Measuring the Social and Externality Benefits of Influenza Vaccination *

Corey White
California Polytechnic State University, San Luis Obispo
cwhite46@calpoly.edu

This Version: July 21, 2017

Abstract

Vaccination represents a canonical example of externalities in economics, yet there are few estimates of their magnitudes. I provide evidence on the social and externality benefits of influenza vaccination in two settings. First, using pre-existing differences in state-level vaccination rates interacted with exogenous variation in vaccine quality, I provide causal estimates of the impacts of aggregate vaccination rates on mortality and work absences in the United States. Scaled nationally, I find that a one percentage point increase in the vaccination rate results in 985 fewer deaths and 7.5 million fewer work hours lost due to illness each year. The mortality effects are concentrated among individuals 75 and older, but 35-85% of the benefits are driven by the vaccination of people under 75, suggesting a considerable externality effect. Second, I examine a setting in which vaccination is targeted at a group with high externality benefits: influenza vaccination mandates for health care workers. I estimate that these mandates lead to a 17% decrease in hospital admissions with an influenza diagnosis. For both the general population and the population of health care workers, I calculate the monetary benefit per vaccination and find that these benefits are large in comparison to the costs of vaccine administration.

JEL Codes: I12, I18, D62, H23

Keywords: Vaccine, Vaccination, Influenza, Flu, Externality, Health

*This paper has benefitted greatly from conversations with Olivier Deschênes, Peter Kuhn, Maya Rossin-Slater, and Heather Royer, and from participants at the Western Economic Association Conference and the UCSB labor lunch seminar series. I thank the UCSB Department of Economics and the Broom Center for Demography for providing funding for this research, and Betty Henderson-Sparks for facilitating access to restricted data through California's Office of Statewide Health Planning and Development. All errors are my own.

1 Introduction

According to the Centers for Disease Control and Prevention (CDC), between 5% and 20% of the U.S. population are infected with influenza each year; these infections result in an average of approximately 200,000 hospitalizations and over 20,000 deaths.[1] Influenza is considered to be a vaccine-preventable disease, yet vaccination rates for influenza are substantially lower than vaccination rates for other vaccine-preventable diseases. This is largely due to the fact that the vaccine has to be received annually (and thus the cost of vaccination is relatively high) and due to the lack of public policy incentivizing vaccination.

Vaccination serves as a canonical example of positive externalities in economics. Those who receive the vaccine incur some cost (monetary or otherwise) and experience a *private* benefit through the reduced risk of becoming ill; the *externality* benefit comes through the reduced risk of spreading the disease to others and the *social* benefit is the sum of the two. Because the benefits of vaccination are not fully internalized by the recipient, vaccines will be under-utilized relative to the social optimum in the absence of policy. This feature of vaccination has long been recognized by economists, and many theorists have considered how the socially optimal level of vaccination can be reached.[2] Achieving a social optimum requires information on both the marginal cost and the marginal social benefit of vaccination. While the private benefits of vaccination can be measured to some extent through the use of randomized controlled trials (RCTs), estimating the full extent of the social benefits requires an analysis at the population level.

This paper measures the marginal social benefit of influenza vaccination in two settings. First, I estimate the effects of state-level vaccination rates on influenza-related mortality and work absences in the United States. This portion of the analysis addresses the social benefits of vaccination in the general population. Second, I consider the efficiency gains to be had through *targeted* vaccination by examining a situation in which the externality benefits of vaccination are likely to be especially important. I analyze the impacts of influenza vaccination mandates that apply to health care workers (HCWs) in California.

I measure the causal impacts of state-level vaccination rates by interacting pre-existing state-level differences in vaccination rates with year-to-year variation in the effectiveness of the vaccine. Vaccine effectiveness is measured as the extent to which the strains included in the season's vaccine match the strains that end up circulating. Mis-matches occur because of unpredictable genetic changes in the virus, and the prominence of these mis-matched viruses is not known until after vaccines have been distributed. Mis-matches provide an exogenous

[1] Source: http://www.cdc.gov/flu/about/qa/disease.htm.

[2] For example: Stiglitz (1988); Brito et al. (1991); Francis (1997); Geoffard and Philipson (1997); Francis (2004); Boulier et al. (2007); Althouse et al. (2010); Manski (2010, 2016).

source of variation in *effective* vaccination while allowing for *actual* vaccination rates to be held constant.

I find that increases in influenza vaccination rates lead to significant reductions in influenza-related mortality. Scaled nationally, I find that a one percentage point increase in the U.S. vaccination rate would result in approximately 985 fewer deaths per year. The mortality benefits primarily accrue to individuals 75 and older, though 35-85% of this benefit is driven by the vaccination of people under 75, suggesting that externalities play an important role. I also find that influenza vaccination is associated with reductions in the probability of illness-related work absences. The estimates indicate that a one percentage point increase in the U.S. vaccination rate would result in approximately 7.5 million fewer work hours lost due to illness annually. I find no impacts on either outcome during periods in which there is no influenza circulating and no impacts on outcomes that are implausibly related to influenza during any period.

I translate these impacts into monetary estimates of the marginal social benefits of vaccination. Using an age-adjusted value of statistical life (VSL), I estimate that each vaccination confers between $67 and $254 in benefits, depending on the VSL figure considered, due to reduced mortality among individuals 75 and older. This benefit grows considerably if younger individuals are considered as well, though the estimates are only statistically different from zero at conventional levels for the elderly. On the dimension of illness absences, I find that each vaccination confers benefits of approximately $49.

Because the first component of the analysis exploits natural variation in vaccination rates, the estimates can be interpreted as the impacts of increasing vaccination among individuals who are on the margin of the decision to vaccinate. The social benefits of vaccination are likely to be heterogeneous depending on who is vaccinated, and it is not necessarily those who are on the margin whose potential externality benefit is largest. In the second component of the analysis, I consider vaccination policy targeted at individuals with a high potential externality benefit by exploiting the roll-out of county-level influenza vaccination mandates that apply to health care workers in California. This setting also provides a distinct advantage in measuring externality impacts as there exists clear link between those who receive the treatment (health care workers) and those who benefit from the externality (their patients). Most of these mandates apply to all licensed health care facilities in a county, and thus there is potential for these mandates to reduce the spread of influenza both within the hospital (the unit of analysis) and in other health care settings (e.g., primary care offices). The main finding is a reduction in admissions with an influenza diagnosis by approximately 17%. This effect is largest for children (especially infants), and is driven by both infections that were acquired prior to admission and during the hospital stay. Using only the reduction

in new admissions (i.e., infections acquired outside of the hospital), I estimate the marginal benefit of vaccination in terms of health care costs to be approximately $172 per vaccination.

For both vaccination of the general population and of health care workers, the estimated marginal benefits of vaccination are large in comparison to the cost of vaccine administration, suggesting that programs seeking to increase vaccination in either population are likely to be cost-effective.

The primary contribution of this paper is to provide causal estimates of the effects of increasing population-level vaccination rates. While a large medical literature evaluates the benefits of influenza vaccination, much of existing evidence on these benefits is derived from RCTs in which vaccination is randomized within a given group and outcomes are compared across individuals assigned to the treatment or control group.[3] Such a design cannot capture externality effects, and estimates of the direct effects may be attenuated if members of the control group benefit from the vaccination of others. There are a limited number of studies in the RCT literature that directly evaluate externality effects by randomizing across groups rather than individuals (i.e., cluster RCTs). For example, Loeb et al. (2010) employ such a design, randomizing across isolated communities in Canada. In their study, influenza vaccinations were provided to children in the treatment communities and placebo vaccinations were provided to children in control communities. The authors find that vaccinating children led to reductions in laboratory-confirmed influenza for both children and adults in the treated communities, providing evidence of an externality benefit.

While it is possible to identify the presence of externalities in the context of an RCT, it is exceedingly difficult to identify the benefits of vaccination on severe outcomes such as mortality. The relative infrequency of the outcome would necessitate an extremely large-scale study; furthermore, ethical concerns over providing placebo vaccinations to high risk groups essentially relegates the study of *any* benefits (i.e., not only mortality) of influenza vaccination in the elderly population to an observational setting. The potential for bias in existing observational studies is large: a review of the evidence on vaccination in the elderly population noted implausibly large effects of vaccination on all-cause mortality, explaining that these results were likely due to, "systematic differences between the intervention and control arms" (Jefferson et al., 2010).

To my knowledge, there are few examples of papers that effectively circumvent this endogeneity issue. Ward (2014) uses exogenous variation in vaccine effectiveness to evaluate the impacts of a regional influenza vaccination campaign in Ontario, Canada. The author

[3]Reviews of this evidence are available from several sources, including the annual Recommendations of the Advisory Committee on Immunization Practices provided by the CDC (Grohskopf et al., 2014), a number of Cochrane reviews (Jefferson et al., 2010, 2012; Demicheli et al., 2014), and others (Osterholm et al., 2012).

finds that the program increased vaccination rates for non-elderly adults by approximately 10.8 percentage points (the post-treatment vaccination rates was approximately 33.3%) and resulted in a near elimination of influenza infection, a 92% reduction. The results suggest that Ontario reached a threshold level of vaccination beyond which the marginal benefits of vaccination fall to near zero. Models of influenza dynamics suggest the existence of such a threshold (Boulier et al., 2007). That being said, the fact that an annual epidemic is still experienced each year in the U.S. despite vaccination rates well above those during the study period in Ontario suggests that such a threshold has not been reached in the U.S. and that the results of the program in Ontario may have been specific to the location or period of analysis.

Similar to Ward (2014), my identification strategy relies on exogenous year-to-year variation in vaccine effectiveness. My strategy, however, has the advantage of exploiting variation in vaccination rates and outcomes across 50 states and 21 influenza seasons. As such, the average impacts that I estimate are not strongly influenced by the experience in any one location or period. The estimates presented here are somewhat smaller in magnitude, though they still suggest large benefits of vaccination: my estimates suggest that a similar increase in the influenza vaccination rate (10.8 percentage points) would decrease influenza mortality by up to half.[4]

This paper also provides the first large-scale evidence on the impacts of influenza vaccination mandates for health care workers. This is an important contribution as such policies are actively being considered by regional public health departments. This is underscored by editorial articles published in several prominent medical journals that call for the adoption of such requirements (Stewart, 2009; Caplan, 2011; Hooper et al., 2014). The existing evidence on the benefits of such policies is derived from a relatively small number of studies that assess the impacts of vaccination requirements primarily for employees of long-term care facilities. A recent meta-analysis rates the overall quality of evidence on the subject as either "low" or "very low" (Thomas et al., 2016). Furthermore, these studies have limited scope for identifying the impacts of HCW vaccination on low-probability outcomes, and any impacts in settings other than a long-term care facility. The latter point is especially important in light of the findings presented here, indicating that the largest benefits of these policies accrue to infants.

