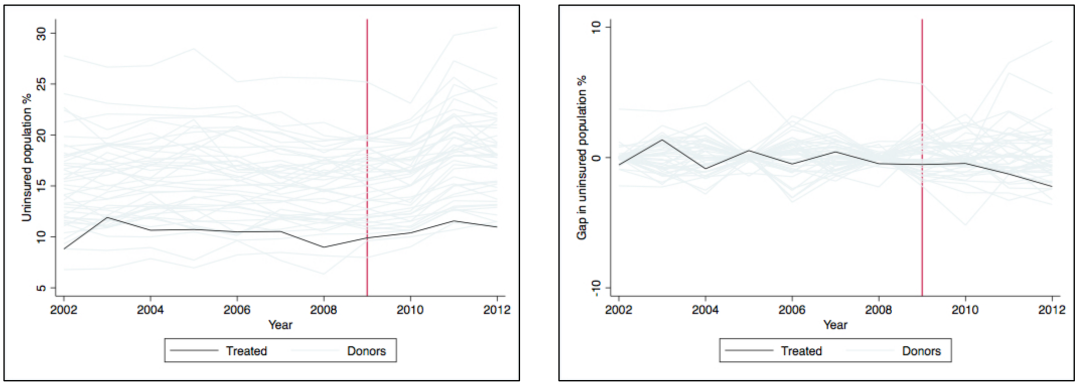
Table B.3. State weights in the synthetic Iowa for total BRFSS population.

State	Weight	State	Weight
Alabama	0	Montana	0
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	-
California	0	New Jersey	0
Colorado	0	New Mexico	0
Connecticut	-	New York	-
Delaware	0.417	North Carolina	0
District of Columbia	-	North Dakota	0.435
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	*	South Dakota	0
Kansas	0	Tennessee	0
Kentucky	0.148	Texas	0
Louisiana	0	Utah	0
Maine	0	Vermont	-
Maryland	0	Virginia	0
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0
Mississippi	0	Wyoming	0
Missouri	0		
RMSPE:		0.742	

Table B.4. Predictor means for Iowa, total population.

	Iowa	
Variables	Real	Synthetic
Unemployment rate	4.129	3.884
Personal income per capita	33,148	35,672
Hispanic population %	4.867	4.485
Population % with college education	23.642	22.901
Population % with a college degree	17.254	19.065
% uninsured 2008	8.980	9.447
% uninsured 2005	10.722	10.184
% uninsured 2002	8.811	9.375

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2008, variables. See Table A.1. for the description and units for each variable.



a b

Figure B.2. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for Iowa and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents Iowa.

Vermont

Vermont, same-sex households

Table B.5. State weights in the synthetic Vermont for same-sex households.

State	Weight	State	Weight
Alabama	0	Montana	0
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	-
California	0.05	New Jersey	0
Colorado	0	New Mexico	0
Connecticut	-	New York	-
Delaware	0.169	North Carolina	0
District of Columbia	-	North Dakota	0.015
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0.012
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	-	South Dakota	0
Kansas	0	Tennessee	0
Kentucky	0	Texas	0
Louisiana	0	Utah	0
Maine	0.261	Vermont	*
Maryland	0	Virginia	0.493
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0
Mississippi	0	Wyoming	0
Missouri	0		
RMSPE:		2.946	

Table B.6. Predictor means for Vermont, same-sex households.

	Vermont	
Variables	Real	Synthetic
Unemployment rate	3.986	4.181
Personal income per capita	35,668	38,044
Hispanic population %	0.835	6.292
Population % with college education	19.103	19.943
Population % with a college degree	24.503	20.847
% uninsured 2008	14.177	15.105
% uninsured 2005	17.245	17.151
% uninsured 2002	16.205	16.611

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2008, variables. See Table A.1. for the description and units for each variable.

a b

Figure B.3. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for Vermont and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents Vermont.

Vermont, total population

Table B.7. State weights in the synthetic Vermont for total BRFSS population.

State	Weight	State	Weight
Alabama	0	Montana	0.048
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	-
California	0	New Jersey	0
Colorado	0	New Mexico	0
Connecticut	-	New York	-
Delaware	0.14	North Carolina	0
District of Columbia	-	North Dakota	0
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	-	South Dakota	0
Kansas	0.445	Tennessee	0
Kentucky	0	Texas	0
Louisiana	0	Utah	0
Maine	0	Vermont	*
Maryland	0	Virginia	0.367
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0
Mississippi	0	Wyoming	0
Missouri	0		
RMSPE:		0.865	

Table B.8. Predictor means for Vermont, total population.