Finally, this paper contributes more generally to a literature in economics that seeks to empirically identify externality impacts in a variety of settings. This literature is especially

[4]This is inferred from estimates of average annual influenza mortality; if lower-bound figures (near 20,000) are used then the estimates suggest a reduction in mortality by approximately one half. If upper-bound figures are used (near 50,000), then the estimates suggest a reduction in mortality by less than a quarter.

prominent in environmental economics, where many papers have sought to measure the impacts of pollution (broadly defined) on a variety of outcomes.[5] Other examples include the evaluation of externality impacts of de-worming programs on health and schooling outcomes (Miguel and Kremer, 2004) and the estimation of displacement effects in job placement programs (Crépon et al., 2013). Notably, there are few papers to my knowledge that seek to empirically identify externality or social impacts of vaccination, despite the fact that vaccines are often regarded as the "textbook" example of a positive externality (Stiglitz, 1988).

The remainder of the paper is structured as follows. Section 2 provides background information on influenza and influenza vaccination, as well as a conceptual discussion that is helpful for interpreting the results of the empirical analysis to follow. Section 3 describes the data, empirical strategy and the results for the analysis of aggregate vaccination rates. Section 4 describes the data, institutional setting, empirical strategy, and results for the analysis of HCW mandates in California. Finally, Section 5 offers a discussion of the estimates and concludes.

2 Background

In this section, I provide a brief overview of several points regarding influenza and influenza vaccination that are necessary for interpreting the results of the empirical analysis. I also provide a conceptual discussion of the benefits of vaccination, focusing on the theoretical shape of the marginal social benefit curve in the specific case of influenza vaccination.

2.1 Influenza and Influenza Vaccination

There are three key points regarding influenza for which I provide an overview in this section. First, I discuss the burden of influenza; specifically, it is important to understand the specific ways in which different groups are affected by the disease. Second, I discuss influenza vaccination, summarizing the current state of knowledge regarding vaccine efficacy. Third, I discuss in more detail the importance of vaccine match, as an understanding of the causes and consequences of vaccine mis-match are key to understanding the identification strategy used in the analysis to follow.

The total burden of influenza illness is large and crosses all demographic groups, though there is substantial heterogeneity in how groups are affected. For the sake of this analysis, I focus on age as the primary dimension of heterogeneity. This discussion reflects the findings

[5] These papers study the consequences of a variety of pollutants including regional air pollution, water pollution, the potential impacts of climate change (see Graff Zivin and Neidell (2013) for a review).

of the CDC's Recommendation of the Advisory Committee on Immunization Practices, which summarizes the general findings from an extensive list of references (Grohskopf et al., 2014).

For children, influenza is responsible for large number of outpatient visits and hospitalizations, and this is especially true for infants (children under one). Neuzil et al. (2000) find that influenza was responsible for an annual average of 6-15 outpatient visits per 100 children under 15. Additionally, Zhou et al. (2012) estimate annual influenza-related hospitalization rates (per 100,000) equal to 151 for infants, 38.8 for children aged 1-4, and 16.6 for individuals 5-49. While outpatient visits and hospitalization are fairly common, death attributable to influenza among children is relatively rare. For non-elderly adults, influenza infection is typically less severe and less likely to result in hospitalization or death. While severe outcomes are less likely, the burden of influenza is still significant, often resulting in outpatient visits and worker absenteeism (Molinari et al., 2007). Influenza infection in elderly adults is the most severe. The majority of deaths related to influenza occur in individuals at least 65 years old. The CDC estimates that between 1976-2007, average annual deaths attributable to influenza were 21,098 for individuals 65 and older, 2,385 for individuals 19-64, and 124 for individuals under 19 (Thompson et al., 2010). As such, these estimates indicate that elderly individuals account for approximately 90% of all influenza-related deaths. It should be noted that due to difficulties in reporting and diagnosis, there is no consensus as to the number of deaths that are caused by influenza in each year. Other evidence suggests that the true number could be much larger; Dushoff et al. (2006), for instance, estimate annual average deaths equal to 41,400 for the period 1979-2001.

Influenza vaccination effectiveness – the extent to which vaccination protects against laboratory-confirmed influenza – is determined by several factors. Vaccine match is an especially important factor and will be discussed more fully below, but it is important to note that even when the vaccine is perfectly matched it is not 100% effective. Vaccine effectiveness also varies with age; diminished immune response in elderly individuals means that they are less able to create the antibodies needed to gain immunity. Estimates of vaccine effectiveness in the prime-age population vary, though several studies find values in the range of 50-60% in a well-matched season (Demicheli et al., 2014; Grohskopf et al., 2014). Estimates of vaccine effectiveness in the elderly population are more contentious; this is primarily due to the fact that ethical concerns over providing placebo vaccinations to high-risk populations limit the ability of researchers to use RCTs. There is some debate as to whether the vaccine provides *any* protective benefits among the elderly (Simonsen et al., 2007), though a recent study reported by the CDC indicated effectiveness of approximately 26% among people 65 and older (McLean et al., 2014) during a well-matched season.

Vaccine match – the degree to which the strains included in the vaccine match the strains

that end up circulating – is an especially important determinant of vaccine effectiveness; studies of vaccine effectiveness find that the vaccine is much less effective when at least one of the dominant circulating strains is not included in the vaccine (Jefferson et al., 2010). Understanding the identification strategy in the main analysis requires understanding the process by which a vaccine mis-match occurs. For the North American vaccine, this process begins in early Spring, when the World Health Organization (WHO) convenes a meeting in order to make recommendations regarding the composition of the following season's influenza vaccine. The vaccine includes three (trivalent) or four (quadrivalent) strains, and the decision as to which strains to include in the vaccine is primarily based on which strains were circulating most recently.[6] The Food and Drug Administration makes the ultimate decision regarding vaccine composition in the U.S., and vaccine composition is common across all states. Due to the time it takes to produce and distribute the vaccine, this decision must be made in early Spring so that vaccines can be administered in the Fall. The influenza virus itself undergoes constant genetic change ("antigenic drift") such that there are always viruses in existence that are genetically distinct from the dominant strains; vaccines may not provide protection against these genetically distinct viruses. Significant mis-matches occur when one or more of these genetically distinct viruses spreads and becomes one of the dominant strains in a given season. Such mis-matches are essentially impossible to predict prior to the start of influenza season. It is also important to note that not only is the vaccine formulated well before influenza season begins, but individuals typically have no information regarding vaccine match at the time of vaccination. I provide direct evidence of this in Section 3.2.

2.2 Marginal Social Benefits of Vaccination

The goal of this paper is to estimate the marginal social benefits of influenza vaccination. It is useful to consider this in the context of a simple economic framework of externalities. In this framework, there is a marginal private benefit of vaccination (MPB) and a marginal social benefit of vaccination (MSB); these benefits are assumed to be decreasing in the vaccination rate and the MSB is at least as large as the MPB at all points (i.e., the externality is non-negative). In a competitive equilibrium, consumers purchase vaccines such that the MPB equals the marginal private cost (MPC), and the vaccine will be under-provided relative to a social optimum. The economic intuition is straightforward and is the basis for the analysis conducted in this paper. Considering the shape of the benefits curve in the specific context of influenza vaccination provides additional insight.

[6]The quadrivalent vaccine was introduced in 2012. This does not present a problem for the analysis as the fourth strain has not circulated widely in the years under study.

Boulier et al. (2007) combine this basic externality theory with a workhorse model of disease dynamics (the susceptible-infected-removed, or "SIR", model) and parameterize the model to the case of influenza in order to derive theoretical predictions for the shape of the marginal benefits curves. Figure 1 reproduces their results, allowing the MPB and MSB to depend on vaccine effectiveness. I have plotted each assuming either 100% or 50% effectiveness (denoted E). Recall that even when the vaccine is well-matched, estimates of vaccine effectiveness are typically in the range of 50-60%. Consider first the case of a perfectly effective vaccine. The gap between the MPB and MSB represents the marginal externality effect; note that the y-axis measures the number of infections such that at a vaccination rate of zero, the model predicts that an additional vaccination will prevent more than 1.5 infections in expectation; 0.5 infections are prevented in private benefits and the remainder are prevented in externality benefits. Measuring infections is equivalent to measuring the cost of disease if it assumed that the cost of infection is homogeneous and equal to one.

As the vaccination rate increases, the MPB decreases but the MSB stays relatively flat (or increases) until a threshold is reached. This threshold represents the point at which a seasonal epidemic fails to emerge. The shape of these curves prior to the threshold is important as they imply that neither the externality nor the social benefit of vaccination decreases prior to this point. Furthermore, it can be inferred that the U.S. is not beyond the threshold, as an influenza epidemic *does* emerge in each season. Current vaccination rates (approximately 43% in 2014) in combination with the persistence of an annual epidemic is at odds with the model that assumes $E = 100\%$ and predicts a threshold level of vaccination between 30% and 40%. If we consider a lower E, the benefits of vaccination fall and the threshold increases. At a more realistic $E = 50\%$, the vaccination threshold beyond which there would be limited marginal benefits of vaccination is over 60%.

I caution that this model depends on a number of parameter choices that are difficult to estimate accurately.[7] That being said, considering the general shape implied by the model helps to guide the interpretation of the results presented in this paper. The model implies relatively constant marginal social benefits of vaccination below a threshold. The implication is that estimates of the social benefits are unlikely to depend strongly on the level of vaccination. In other words, we should expect that the relationship between vaccination rates and the outcome of interest is roughly linear. This applies only until the threshold is reached, though reaching the threshold would be obvious as a seasonal epidemic would fail to form and very few infections would occur.

[7]Important parameters include the vaccine effectiveness (as shown) and the "contact number", which is the number of additional infections that result from a single infection when the entire population is susceptible.

Discussion of this framework also presents the opportunity to discuss potential heterogeneous impacts of vaccination. It is worth considering how three groups in particular may differ from the remainder of the population: infants, the elderly, and health care workers. For infants, the cost of infection is relatively high as infection is often serious enough to result in hospitalization; furthermore, infants under 6 months of age cannot receive the vaccine. For these reasons, infants are a group who benefit substantially from the vaccination of others. A similar intuition holds for the elderly: the cost of infection is high and vaccine effectiveness is relatively low. These factors combined imply that the elderly benefit substantially from the vaccination of others. Health care workers are a group who come in relatively frequent contact with both infected individuals and people with a high cost of illness; as such, the external benefits of health care worker vaccination are likely to be large. This heterogeneity motivates the focus on these subgroups in the empirical analysis to follow.

3 Part I: Aggregate Vaccination Rates

3.1 Data

The analysis of aggregate vaccination rates requires data on mortality by cause of death, illness-related work absences, influenza vaccination rates, information on the timing and magnitude of influenza activity, and the degree to which each year's vaccine matches the circulating strains. The unit of analysis is the state-year-month and the data cover the years 1993-2013.[8]

3.1.1 Mortality Data

Mortality data are derived from the multiple cause of death files from the National Vital Statistics System (NVSS). This is the restricted version of this data that include state identifiers beyond 2005. It is important to note the use of multiple causes of death in classifying mortality as influenza-related rather than the single underlying cause of death. Dushoff et al. (2006) find that a large number of influenza-related deaths are excluded when only the underlying cause of death is used. Accordingly, I classify deaths by diagnosis if any of the (up to 21) diagnosis codes fall into the relevant category. Even using multiple causes of death, it is very rare for a death to be classified as specifically due to influenza. As such, the category with the highest level of specificity that I use in the analysis of mortality, and the primary outcome of interest, is deaths related to pneumonia and influenza (PI). Because

[8]The data source used to measure vaccination rates begins in 1993.

deaths due to influenza often occur as a result of complications or the exacerbation of pre-existing conditions, even PI deaths may exclude deaths that occurred as a result of influenza infection. As such, I also analyze deaths in two higher levels of aggregation: deaths with any respiratory or circulatory diagnosis, and all-cause deaths. Because it is highly unlikely that deaths without a respiratory or circulatory diagnosis occurred as a result of influenza infection, I use these non-respiratory/circulatory deaths as a falsification test.[9]

3.1.2 Labor Market Data

Data on illness absences and hours worked are derived from the Current Population Survey (CPS) basic monthly files. I follow Stearns and White (2016) in constructing the measure of illness absence; while more details can be found in Stearns and White (2016), I describe the important points here. The measure of illness absences is constructed using two questions posed to all individuals who report being employed. First, individuals who report being employed but absent from work for the entire reference week (i.e., worked zero hours) are asked the main reason for their absence from work. Second, individuals who are employed and at work during the reference week report both their usual hours worked per week and the number of hours actually worked in the reference week. Those who work less than 35 hours during the reference week but report that they usually work at least 35 hours per week are asked the main reason for working less than normal. Typically, these are full-time workers who took one or more days off in the reference week, yet worked a non-zero number of hours (i.e., were not absent the entire week).

Each of these two questions lists "own illness" as one possible reason for missing work and is the reason given for approximately 19% of absences (for both entire-week and partial-week absences). Leave-taking for own illness is the primary labor market outcome of interest, though I also assess whether leave-taking for any other reason is affected as a falsification test. It is important to note that these measures of absence can only be constructed for individuals who work at least 35 hours per week. This is not true of data on hours worked, for which the full sample of employed workers is used. In the labor market analysis, I also use a standard set of controls including indicators for gender, age (<20, 20-30, 30-40, 40-50, 50-60, >60), marital status (married, widowed/divorced/separated, never married), and education (less than high school, high school diploma, some college, college graduate).

[9]The ICD9 and ICD10 codes used to classify these diagnoses are as follows: Influenza (ICD9: 487, ICD10: J9-J11), Influenza/Pneumonia (ICD9: 480-488, ICD10: J9-J18), Respiratory/Circulatory (ICD9: 390-519, ICD10: I00-I99, J00-J99).

3.1.3 Vaccination Rates

Data on state-level vaccination rates are obtained through the Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS is a large-scale telephone survey that has been conducted at the national level since 1993.[10] The BRFSS asks whether each participant has received an influenza vaccination within the past 12 months.[11] The phrasing of this question is not ideal for determining vaccination rates in each influenza season as there are several months wherein the season to which the vaccine applies is ambiguous. As such, I drop all survey responses taken from September through December, as these are the months in which the vast majority of influenza vaccinations are received. Vaccination rates for the 2012-13 influenza season, for example, are calculated as the percentage of respondents surveyed between January and August of 2013 who report having received an influenza vaccination within the past 12 months. Information on vaccination rates was not collected in the survey years 1994, 1996, 1998 or 2000. To fill in these gaps, I linearly interpolate vaccination rates in these years for each state. Because vaccination rates evolve relatively smoothly over time, and because the identification strategy used in this analysis relies not on year-to-year changes in vaccination rates but on baseline differences in the level of vaccination across states, such a procedure is not a concern for identification.[12]

3.1.4 Match Rates

Data on vaccine effectiveness are derived from annual influenza season summaries, which consist of data compiled from the CDC's virologic surveillance system.[13] This system consists of approximately 110 World Health Organization (WHO) laboratories and 240 National Respiratory and Enteric Virus Surveillance System (NREVSS) laboratories located throughout the country. These laboratories test respiratory specimens for the presence of any influenza virus and characterize viruses according to the exact strain. The annual summary data contain information on the number of viruses by strain and information indicating which strains the season's vaccine protects against. The match rate for each season is defined as the percentage of characterized viruses that match the strains contained in that season's vaccine. It is possible that the vaccine can offer some level of protection against strains that are not perfectly matched if the strain in the vaccine and the strain circulating are similar,

[10] The survey began in 1984, though it was conducted in a limited number of states.

[11] The exact phrasing of this question varies slightly from year-to-year. In more recent years, for instance, the survey asks about various types of vaccination (i.e., flu shots or spray). I classify each individual as having received an influenza vaccination if they received at least one dose of any type of influenza vaccination.

[12] Table A1 demonstrates that the main results are not sensitive to excluding the years in which interpolated vaccination rates were used.

[13] These data are available at: http://www.cdc.gov/flu/weekly/pastreports.htm.

and this information is indicated in the data. I construct two versions of the match rate, one in which strains are characterized as matched only if it is the exact strain contained in the vaccine, and one in which strains are characterized as matched if the vaccine offers some level of protection. The main specification uses the average of these two measures.[14]

3.1.5 Influenza Activity

Data on the timing and magnitude of influenza activity are also obtained from the CDC's virologic surveillance system. The primary measure of influenza activity is the percentage of tests that are positive for any type of influenza. I re-scale this variable to range between zero and one within the sample by dividing by the maximum observed value such that the measure can be interpreted as an index with a value of one representing severe (but observed) influenza activity. This data is used in both the analysis of aggregate vaccination rates and and the analysis of health care workers. For the analysis of aggregate vaccination rates, I use data from all U.S. laboratories, and for the analysis that is specific to California, I use data from laboratories in the western region of the U.S. (the finest level of geographic specificity).

3.2 Empirical Framework

Estimating the impacts of population-level vaccination rates is an empirically difficult task. To illustrate this difficulty, consider the following empirical equation:

$$Y_{smy} = \beta V_{sy} + \delta_{my} + \varepsilon_{smy} \tag{1}$$

Y_{smy} is the mortality rate in state s in flu-year y and month m. Note that I define "flu-years" as years running from July through June so that each flu-year represents a distinct influenza season. V_{sy} is the vaccination rate for state s and flu-year y. The regressor of interest in this equation (V_{sy}) is endogenous. It is certainly plausible, for example, that individuals in states that are more affected by influenza are more likely to receive an influenza vaccination. Figure 2 maps vaccination rates along with PI mortality rates in the U.S. The endogeneity of vaccination is apparent, as there appears to be a positive correlation between

[14]The exact process by which I calculate the match rate is slightly more complicated than I have laid out above. Each positive test received by the CDC is classified as either influenza A or influenza B. A subset of influenza A viruses are sub-typed (H1N1 or H2N3) and a subset of each subtype are then characterized to determine the exact strain. A subset of influenza B viruses are characterized to determine the exact strain. The annual summaries contain information on the number of total tests, the number of positive tests, the number of A and B viruses, the number of viruses sub-typed, the number of each subtype characterized, and the number of viruses belonging to a specific strain. Though relatively straightforward, I have developed a calculator that takes these numbers as inputs and outputs the match rate for each season. This calculator is available upon request.

vaccination rates and PI mortality. My strategy is to interact differences in vaccination rates with year-to-year variation in vaccine effectiveness to generate plausibly exogenous variation in the *effective* vaccination rate.

$$Y_{smy} = \gamma_1(V_{sy} * M_y) + \gamma_2 V_{sy} + \delta_{my} + \varepsilon_{smy} \tag{2}$$

Equation (2) describes this difference-in-differences approach. In this equation, M_y is the match-rate, measured nationally for each influenza season. A match rate of zero implies that the vaccine is minimally effective whereas a match rate of one implies maximum effectiveness.[15] γ_2 represents the potentially endogenous component of the relationship between vaccination rates and the outcome of interest: this measures the relationship between vaccination rates and and the outcome in seasons in which the vaccine is minimally effective (i.e., zero match rate). γ_1 is the object of interest, and represents the differential effect of an increase in the vaccination rate between years when the vaccine is at maximum versus minimum effectiveness. Intuitively, γ_1 picks up the impact of *effective* vaccination (i.e., the causal effect of vaccination), but not the component of the relationship between vaccination rates and the outcome that persists in seasons when the vaccine is ineffective.

Identification relies on the assumption that the match rate is exogenous from year to year and that unobserved factors that are correlated with vaccination rates are unrelated to the match rate. Of potential concern is that there may be differential responses to match rates in terms of vaccination behavior across states with different vaccination rates. The process by which the strains are chosen for inclusion in the following season's vaccine formulation, described in Section 2.1, supports the notion that the match rate is effectively random from year to year. Further, I provide evidence that there is limited scope for individuals to respond to vaccine effectiveness in terms of vaccination behavior.

From 2007 onward, the BRFSS has asked respondents not only whether they received a vaccination, but the month in which they were vaccinated. Figure 3 plots the average (across years) cumulative vaccination rate by month for the years in which this data is available. Additionally, this figure displays average influenza intensity by month. Together, this figure shows that in a typical influenza season, nearly all vaccinations are distributed before the onset of the season's influenza outbreak. Because information on the vaccine's match cannot be determined until a significant number of individuals are infected, this plot suggests that there is very limited scope for responding to the match rate at all, much less

[15] A zero match rate does not necessarily imply that the vaccine is completely ineffective as vaccination can provide some level of protection against non-matched strains (especially if the strain is similar to that included in the vaccine). Similarly, maximum effectiveness does not imply that the vaccine is perfectly effective. In fact, maximum efficacy is approximately 60% effectiveness in a typical year.

differentially across states.[16] Table 2 provides additional evidence on this point. Column one reports estimates of a regression of the vaccination rates on match rates. The estimate, while statistically different from zero, is *negative* and close to zero. The estimate implies that the vaccination rates are 0.5 percentage points lower in a 100% match season compared to a 0% match season. If individuals were truly responding to match rates, it is likely that vaccination rates would be *higher* rather than lower in high match seasons. Column two presents estimates of a test for differential responses to vaccine effectiveness by interacting the match rate with mean vaccination rate for each state; the results suggest that there is no differential response among states that tend to have higher or lower vaccination rates.