	Vermont	
Variables	Real	Synthetic
Unemployment rate	3.986	4.303
Personal income per capita	35,668	36,743
Hispanic population %	0.835	6.407
Population % with college education	19.103	21.171
Population % with a college degree	24.503	21.205
% uninsured 2008	10.629	10.782
% uninsured 2005	12.198	12.061
% uninsured 2002	11.571	11.618

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2008, variables. See Table A.1. for the description and units for each variable.

a b

Figure B.4. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for Vermont and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents Vermont.

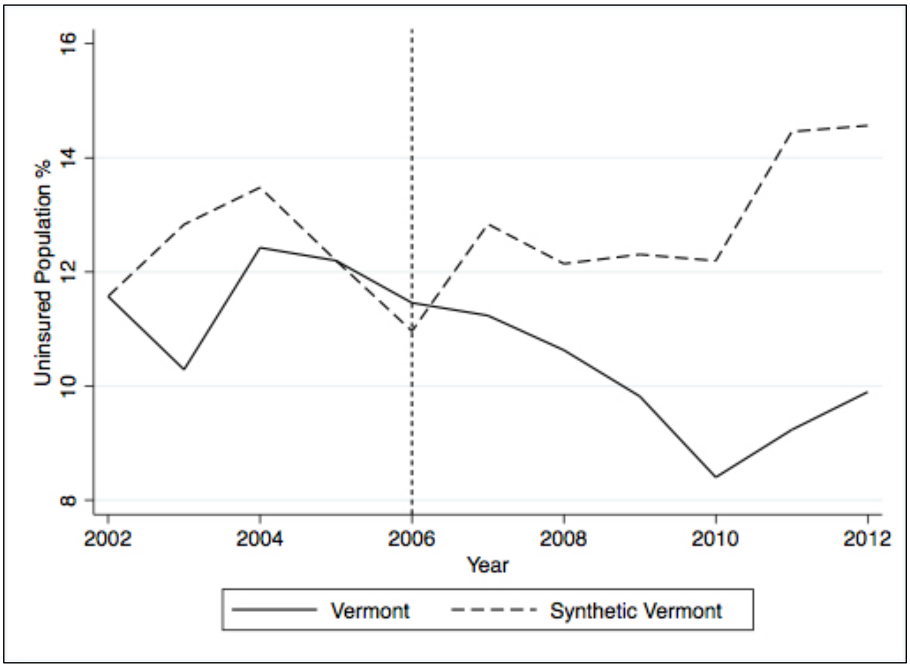


Figure B.5. Trends in the uninsured population percentage: Vermont vs. Synthetic Vermont, total population. Treatment backdated to 2006.

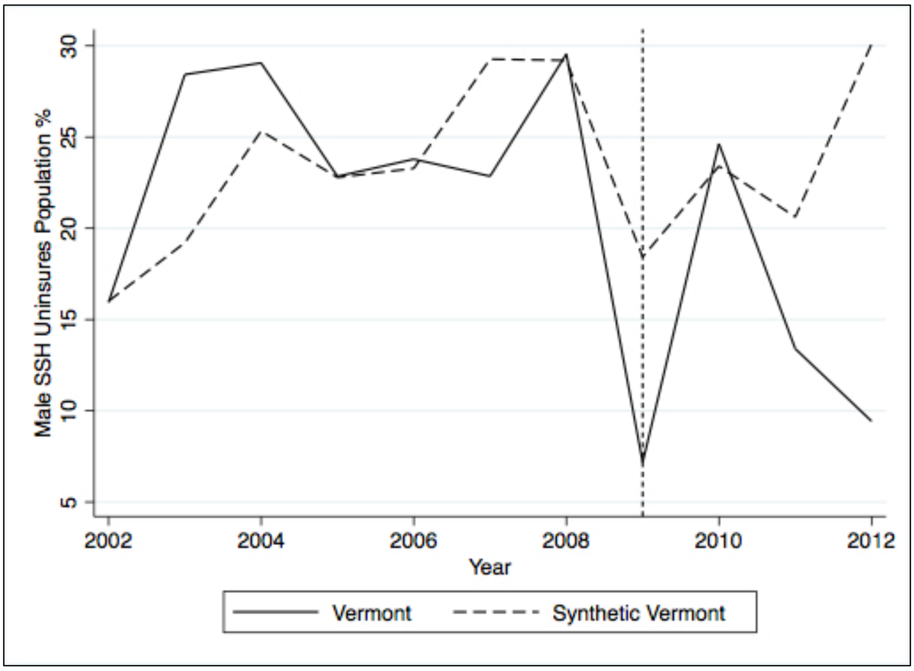


Figure B.6. Trends in the uninsured population percentage: Vermont vs. Synthetic Vermont for male same-sex households.