Influenza vaccination rates will only have a causal effect on mortality during periods in which influenza is circulating. As a third source of variation, I exploit within-year variation in the timing and magnitude of influenza activity. I exploit this variation in two ways. First, I estimate Equation (2) in periods of relatively high and low influenza activity. The coefficient of interest, γ_1, is expected to be relatively close to zero in periods of low influenza activity. Second, I conduct this exercise more formally in a triple-difference approach.

$$\begin{aligned} Y_{smy} = \phi_1(V_{sy} * M_y * A_{my}) + \phi_2(V_{sy} * M_y) \\ + \phi_3(V_{sy} * A_{my}) + \phi_4 V_{sy} + \delta_{my} + \varepsilon_{smy} \end{aligned} \tag{3}$$

A_{my} is a measure of influenza activity common to all states; this is an index that is scaled to equal one during the month in the sample with the maximum influenza activity and zero when there is no influenza activity. As such, ϕ_1 measures the difference in γ_1, from Equation (2), between periods of maximum activity and periods of zero activity. The advantage of this approach is that the linear measure of influenza activity provides a more accurate measure relative to the somewhat arbitrary classification of months into "high" and "low" activity. Estimates of ϕ_1 are my preferred estimates in this analysis.

I have not to this point discussed the inclusion of additional fixed effects (e.g., state fixed effects) as the identification strategy does not rely on their inclusion. The inclusion of state-specific effects changes the intuition as to what variation is being exploited in estimating these models; specifically, the inclusion of state fixed effects implies that the estimates rely on within-state differences over time in the vaccination rate rather than cross-state differences.

[16] This plot illustrate another important point as well. For most years, there does not exist data on the timing of vaccination and as such the main specification defines vaccination rates at the state-by-season level. Figure 3 provides support for the idea that this is a reasonable strategy as vaccinations are typically received before the onset of any influenza activity. In other words, I am not measuring the effects of vaccines received in February on influenza-related mortality in January.

In either case, the interaction with random year-to-year variation in vaccine match provides a plausibly exogenous variation in effective vaccination. I explore a range of fixed-effects specifications, though my preferred specification includes the year-by-month effects described above as well as state-by-month effects that not only account for state-specific factors, but allow seasonality in the outcome to vary by state. Note that I do not include any state-level time varying characteristics; instead, I recognize that Equation (3) allows for the inclusion of state-by-year fixed effects and demonstrate that the results are not sensitive their inclusion.

Finally, to assuage any remaining concerns that the results are driven by changes in vaccination behavior in response to current conditions, I estimate an additional specification wherein only *pre-existing* variation in vaccination rates is employed. More specifically, I re-estimate Equation (3) where the vaccination rate and each of its interactions are instrumented using the vaccination rate from three seasons prior and the corresponding interactions.

3.3 Results

In interpreting the estimates, note that the three regressors (*Vaccination*, *Match*, and *Activity*) are all linear measures. *Vaccination* is measured on a scale of zero to 100 such that the estimates can be interpreted as a one percentage point increase in the vaccination rate, while the other two regressors are measured on a scale of zero to one. This section primarily focuses on mortality as the outcome, which is measured as the PI mortality rate per 100,000 individuals.

3.3.1 Main Results

Table 3 provides estimates of the difference-in-differences equation described in Equation (2). Similar to the following tables, the columns represent estimates from models specified in different ways. The first includes only month-by-year fixed effects, the second adds state-by-month fixed effects, and the third additionally uses the vaccination rate from the prior season to instrument for the current vaccination rate. In Table 3, Panel A presents estimates that include all months, where Panel B restricts the sample to months in which the influenza activity index is greater than 0.5 (relatively high activity), and Panel C restricts the sample to months in which the activity index is less than 0.5 (relatively low activity).

Let us first consider Panel A. In these difference-in-differences estimates, the coefficient on "Vacc" represents the relationship between vaccination rates and PI mortality when the match rate is zero. The first column indicates that the estimate of γ_2 is positive and significant when no state-by-month fixed effects are included in the model. Taken at face

value, this implies that states with higher vaccination rates experience *more* PI deaths. This illustrates the positive correlation suggested by Figure 2.

The coefficient on the interaction term is the estimate of interest, as this represents the differential effect of an increase in the vaccination rate between years in which the vaccine is perfectly matched and years in which the match rate is zero. In the model without state fixed effects, the estimate of the interaction is not statistically different from zero. The estimates are negative and statistically different from zero in columns two and three. The estimates from Panel A include many months in which there is no influenza activity and thus the estimates are likely to be dampened by their inclusion. The estimates of the interaction in Panel B, in which the sample is limited to months with relatively high influenza activity are negative, significant, and similar in magnitude across all three models. The interpretation of the coefficient reported in column one (-0.135) is that a one percentage point increase in the vaccination rate leads to a decrease in influenza mortality of 0.135 per 100,000 residents during seasons in which the vaccine is perfectly matched and months with relatively high influenza activity. Note that I will provide a more intuitive interpretation of the results in discussion of the triple difference model to come.

The coefficient estimates on the un-interacted term in Panel B are worth noting as well. These estimates are somewhat larger than the estimates for all months. The implication is that it is not only people in states with higher PI mortality in general that are more likely to be vaccinated against influenza, but people in states with higher mortality that is specifically related to influenza. Finally, Panel C reports estimates for months with relatively low influenza activity. The estimates indicate that influenza vaccination has little impact on PI mortality in low activity months. Note that the estimates of the interaction terms in models with state fixed effects are statistically different from zero, though quite small in comparison to the estimates for months with high activity. This negative estimate is not unexpected as low activity months are simply defined as months with *relatively* little influenza activity rather than zero activity.

The estimates reported in Table 4 formalize the comparison between months with high and low influenza activity in a triple difference context. The triple interaction is the coefficient of interest in this model. The estimated coefficients are somewhat larger than the coefficients estimated for high activity months in the diff-in-diff specification. This is unsurprising given that these estimates represent a comparison between months with maximum influenza activity and zero influenza activity rather than an estimate representing the differential effect between relatively high and low activity months. A point estimate of -0.196 implies that that a one percentage point increase in the influenza vaccination rate will decrease the PI mortality rate by 0.196 per 100,000 individuals during months with maximum

influenza activity compared to months with no influenza activity, and during seasons in which the vaccine is perfectly matched.[17]

Because the interpretation of this estimate is not the most transparent, I also report an estimate of the expected annual benefit of vaccination. To calculate the expected annual benefit, I first create expected monthly benefits and then sum across months. The expected monthly benefit is calculated as the product of the coefficient estimate, mean month-specific influenza activity, and the mean match rate. The estimates in the second column indicate that the expected annual mortality reduction due to a one percentage point increase in the vaccination rate is approximately 0.309 per 100,000 residents. Scaling this to the U.S. population, this implies that a one percentage point increase in the national vaccination rate would result in approximately 985 fewer deaths due to influenza. This is a significant, albeit reasonable number given that estimates of average annual deaths due to influenza lie in the range of approximately 20,000-40,000.

Before moving on, I first explore the extent to which PI deaths capture the total benefits of influenza vaccination in terms of mortality. In Table 5, I re-estimate the triple difference specification in models where the outcome is constructed using a broader definition. The two broader definitions are deaths with any respiratory or circulatory diagnosis and all-cause deaths. The estimates indicate that the PI categorization captures the majority of the benefits realized by influenza vaccination, though the point estimates do grow slightly (for example, from -0.196 using only PI deaths to -0.242 using all respiratory and circulatory in one specification). While these estimates are slightly larger, the confidence intervals are larger as well; the precision afforded by the more specific PI definition is preferred for the remainder of the analysis. Table 5 also reports estimates from an important falsification test: defining the outcome as deaths that should not plausibly be related to influenza infection (deaths without any respiratory or cardiovascular diagnosis). The estimates indicate small and insignificant effects of influenza vaccination on these deaths.

3.3.2 Age Heterogeneity

A natural question to ask at this point is who reaps the benefit of reduced influenza mortality. I decompose the main effect by age in Table 6, where I report the triple-difference estimate

[17] The other estimates reported in Table 4 are worth noting as well. The second row (the vaccination rate interacted with the match rate), because of the interaction with the exogenous match rate, can be considered the causal effect of increasing influenza vaccination rates on PI mortality during months in which there is no influenza activity. Reassuringly, these estimates are not statistically different from zero. The estimates in the third row (the vaccination rate interacted with influenza activity) again illustrate the positive relationship between influenza-specific mortality and vaccination rates. The estimates in the final row (the vaccination rate un-interacted) essentially represents the relationship between PI mortality that is not influenza-specific and vaccination rates.

for five age groups (infants under 1, 1-10, 10-64, ≥64 and ≥74). In Table 6, I calculate age-specific mortality rates by dividing by the *total* state population rather than the age-specific population. As such, the estimates can be interpreted as an accounting of the total benefits of influenza vaccination – the sum of the impacts for the mutually exclusive age groups equals the total effect. In this table, I report in brackets the percentage of benefits that accrue to each age group. This is calculated as the age-specific coefficient divided by the corresponding all-age coefficient from Table 4. The disadvantage of this approach is that it does not account for the size of the population in each age group; I provide an additional set of estimates in Table A2 that uses the age-specific population in constructing mortality rates.

The age-specific estimates indicate that the vast majority of the mortality benefits of influenza vaccination accrue to the elderly population. This is not a surprising result given that estimates of influenza-related mortality are heavily concentrated among the elderly population. Grohskopf et al. (2014) reports that the ≥ 65 population accounted for 90% of all influenza-related deaths between 1976 and 2007. The estimates presented here accord with these findings: I estimate that between 89% and 93% of the reduction in mortality due to influenza vaccination is experienced among the ≥ 65 population (depending on the model).

While analysis of the ≥ 65 population is useful for comparison with other studies of influenza-related mortality, I find present additional results for an older age group and find that almost all of the benefits are concentrated among the ≥ 75 population. The specification in column 2 indicates that 89% of the total benefits accrue to the ≥ 65 population and 83% to the ≥ 75 population.

3.3.3 Externality Effects

To evaluate the extent to which the mortality benefits of influenza vaccination operate through an externality effect, I use the notion that the vast majority of benefits accrue to individuals who are at least 75 years of age and separately estimate the effects of vaccination rates for individuals who are either within or outside of that age group. More specifically, I estimate the following equation:

$$Y_{smy}^{O75} = \psi(V_{sy}^{O75} * M_y * A_{my}) + \omega(V_{sy}^{U75} * M_y * A_{my}) \\ + \text{Other Interactions \& Fixed Effects} + \varepsilon_{smy} \tag{4}$$

In Equation (4), I include the full set of interactions described in Equation (3) for *both*

people who are under 75 and those who are at least 75. As such, ψ represents a combination of direct and externality effects, where the externality effects are limited to capturing the spread of influenza among people who are in the age 75 and older group. The other coefficient of interest, ω, represents the effect of vaccination among people under 75 on influenza mortality for individuals who are at least 75; this represents a pure externality effect. The results of this exercise are presented in Table 7. The estimates in the first two columns suggest that the majority of the mortality benefits for individuals at least 75 result from the vaccination of people who are under 75. In the IV estimates, the dominant magnitude flips to at least 75 group, though the standard errors grow substantially. The figure reported in brackets indicates that vaccination of those under 75 accounts for between 35% and 85% of the total mortality benefits that accrue to those at least 75. These results suggest a substantial role for externality effects in the mortality benefits of vaccination. Given the relatively low efficacy of influenza vaccinations in older individuals, this is an important though not necessarily surprising finding.

3.3.4 Evaluating Mortality Benefits

The monetary benefits of influenza vaccination in terms of mortality depend on the value of a statistical life (VSL). Because the mortality benefits are concentrated among individuals at least 75 years of age, it is especially important that VSL estimate is age-adjusted. I use the method of Murphy and Topel (2006), who develop a framework for estimating the value of *remaining* life given a standard VSL figure that is evaluated using mortality risk reductions from working-age adults. I apply two such figures: estimates from Ashenfelter and Greenstone (2004) of \$2.3 million (denoted "AG") as a lower bound, and the current standard from the EPA of \$8.8 million as an upper bound.[18] The Murphy and Topel (2006) framework provides VSL estimates for single years of age; I follow the method of Barreca et al. (2016) to calculate a VSL estimate for more aggregated age groups, taking a weighted average of single-year VSL estimates where the weight is the share of deaths from each single-year age group. The single-year VSL figures, derived from Murphy and Topel (2006), as well as the weights used to calculate VSL estimates for age groups, are described in Figure 5. For the benefits calculations, I consider the following age groups: < 10, 10-64, 65-74 and ≥ 75. The VSL estimates for these groups are highly dependent on age. These VSL estimates, along with estimates of the monetary benefits of mortality reduction are presented in Table 8.

For each age group, I consider the annual benefits of a policy that increases the national

[18] Each VSL figure is reported in 2016 dollars.

influenza vaccination rate by one percentage point.[19] I use estimates of age-specific reductions in mortality; these are similar to those presented in Table 6.[20] I then estimate the annual number of deaths avoided, as well as the monetary benefits. Further, I estimate the benefits per vaccination (dividing by 3.19 million, as that is approximately one percent of the U.S. population). While the reductions in mortality are concentrated among individuals at least 75, the monetary benefits are more equally distributed given the relatively low VSL for older individuals. Considering all ages, I find that the benefit of an additional vaccination to be between $134 and $511, depending on whether VSL estimates from the EPA or AG are used. That being said, the only age-specific estimates considered here that are significantly different from zero are the estimates for the at least 75 group; as such, my preferred estimates of the social benefits of vaccination in terms of mortality are limited to this group: $67 using the AG VSL and $254 using the EPA VSL.

3.3.5 Work Absences

Mortality represents the most severe outcome associated with influenza infection; additionally, the lack of quality evidence on the mortality benefits of vaccination makes it a particularly important outcome to examine. While I do provide estimates of the monetary benefits of influenza vaccination on the dimension of mortality, such estimates are often subject to considerable controversy. Since one of the primary purposes of this analysis is to determine the marginal social benefit of vaccination, it would be useful to analyze an outcome for which calculation of the monetary benefits is subject to less debate. In this section, I re-estimate the triple-difference specification (described in Equation (3)) with three different outcomes: the average share of employed individuals absent from work due to illness, average hours worked, and the average share of workers absent from work due to reasons other than illness (as a falsification test).

The results of this exercise are described in Table 9. The main finding is that increasing state-level vaccination rates by one percentage decreases the share of workers absent from work in that month by approximately 0.05 percentage points during maximum influenza months and during a good match year (an approximate 2% effect from the mean). While the estimates for average hours worked are positive (the expected direction), they lack precision, and none of these estimates are statistically different from zero. It is not surprising that a statistically significant effect would not be detected for average hours since this variable

[19] I implicitly assume that the distribution of the increased vaccination rates across ages is proportional to the current age distribution of vaccination rates.

[20] Note that these age groups are slightly distinct from those considered previously; the age groups used here need to be mutually exclusive, and should include the ≥ 75 category as that is where most of the benefits lie. The estimates for the modified age groups are presented in Table A8.

is non-specific relative to illness absences; put another way, a far greater proportion of the variation in this variable comes through factors unrelated to influenza.[21] Reassuringly, I find small and insignificant impacts on absences for reasons other than illness.

Because the estimates for average hours worked lack precision, I require an estimate of the number of hours lost per absence so that I can use these estimates to calculate a monetary benefit of increasing the vaccination rate. Consider the average number of hours worked for three groups: workers who take no absence (43.29), workers reporting any illness absence (18.31), and workers reporting illness absence for only part of the reference week (24.92). The difference in average hours between the no-absence workers and the illness absence workers are my estimates of the number of hours lost per reported absence. I prefer the more conservative definition which excludes entire week absences. The differences are 17.0 hours (partial week illness absences) and 24.9 (all illness absences). These estimates are in line with estimates from research analyzing working days lost due to influenza infection; Keech and Beardsworth (2008) is a meta-analysis of such studies, reporting that the mean number of days lost due to influenza infection ranges from 1.5 to 4.9, depending on the study.

The goal is to convert the estimates in Table 9 into an estimate of the expected annual benefit per vaccination in monetary terms. To start, I multiply the effect on work absences by the number of hours lost per absence to arrive at the number of work hours lost per week, in maximum influenza months and perfectly matched seasons. I then multiply by the number of weeks per month (30.5/7) to convert this to a monthly effect. I then multiply by the same factor that I used to scale the mortality estimates to an expected annual benefit ($\overline{Match} \times \sum_m \overline{Activity}_m$). The resulting figure is 0.06 hours gained per worker, per year in expectation. Scaling this to a national level, I multiply by the number of full time workers in the U.S. (126 million) to arrive at a figure of 7.47 million hours gained; multiplying by the median hourly wage for full-time workers ($20.58) results in expected annual benefits of $156.8 million.[22] Since it would take approximately 3.19 million additional vaccinations (1% of the U.S. population), this implies that the marginal social benefit of vaccination in terms of productive hours gained is approximately $49.2. If I assume that each reported weekly absence results in 24.9 hours lost, this figure is $70.7 per vaccination.

[21] It is worth noting that the magnitude of these estimates is substantially larger than would be implied by the effect on illness absences, yet it is still insignificant.

[22] Median hourly wage for full-time workers is calculated as median usual weekly earnings ($824) divided by median usual hours (40).

3.4 Robustness Checks

In this section, I provide a brief discussion of several robustness checks. All estimates for these robustness checks are reported in the appendix. The first check I present is a specification test in order to determine the extent to which the estimates rely on the set of fixed effects included in the analysis. I emphasize in the main analysis that identification does not rely on the inclusion of state fixed effects, though the only estimates I present in the main analysis either exclude state fixed effects altogether or include highly flexible state-by-month effects. In Table A3, I present several additional specifications. The results reported here indicate that the main effect changes little across a wide variety of specifications. Note that I do not include any state-level time-varying controls in the main analysis. One advantage of the triple-difference specification – which emphasizes the idea that influenza vaccination has no casual effect on mortality (or any other outcome) during periods in which there is no influenza circulating – is that it allows for the inclusion of state-by-year fixed effects. State-by-year fixed effects control for all factors common to all months within a given state-year, effectively controlling for any state-specific factors that vary at the annual level. The useful variation in vaccination rates is preserved in this setting, as the effect of vaccination depends on the level of influenza activity, which is not constant throughout the year. The main results are not sensitive to the inclusion of these fixed effects, which lends support to the identification strategy in general and obviates the need for state-specific time-varying controls.

Another concern is the lack of regional variation in match rates in the main analysis. Although the strains included in the vaccine are identical across North America, it is possible that regional variation in the circulating strains generates regional variation in the match rate. If a subset of regions tend to drive variation in match rates at the national level, and this variation is correlated with vaccination rates, then the inability to use regional variation in match rates could bias the estimates. To determine the extent to which this may be the case, I utilize data on regional variation in the strains of influenza that circulate in a given year to construct match rates at the regional level (census divisions) for the years in which this data are available. While there is variation across regions in the match rate, I find that the degree of variation tends to be minimal. I provide estimates of the main results that include the regional measure of match rates and present the estimates in Table A4. For the sake of comparison, I first present estimates that exclude years prior to 1998, but do not use regional variation in the match rate – I am not able to construct regional measures of the match rate prior to this year. This turns out to be unimportant, however, as the main estimates are not sensitive to the exclusion of earlier years. The estimates are also insensitive

to the use of regional match rates.[23]

Next, I address a concern that by comparing outcomes in months that have relatively high or low influenza activity, that I am effectively conditioning on an outcome. In the main analysis, influenza activity is defined at the national level, and I argue that it is highly unlikely that local conditions affect the timing and magnitude of influenza activity at the national level on such a scale that the results would be meaningfully affected. In any case, I present two additional sets of results to support this claim. First, using regional data on influenza activity (for years 2008 and beyond), I define influenza activity for each state as the average level of activity in all regions *except* the state of interest. In this specification, it is even less likely that the measure of influenza activity would be meaningfully affected by local conditions as the measure excludes all states in the region of the state of interest. These results are presented in Table A5; the estimates indicate that the results are insensitive to the use of this leave-one-out measure of activity.

Second, instead of using actual influenza activity, I defined the months December through March as months that typically have high influenza activity and all other months as months with typically low activity, and use this consistent set of months across all years in the analysis. The results from this analysis are presented in Table A6. As expected, these estimates are smaller in magnitude – these estimates can essentially be interpreted as an intent-to-treat effect, as the comparison is no longer between months with *actual* high or low levels of influenza activity, but a comparison between months with *typically* high or low influenza activity. That being said, the estimates can be scaled up for comparison with the main estimates; the average levels of influenza activity in the typically high and low months are 0.38 and 0.09, respectively. As such, I scale the estimates by a factor of 1/(0.38-0.09) and find that the estimates are similar to those in the main analysis.

Finally, to test for lagged impacts of influenza mortality, I replicate the estimates presented in Table 4, but additionally include one and two month lags in the interactions that include influenza activity (as these are the only interactions that vary across months within a year). The estimates reported in Table A7 represent the sum of the contemporaneous and lagged coefficients. The estimates indicate that allowing for lagged effects of mortality does not substantially alter the results. The estimates that include lagged coefficients are slightly larger in some specifications, but these differences are small.

[23]Note that this model includes the Match × Activity interaction and the main effect for Match, as these are no longer absorbed by the month-year fixed effects; the coefficients on these terms are not easily interpreted, as they represent out-of-sample predictions (i.e., a vaccination rate of zero).