New Hampshire

New Hampshire, same-sex households

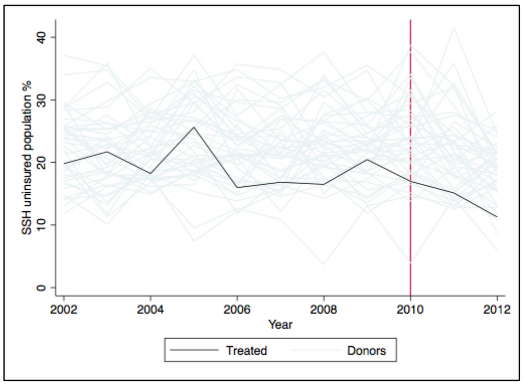
Table B.8. State weights in the synthetic New Hampshire for same-sex households population.

State	Weight	State	Weight
Alabama	0.133	Montana	0
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	*
California	0	New Jersey	0.16
Colorado	0	New Mexico	0
Connecticut	-	New York	-
Delaware	0.013	North Carolina	0
District of Columbia	-	North Dakota	0
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	-	South Dakota	0
Kansas	0	Tennessee	0
Kentucky	0	Texas	0
Louisiana	0	Utah	0.188
Maine	0	Vermont	-
Maryland	0.464	Virginia	0
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0
Mississippi	0	Wyoming	0.042
Missouri	0		
RMSPE:		3.182	

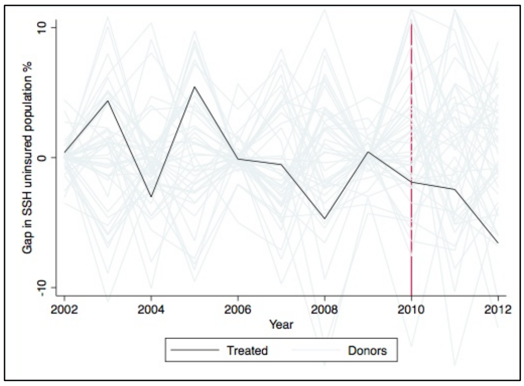
Table B.9. Predictor means for New Hampshire, same-sex households.

	New Hampshire	
Variables	Real	Synthetic
Unemployment rate	4.150	4.833
Personal income per capita	41,666	39,774
Hispanic population %	2.063	8.584
Population % with college education	21.987	19.882
Population % with a college degree	23.879	21.708
% uninsured 2009	20.451	20.001
% uninsured 2006	15.975	16.083
% uninsured 2002	19.798	19.398

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2009, variables. See Table A.1. for the description and units for each variable.



a



b

Figure B.7. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for New Hampshire and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents New Hampshire.

New Hampshire, total population

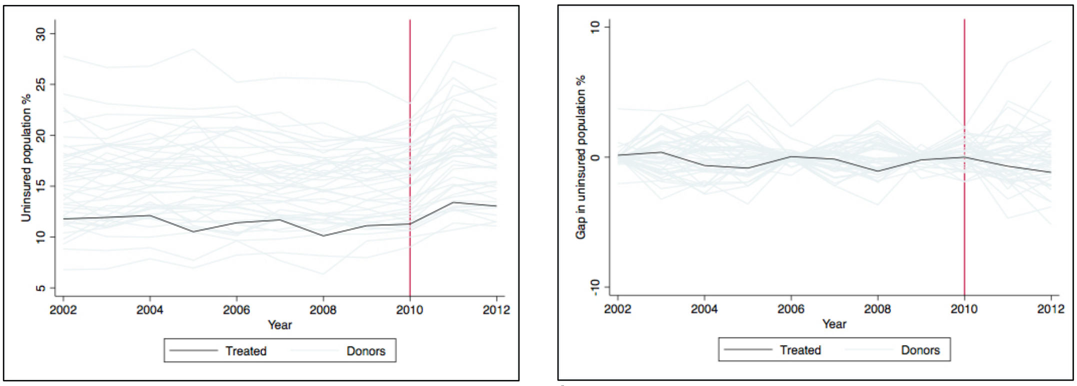
Table B.10. State weights in the synthetic New Hampshire for total BRFSS population.