4 Part II: Health Care Worker Mandates

The analysis conducted in Section 3 was intended to estimate the benefits of influenza vaccination in the general population. The estimates are relevant to a policy that would increase vaccination rates in proportion to the current distribution of vaccination rates. In other words, a hypothetical policy would target individuals who are closest to the margin of choosing whether to receive a vaccination. In this section, I recognize that there is likely to be substantial heterogeneity in benefits depending on who receives the vaccine. Health care workers (HCWs) come in relatively frequent contact with infected individuals *and* individuals whose cost of infection is high. As such, HCWs are a group for whom the external benefits of vaccination are large. I examine the effects of mandates requiring health care workers be vaccinated against influenza on the outcomes of their patients. To begin, I describe the institutional background and the roll-out of these policies.

4.1 Institutional Background

On September 28, 2006, the Governor of California signed into law Senate Bill 739, requiring that health facilities implement various measures to protect against the spread of infection within these facilities. One component of this law was to require all licensed hospitals in California to report to the Department of Public Health on the percentage of HCWs (employees and non-employee personnel) vaccinated against influenza in each season beginning in 2008/2009; fortunately, this allows for the measurement of vaccination rates within hospitals in California. This law would also require that all health facilities offer influenza vaccinations free of charge to all HCWs and require that employees sign a statement declaring that he or she had declined vaccination. Though detailed data on vaccination rates prior to this policy are not available, it is likely that these policies increased vaccination rates of HCWs. This is only relevant for this study in that baseline levels of vaccination are likely higher, however, as the variation in influenza *mandates* does not begin until 2009, when hospital-level vaccination rates are observed.

In May of 2009, the H1N1 influenza pandemic began. In response to the pandemic, several individual hospitals began requiring influenza vaccination for their workers. Because these mandates were implemented in response to the crisis, the timing of vaccination relative to the timing of the pandemic is unclear. For this and other reasons, I treat 2009 differently than other years and I will elaborate on this in more detail in the following section. After the 2009 pandemic, these hospitals continued requiring annual influenza vaccinations for their workers and in following influenza seasons several other hospitals began introducing their own vaccination requirements. Beginning in the 2011-12 influenza season, three counties

implemented county-wide influenza vaccination mandates. In the 2012-13 and 2013-14 influenza seasons, several more counties began implementing vaccination requirements. The influenza season in which these policies were implemented is indicated for all hospital- and county-level policies in Table A9. More than half of the hospitals in the state were subject to an influenza vaccination mandate in the 2013-14 influenza season. Because these mandates typically required HCWs to be vaccinated against influenza by the start of December at the latest, and since all influenza seasons in the sample past 2009 did begin until well into December or later, the timing problem that exists for the 2009 pandemic does not exist in the following influenza seasons.

The county-level policies were not all implemented in exactly the same fashion. Specifically, a limited number of these policies only applied only to hospitals, while others applied more broadly (i.e., all licensed health care facilities). Figure 6 maps the implementation of these policies over time and distinguishes between the type of county-level mandate. As is clear from this map, much of the variation in influenza vaccination requirements comes in the last two seasons in the sample, and most of the county-level policies apply to all licensed health care facilities.

4.2 Data

To estimate the effects of HCW vaccination mandates on patient outcomes, I make use of data on the timing of the mandates themselves, vaccination rates for HCWs and outcomes observed at the hospital.

4.2.1 Mandates

Data on the timing of mandates is compiled from several sources. Information on hospital-level mandates comes largely from the Immunization Action Coalition (IAC), a non-profit immunization activist group. IAC's "Influenza Vaccination Honor Roll" lists health care organizations across the U.S. that mandate influenza vaccination for their workers as well as the dates these policies were implemented.[24] This is not a comprehensive list of vaccination mandates, and it is possible that there are hospitals in the sample that do indeed have a vaccination mandate but are not classified as such in my data. The California Department of Public Health (CDPH) maintains a list of counties that require influenza vaccination mandates with information on the implementation date of each.[25] This list, however is not completely accurate with respect to the implementation dates. Through a process of search-

[24]Source:http://www.immunize.org/honor-roll/influenza-mandates/.
[25]Source: https://www.cdph.ca.gov/.

ing for county-level public health orders and identifying the initial date of implementation, I have either verified or amended the dates of nearly all counties on the list provided by the CDPH.[26] Summary statistics reported in Table 1 indicate the number and percentage of hospitals subject to vaccination requirements in each influenza season.

4.2.2 HCW Vaccination Rates

As required by California law, all licensed hospitals report, for each influenza season, information on their vaccination rates to the CDPH. This information is compiled in their annual Hospital Employee Influenza Vaccination Reports. Though all hospitals provide information on vaccination rates, the within-hospital response rate is not 100% (though most are well over 90%). The vaccination rate for each hospital and influenza season is calculated as the percentage of responding employees who are vaccinated. In estimates of the effects of vaccination mandates on vaccination rates (i.e., the first stage), I only use hospitals that have response rates of at least 80% in all years. In part due to a statewide law implemented in 2006 that encouraged but did not require vaccination, the baseline level of vaccination was relatively high for many hospitals even in the absence of a mandate. As indicated in Table 1, the average vaccination rate for hospital-years without a mandate is approximately 73%, compared to an average of 91% for hospital-years with a mandate in place.

4.2.3 Hospital Patient Outcomes

The primary data source on outcomes is restricted data on the universe of inpatient hospital admissions in California between 2005 and 2014, obtained through California's Office of Statewide Health Planning and Development (OSHPD). For each admission, I observe the hospital and date of admission, allowing the data to be merged with information on vaccination mandates and influenza activity.

Unlike the mortality data, in which influenza is rarely indicated as a cause of death, hospital patients routinely receive diagnoses specifically for influenza. This allows the outcome measure to be more specific in nature. The primary outcome of interest is admissions with a diagnosis specifically for influenza; admissions are classified as such if any of the up to 25 diagnoses are for influenza.

I also examine a set of less specific outcomes that may be affected during periods of high influenza activity. These outcomes include a more aggregate classification for influenza-related diagnoses (pneumonia/influenza), the patient's length of stay, and the patient's total hospital charges. Data on all outcomes is collapsed to the monthly level, where the admissions

[26] I have compiled a number of public health orders and these documents are available upon request.

are hospital-by-month counts of influenza-related hospital admissions, and the less specific measures are average length of stay, and average charges. Summary statistics for all outcomes are presented in Table 1. In addition to these hospital-level measures, I also examine PI mortality, which is observed at the county level using restricted data files from the National Vital Statistics System (i.e., the same outcome as in the analysis in Part I, though at the *county* level).

4.3 Empirical Framework

I estimate the impacts of these vaccination requirements using a standard difference-in-differences framework that exploits the roll-out of these policies over time. The data used in this analysis is described in Section 4.2. Similar to the mortality analysis, I also use a triple-difference framework that additionally exploits the timing and magnitude of influenza activity for outcomes that exhibit variation throughout the year. Because HCW vaccination rates are measured annually, the triple-difference strategy is not necessary in estimates of the first-stage. Furthermore, the primary outcomes of interest is influenza-specific diagnoses (i.e., more specific than PI diagnoses). Because there is no variation in diagnoses that are specific to influenza in months with no influenza circulating, the triple-difference strategy is not be necessary for this outcome.[27] Consider the following difference-in-differences equation to be estimated at the annual level:

$$\mathrm{Y}_{hy} = \alpha + \pi \mathrm{Required}_{hy} + \delta_h + \delta_y + \varepsilon_{hy} \tag{5}$$

In Equation (5), Y_{hy} represents either vaccination rates (first stage), or the number of influenza diagnoses (reduced form) at hospital h in year y. $\mathrm{Required}_{hy}$ is a variable indicating whether there is a vaccination requirement in effect; δ_h and δ_y are hospital and year fixed effects.

The coefficient of interest, π, is identified under the assumption that variation in the timing of influenza vaccination mandates is uncorrelated with other unobserved time-varying determinants of the outcomes of interest. The hospital fixed effects control for all hospital-specific and time invariant factors such as hospital size or hospital type (e.g., research hospitals). The year fixed effects account for factors common to all hospitals and specific to a given year, such as changes in state or national health policy. While the identifying assumption is fundamentally un-testable, I provide evidence from a number of indirect tests that support the assumption. Importantly, in the discussion of results I provide an event

[27] Estimation of the triple-difference specification is possible for this outcome, but results in extremely large standard errors.

study version of Equation (5); this exercise indicates that changes in the outcomes of interest coincide with the implementation of the policy, and that treatment effects are not identified off of differential trends between treatment and control hospitals.

Because influenza diagnoses represent a highly specific outcome, many hospital-years have zero admissions with influenza diagnoses (especially small hospitals). Figure 7 displays the distribution of this outcome, which clearly indicates that a count model would be preferred to an OLS estimator. There are several possible count models available, and in the case of panel data requiring fixed effects (as here) the choice is not trivial.[28] The workhorse count model that allows for fixed effects is the Poisson fixed effects estimator (Hausman et al., 1984; Wooldridge, 1999); this estimator, unlike many nonlinear models, provides consistent estimates of the slope parameters in the presence of fixed effects. A deficiency of the Poisson model, however, is that it assumes that the variance and mean of the outcome are equal (i.e., equi-dispersion). Table 1 clearly indicates the presence of over-dispersion in the hospitalization counts (i.e., the variance is greater than the mean). The usual solution is to use a negative binomial in place of a Poisson model, which allows for over-dispersion in the data. Hausman et al. (1984) offer a fixed-effects version of the negative binomial, but subsequent work has pointed out that this model requires an additional and often unrealistic assumption regarding the relationship between the fixed effects and the over-dispersion parameter (Allison and Waterman, 2002; Guimaraes, 2008). An alternative strategy is to estimate a standard negative binomial model with a full set of indicators as fixed effects. In nonlinear models using short panels, this leads to biased and inconsistent estimates of the slope parameters due to an incidental parameters problem. That being said, Allison and Waterman (2002) provide evidence from Monte Carlo simulations that suggests little bias resulting from the incidental parameters problem in the case of the negative binomial model with indicator as fixed effects. I adopt the negative binomial with indicators as fixed effects as the main specification, though I also show that the results are not sensitive to the choice of count model.[29] With a count model it is important to allow the probability of an event to occur (i.e., an influenza-related diagnosis) to differ by hospital size. This is done through the use of an exposure variable, which I set to be the total number of all-cause admissions at a given hospital over the entire sample period.

For outcomes that vary across all months of the year, the preferred specification is a triple difference, estimated at the monthly level and taking the following form:

[28] See Cameron and Trivedi (2013a,b) for a review of count models in general and specifically for panel data.

[29] Estimates for a fixed effects Poisson model and a zero-inflated negative binomial model are presented in Table A10.

$$Y_{hym} = \alpha + \theta_1(\text{Required}_{hym} * \text{Activity}_{ym}) + \theta_2 \text{Required}_{hym} + \delta_h + \delta_{ym} + \varepsilon_{hym} \quad (6)$$

In Equation (6), the policy indicator is interacted with an index of influenza activity, Activity_{ym}. This measure, described previously, is an index that measures influenza activity at any particular time and ranges from zero to one (where one is the maximum observed value in the sample). As such, θ_1 measures the effect of influenza vaccination mandates during a time of peak influenza activity relative to a period with zero influenza activity. Further, θ_2 measures the effect of influenza vaccination mandates during times of very low influenza activity and is expected to be near zero. This in itself provides an additional test of the validity of the (stronger) assumption required for the difference-in-differences model, and demonstrating that estimates of θ_2 are near zero lends additional credence to the validity of the assumptions behind the empirical model described in Equation (5).

Because the outcomes of interest here (average charges, average length of stay, PI diagnoses, and PI mortality) are less specific in nature, vary throughout all months of the year, and are rarely equal to zero, a count model is not necessary for the analysis of these variables.[30] The distributions for length of stay and charges at the micro-level (i.e., before collapsing to the hospital-month level) have extremely long tails. To ensure that my estimates are not driven by these outliers, I exclude micro-level observations that are above the 99th percentile of each variable's distribution before calculating monthly averages. Additionally, charges are not reported for all inpatient visits. Some hospitals in particular consistently fail to report charges. Because these observations are unlikely to be missing randomly, I exclude hospitals that do not report charges for at least 95% of their patients over the sample period in the analysis of average charges.[31]

Finally, the data on mortality are derived from a different data source than the hospital-level measures. The mortality data are only available at the county level and as such only county-level vaccination mandates are used in the analysis of this outcome. Furthermore, it should be noted that the estimating equation is slightly distinct in that county fixed effects are used in place of hospital fixed effects.

[30] While PI diagnoses still represent a count, these are not typically near zero and thus do not require a count model.

[31] This consists of approximately 13% of the hospitals in the sample.

4.4 Results

The estimates for the first stage and influenza-specific diagnoses are derived using a model estimated at the annual level (described in Equation (5)) and are presented in Table 10. The first-stage estimates indicate that influenza vaccination mandates led to increases in the HCW vaccination rate of approximately 9.9 percentage points. It is important to keep in mind that the first-stage estimates only represent vaccination rates for hospital workers. This is especially important in considering the county-level requirements, which often apply far more broadly than to just hospital workers. Because I do not observe the vaccination rates of other HCWs, I prefer to display the results in the remainder of the paper as reduced-form policy estimates rather than providing results in an instrumental variables framework. That being said, there is reason to believe that the first-stage effect for non-hospital HCWs may be larger than that of hospital HCWs. The CDC conducts an online survey that provides national estimates of influenza vaccination for HCWs by place of work. The survey applying to the 2014-15 influenza season indicated vaccination rates of 78.7%, 66.3%, 54.4% and 55.7% for HCWs in hospital, ambulatory care, long-term care, and other settings, respectively. Because hospital workers tend to have the highest baseline vaccination rate, it is likely that influenza vaccination requirements have a larger effect on workers in settings with a lower baseline level. That being said, it is also possible that enforcement is weaker in non-hospital settings; unfortunately I cannot test this with the data available.

The reduced-form estimates presented in Table 10 indicate that HCW vaccination mandates lead to an approximate 17% reduction in hospital visits that include a diagnosis for influenza. Recall that these estimates are derived using a negative binomial model; and the coefficients reported may be interpreted as percent changes. For all outcomes I present results that include all years, and results that exclude the 2008-09 and 2009-10 influenza seasons; this is to ensure that it is not conditions specific to the H1N1 pandemic that are driving the results. In general, the results are quite similar regardless of this exclusion.

At this point, it is not clear whether the reduction in influenza diagnoses is due to a reduction in new admissions (through influenza acquired in a non-hospital health care setting) or a reduction in influenza acquired in the hospital; I explore this further in Section 4.4.1. For now, however, I focus on establishing that the observed reduction is indeed a causal effect.

The results presented in Table 10 are supplemented with event study versions of both the first-stage and reduced form estimates in Figures 8 and 9. The event-study analysis for the first-stage reveals no statistically significant effects on influenza vaccination in the years prior to the implementation of the policy, and highly significant positive effects in the first year that the mandates are implemented and each year thereafter. For the reduced-form

event study, I again present versions that include or exclude H1N1 pandemic seasons. The event-study estimates reveal that there are substantial pre-treatment trends in the outcomes in treated hospitals compared to the control hospitals. Specifically, it appears as though influenza diagnoses were increasing in hospitals that adopted these mandates relative to hospitals that did not. In the first year of policy implementation, however, there exists a dramatic decrease in the adopting hospitals. This decrease appears to be sustained over time, though the differential pre-treatment trend appears to persist in the post-period as well. Because the trends in the outcome are in the opposite direction of the treatment effects, this implies that estimates of the treatment effects that do not account for these trends are biased towards zero. Indeed, specifications that include hospital-specific time trends exhibit somewhat larger point estimates. These estimates are provided in Table A10.

Examining the age groups affected by these policies reveals strikingly different results relative to the estimates of influenza vaccination rates on mortality. Table 11 reports results for the following groups: under 1, 1-9, 10-64 and over 64. The estimates indicate that treatment effects are strongest for children, with infants (under 1) being affected most. In general, it is not surprising that infants would benefit more on the dimension of hospitalization relative to mortality, as infants are quite likely to be hospitalized for influenza, but unlikely to die as a result of infection. The point on hospitalizations is emphasized in Figure 10, which plots the distribution of influenza diagnoses by single year of age; the plot indicates that infants are unique in that the number of hospitalizations for this group is much larger than for any other single year of age.

While the age distribution indicates that infants are more likely to be hospitalized in general, the estimates indicate that infants receive larger benefits of HCW vaccination compared to other age groups. There are several potential reasons why this could be the case. First, many hospitalizations for children occur in children's hospitals; if infectious diseases are more likely to spread in a children's hospital setting, then it is likely that increasing the vaccination rate of employees in these hospitals would have a relatively large effect. Also note that much of the variation in these policies is identified off of mandates that apply to all health care facilities in a county, including primary care settings; if pediatric primary care offices act as an important vector for disease, then it is again likely that increasing vaccination in these settings would have a disproportionately large effect. In each case, it is not unreasonable to argue that the spread of disease is more likely among children relative to adults.

Next, I move on to discussing the results for the less specific outcomes estimated in the triple-difference framework. Specifically, the results for PI diagnoses, average charges, average length of stay, and PI mortality are presented in Table 12. Note that all outcomes

are measured at the hospital level with the exception of PI mortality. The data on mortality come from the National Vital Statistics System and are measured at the county level, and thus only county-level HCW vaccination mandates are relevant to these estimates. While the estimates for PI diagnoses and PI mortality are presented for completeness, these are not precisely estimated.[32]

For both average length of stay and average charges, the estimates exhibit greater precision and indicate that HCW vaccination leads to improvements during months of high influenza activity. For both, there is no statistically significant effect during periods in which influenza is not circulating. The estimates for both length of stay and charges are statistically different from zero at the 10% level in models that include all years, and at the 5% level in models that exclude H1N1 pandemic years.

4.4.1 Mechanisms

To this point, I have provided evidence that patients benefit from HCW vaccination. What is not yet clear is the mechanism through which these benefits are realized. In this section, I attempt to discern whether these impacts are driven by one of two alternative (though not mutually exclusive) mechanisms. The first potential mechanism is that the benefits are driven by factors *within* the hospital. The most obvious is the reduced probability that a patient is infected with influenza during their stay (i.e., hospital-acquired infection). It is also possible, however, that by vaccinating health care workers and improving the health of the hospital staff, the hospital is less likely to be understaffed or is otherwise operating at greater capacity. The second potential mechanism is a reduced probability of infection in a non-hospital health care setting. In this case, the reduction in influenza diagnoses is a result of fewer admissions rather than fewer in-hospital infections. This is certainly possible as most HCW mandates apply not only to hospitals, but all licensed health care facilities within a county.

The results indicating that HCW vaccination mandates lead to shorter stays with lower charges support the idea that at least some of the benefits are derived through within-hospital factors. To test this more directly, however, I make use of a feature of the hospitalization

[32]For example, the 95% confidence interval for the coefficient on PI diagnoses is (-9.6%, 12.9%). Multiplying by average monthly influenza activity gives the annual effect and associated confidence interval (2.4%, 3.3%). Note that for every influenza-specific diagnosis there are approximately 27 PI diagnoses; if the influenza-specific diagnoses captured the full impact of these mandates (which is unlikely), then the implied effect on PI admissions (based on the 17% reduction from Table 10) would be approximately -0.6% annually. If only one in four influenza illnesses were correctly diagnosed with an influenza-specific ICD code (rather than a more general PI code), the implied effect on PI admissions would still be within the confidence interval for the estimates using PI diagnoses. Due to the lack of precision in these estimates, I defer to the estimates using influenza-specific diagnoses as my preferred estimate for the effects on influenza-related diagnoses.

data that indicates whether each diagnosis was present at the time that the patient was admitted. In Table 13, I separately analyze the effects of HCW mandates on influenza diagnoses that were or were not present at the time of admission. I interpret diagnoses that were present at the time of admission as new admissions, and diagnoses not present at the time of admission as hospital-acquired. Because a large proportion of total influenza diagnoses are classified as present at admission, the estimates for this category are quite similar to those in the main analysis. The estimates for diagnoses that were not present at admission are substantially larger in percentage terms; the estimates indicate that HCW vaccination mandates lead to an approximate 35% reduction in hospital-acquired influenza diagnoses. In general, this evidence suggests that the reduction in influenza diagnoses is driven by both influenza acquired within and outside of the hospital setting.

4.4.2 Estimating Benefits

As with the analysis conducted at the national level, one of the primary purposes of this analysis is to estimate the marginal social benefit of vaccination in monetary terms. Due to relatively imprecise estimates and the fact that it will be necessary to make somewhat stronger assumptions, the figure I derive here should be taken with some degree of caution relative to the estimates for the national-level analysis. I believe that this back-of-the-envelope calculation is informative nonetheless.

To estimate the marginal social benefit of HCW vaccination, I use the estimate corresponding to the reduction in influenza diagnoses that were present on admission, assuming that these represent new admissions (i.e., the admissions would not have taken place in absence of the policy). To arrive at a figure of benefits per vaccination, I use figures corresponding to California as a whole and consider a hypothetical policy that considers a statewide HCW vaccination mandates (relative to no mandate). First, the mean annual number of admissions with an influenza diagnoses that was present on admission is 6,451. This figure excludes the H1N1 pandemic season, during which there were significantly more admissions relative to a typical season. Furthermore, this figure includes the years in which mandates were in place for some hospitals, and thus represents a lower bound on the counterfactual in which HCW mandates were never introduced. Multiplying this figure by the estimated reduction indicates 1,097 fewer admissions each year. The average charges for these influenza-related admissions is $58,803. Since charges do not necessarily represent the cost of providing services or the willingness to pay to avoid an illness, I prefer to use a measure of the cost of providing hospital services. This measure is obtained by multiplying charges by a cost-to-charge ratio. This ratio (0.288) is specific to hospitals in California and is derived from the National Inpatient Sample provided by the Healthcare Cost and Uti-

lization Project. Average hospital costs for influenza-related admissions are then $16,935. The implied benefits of these mandates are approximately $18.5 million per year if applied statewide. To arrive at a per-vaccination figure, I require an estimate of the number of vaccinations required to achieve this. The Bureau of Labor Statistics estimates there were approximately 1.10 million workers in occupations related to the health care industry in California during 2015. Multiplying by the first-stage estimate of 0.098 implies 108,057 additional vaccinations would be received in the hypothetical statewide policy. The implied benefits are equal to $171.9 per HCW vaccination on the margin of reduced hospital costs.