State	Weight	State	Weight
Alabama	0	Montana	0
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	*
California	0	New Jersey	0.289
Colorado	0	New Mexico	0
Connecticut	-	New York	-
Delaware	0	North Carolina	0
District of Columbia	-	North Dakota	0
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	-	South Dakota	0.1
Kansas	0	Tennessee	0
Kentucky	0	Texas	0
Louisiana	0	Utah	0
Maine	0.008	Vermont	-
Maryland	0	Virginia	0.417
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0.186	Wisconsin	0
Mississippi	0	Wyoming	0
Missouri	0		
RMSPE:		0.561	

Table B.11. Predictor means for New Hampshire, total population.

	New Hampshire	
Variables	Real	Synthetic
Unemployment rate	4.150	4.612
Personal income per capita	41,666	41,303
Hispanic population %	2.063	8.082
Population % with college education	21.987	19.897
Population % with a college degree	23.879	23.217
% uninsured 2009	11.1254	11.335
% uninsured 2006	11.407	11.362
% uninsured 2002	11.785	11.638

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2009, variables. See Table A.1. for the description and units for each variable.



a b

Figure B.8. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for New Hampshire and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents New Hampshire.

New York

New York, same-sex households

Table B.12. State weights in the synthetic New York for same-sex households.

State	Weight	State	Weight
Alabama	0.002	Montana	0
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	-
California	0.138	New Jersey	0
Colorado	0	New Mexico	0
Connecticut	-	New York	*
Delaware	0.095	North Carolina	0
District of Columbia	-	North Dakota	0
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0.637
Indiana	0	South Carolina	0
Iowa	-	South Dakota	0
Kansas	0	Tennessee	0
Kentucky	0	Texas	0
Louisiana	0	Utah	0
Maine	0	Vermont	-
Maryland	0.127	Virginia	0
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0
Mississippi	0	Wyoming	0
Missouri	0		
RMSPE:		1.766	

Table B.13. Predictor means for New York, same-sex households.

	New York	
Variables	Real	Synthetic
Unemployment rate	6.078	6.404
Personal income per capita	43,426	39,646
Hispanic population %	16.198	13.440
Population % with college education	18.476	19.427
Population % with a college degree	22.288	21.709
% uninsured 2010	14.706	14.631
% uninsured 2006	13.078	14.614
% uninsured 2002	16.512	16.358

Note. All variables, except for the lagged Population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2010, variables. See Table A.1. for the description and units for each variable.

a b

Figure B.9. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for New York and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents New York.

New York, total population

Table B.14. State weights in the synthetic New York for total BRFSS population.

State	Weight	State	Weight
Alabama	0	Montana	0
Alaska	0	Nebraska	0
Arizona	0.026	Nevada	0
Arkansas	0	New Hampshire	-
California	0	New Jersey	0.599
Colorado	0	New Mexico	0
Connecticut	-	New York	*
Delaware	0	North Carolina	0
District of Columbia	-	North Dakota	0
Florida	0	Ohio	0
Georgia	0	Oklahoma	0
Hawaii	-	Oregon	0
Idaho	0	Pennsylvania	0
Illinois	0	Rhode Island	0
Indiana	0	South Carolina	0
Iowa	-	South Dakota	0
Kansas	0	Tennessee	0
Kentucky	0	Texas	0.044
Louisiana	0	Utah	0
Maine	0.331	Vermont	-
Maryland	0	Virginia	0
Massachusetts	-	Washington	0
Michigan	0	West Virginia	0
Minnesota	0	Wisconsin	0
Mississippi	0	Wyoming	0
Missouri	0		
RMSPE:		1.007	

Table B.15. Predictor means for New York, total population.

	New York	
Variables	Real	Synthetic
Unemployment rate	6.078	5.839
Personal income per capita	43,426	41,781
Hispanic population %	16.198	12.445
Population % with college education	18.476	18.596
Population % with a college degree	22.288	22.323
% uninsured 2010	11.388	12.243
% uninsured 2006	13.482	13.538
% uninsured 2002	15.046	14.269

Note. All variables, except for the lagged population percentage that is uninsured (% uninsured YYYY), are averaged over the periods 2002 to 2010, variables. See Table A.1. for the description and units for each variable.

a b

Figure B.10. Population percentage that is uninsured (a), and gaps for this variable between the synthetic and real state (b) for New York and all states in the donor pool.

Note. The light lines represent each individual state in the donor pool, the dark line represents New York.