5 Discussion and Conclusion

In this paper, I estimate the marginal social benefits of influenza vaccination for the general population and for the population of health care workers. In both cases, I provide policy-relevant estimates of the benefits per vaccination: these benefits are estimated to be $67 in terms of reduced mortality (at minimum), $49 in terms of work hours gained, and $172 per health care worker vaccination in terms of reduced health care costs. How do these benefits compare to the costs of prospective vaccination programs? Prosser et al. (2008) estimates that the cost of administering a vaccination (including the medicine, labor, overhead, promotion, and other expenses) ranges from $15 in a mass vaccination setting to $37 in a schedule doctor's office visit.[33] Administration costs, however, may only represent a portion of the total private costs of vaccination if there are significant non-monetary costs such as inconvenience or discomfort. Indeed, many choose not to vaccinate despite monetary costs equal to zero (influenza vaccination is covered under Medicare, and many health plans cover vaccination with zero copay). Recognizing these non-monetary costs of vaccination is critical in the development of policies that encourage influenza vaccination.

What do these estimates suggest for vaccination policy? The answer to this question depends on the type of policy under consideration. Let us consider two prospective vaccination policies in turn: a policy to increase vaccination in the general population and a policy to increase the vaccination rates among health care workers.

The analysis of aggregate vaccination rates is relevant to a policy that would increase vaccination rates among the general population by targeting those on the margin of the decision to vaccinate. Such a policy could be accomplished through a number of mechanisms: by providing monetary incentives or by reducing non-monetary costs through increasing accessibility to vaccine providers, for instance. The marginal social benefits that I estimate suggest that vaccination policy resulting in marginal increases in the vaccination rate above

[33]Dollar estimates are converted to $US 2016 dollars

current levels is beneficial so long as the marginal cost curve does not increase steeply at the current level of vaccination. While such a steep increase in the cost curve is conceivable at some level of vaccination, as some individuals are opposed to vaccination on religious grounds or concerns over vaccine safety, that level is likely to be quite high as those individuals represent only a small portion of the total population (Kennedy et al., 2005).

Policymakers may even consider incentive schemes that target the socially optimal level of vaccination in the population (i.e., non-marginal changes). Models of disease dynamics, such as Boulier et al. (2007), suggest that the marginal benefits of vaccination are relatively constant until a threshold level of vaccination is reached and an epidemic is no longer experienced. If the marginal social cost curve does not increase steeply prior to this threshold, the results presented here suggest that targeting vaccination rates at a level near the threshold is likely to be socially optimal. If policymakers are risk averse and the location of the threshold is unknown, it may be the case that targeting a level of vaccination that is safely beyond the threshold is optimal as well, so long as the marginal costs are not too large.

Considering vaccination policy that applies only to health care workers is a somewhat more simple thought experiment: such policy only affects a small portion of the population, and is unlikely to result in such large benefits that a threshold level of vaccination is reached. For health care workers, the estimates presented here are large in comparison to adminstration costs. It is worth noting that many health care facilities employ mass vaccination campaigns that not only reduce the administrative costs of vaccination, but likely reduce any inconvenience costs through making vaccination highly accessible (Prosser et al., 2008; Nowalk et al., 2013). The estimates in this paper are relevant to policies that *mandate* influenza vaccination, creating an extremely high cost for those choosing not to vaccinate.[34] It is possible that other incentive-based programs could achieve a more efficient result if there are individuals for which the marginal cost of vaccination is very high, yet still choose to vaccinate under a mandate given an even higher cost of choosing not to do so. In any case, the social benefits of health care worker vaccination that I estimate are large, indicating that policies increasing vaccination among health care workers are cost-effective under reasonable assumptions regarding the costs.

In summary, I estimate that the social benefits of influenza vaccination are substantial and that much of the total benefits are realized through externality effects. Determining the socially optimal level of vaccination depends critically on the marginal cost of vaccination and the shape of the marginal cost curve – under reasonable assumptions about these marginal costs, the results of this study indicate that policies that encourage take-up of influenza

[34]Most mandates do allow for medical exemptions and for those choosing not to vaccinate to wear a surgical mask through the remainder of influenza season.

vaccination in either the general population or in the population of health care workers are cost-effective.

References

Allison, Paul D and Richard P Waterman, "Fixed-Effects Negative Binomial Regression Models," *Sociological Methodology*, 2002, *32* (1), 247–265.

Althouse, Benjamin M, Theodore C Bergstrom, and Carl T Bergstrom, "A public choice framework for controlling transmissible and evolving diseases," *Proceedings of the National Academy of Sciences*, 2010, *107* (suppl 1), 1696–1701.

Ashenfelter, Orley and Michael Greenstone, "Using Mandated Speed Limits to Measure the Value of a Statistical Life," *Journal of Political Economy*, 2004, *112* (S1), S226–S267.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro, "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century," *Journal of Political Economy*, 2016, *124* (1), 105–159.

Boulier, Bryan L, Tejwant S Datta, and Robert S Goldfarb, "Vaccination externalities," *The BE Journal of Economic Analysis & Policy*, 2007, *7* (1).

Brito, Dagobert L, Eytan Sheshinski, and Michael D Intriligator, "Externalities and compulsary vaccinations," *Journal of Public Economics*, 1991, *45* (1), 69–90.

Cameron, A Colin and Pravin K Trivedi, "Count Panel Data," *Oxford Handbook of Panel Data Econometrics*, 2013.

_ **and** _ , *Regression analysis of count data*, Vol. 53, Cambridge university press, 2013.

Caplan, Arthur, "Time to mandate influenza vaccination in health-care workers," *The Lancet*, 2011, *378* (9788), 310–311.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment*.," *Quarterly Journal of Economics*, 2013, *128* (2).

Demicheli, Vittorio, Tom Jefferson, Lubna A Al-Ansary, Eliana Ferroni, Alessandro Rivetti, and Carlo Di Pietrantonj, "Vaccines for preventing influenza in healthy adults," *The Cochrane Library*, 2014.

Dushoff, Jonathan, Joshua B Plotkin, Cecile Viboud, David JD Earn, and Lone Simonsen, "Mortality due to Influenza in the United States An Annualized Regression Approach Using Multiple-Cause Mortality Data," *American Journal of Epidemiology*, 2006, *163* (2), 181–187.

Francis, Peter J, "Dynamic epidemiology and the market for vaccinations," *Journal of Public Economics*, 1997, *63* (3), 383–406.

_ , "Optimal tax/subsidy combinations for the flu season," *Journal of Economic Dynamics and Control*, 2004, *28* (10), 2037–2054.

Geoffard, Pierre-Yves and Tomas Philipson, "Disease eradication: private versus public vaccination," *The American Economic Review*, 1997, *87* (1), 222–230.

Grohskopf, Lisa A, Sonja J Olsen, Leslie Z Sokolow, Joseph S Bresee, Nancy J Cox, Karen R Broder, Ruth A Karron, and Emmanuel B Walter, "Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP)United States, 2014–15 influenza season," *MMWR Morb Mortal Wkly Rep*, 2014, *63* (32), 691–7.

Guimaraes, Paulo, "The fixed effects negative binomial model revisited," *Economics Letters*, 2008, *99* (1), 63–66.

Hausman, Jerry, Bronwyn H Hall, Zvi Griliches et al., "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica*, 1984, *52* (4), 909–38.

Hooper, CR, A Breathnach, and R Iqbal, "Is there a case for mandating influenza vaccination in healthcare workers?," *Anaesthesia*, 2014, *69* (2), 95–100.

Jefferson, Tom, Alessandro Rivetti, Carlo Di Pietrantonj, Vittorio Demicheli, and Eliana Ferroni, "Vaccines for preventing influenza in healthy children," *The Cochrane Library*, 2012.

_ , Carlo Di Pietrantonj, Lubna A Al-Ansary, Eliana Ferroni, Sarah Thorning, and Roger E Thomas, "Vaccines for preventing influenza in the elderly," *The Cochrane Library*, 2010.

Keech, Martin and Paul Beardsworth, "The impact of influenza on working days lost," *Pharmacoeconomics*, 2008, *26* (11), 911–924.

Kennedy, Allison M, Cedric J Brown, and Deborah A Gust, "Vaccine beliefs of parents who oppose compulsory vaccination.," *Public health reports*, 2005, *120* (3), 252.

Loeb, Mark, Margaret L Russell, Lorraine Moss, Kevin Fonseca, Julie Fox, David JD Earn, Fred Aoki, Gregory Horsman, Paul Van Caeseele, Khami Chokani et al., "Effect of influenza vaccination of children on infection rates in Hutterite communities: a randomized trial," *Journal of the American Medical Association*, 2010, *303* (10), 943–950.

Manski, Charles F, "Vaccination with Partial Knowledge of External Effectiveness," *Proceedings of the National Academy of Sciences*, 2010, *107* (9), 3953–3960.

_ , "Mandating Vaccination with Unknown Indirect Effects," *Working Paper*, 2016.

McLean, Huong Q, Mark G Thompson, Maria E Sundaram, Burney A Kieke, Manjusha Gaglani, Kempapura Murthy, Pedro A Piedra, Richard K Zimmerman, Mary Patricia Nowalk, Jonathan M Raviotta et al., "Influenza vaccine effectiveness in the United States during 2012–2013: variable protection by age and virus type," *Journal of Infectious Diseases*, 2014, p. jiu647.

Miguel, Edward and Michael Kremer, "Worms: identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 2004, *72* (1), 159–217.

Molinari, Noelle-Angelique M, Ismael R Ortega-Sanchez, Mark L Messonnier, William W Thompson, Pascale M Wortley, Eric Weintraub, and Carolyn B Bridges, "The annual impact of seasonal influenza in the US: measuring disease burden and costs," *Vaccine*, 2007, *25* (27), 5086–5096.

Murphy, Kevin M and Robert H Topel, "The Value of Health and Longevity," *Journal of Political Economy*, 2006, *114* (5).

Neuzil, Kathleen Maletic, Beverly G Mellen, Peter F Wright, Edward F Mitchel Jr, and Marie R Griffin, "The effect of influenza on hospitalizations, outpatient visits, and courses of antibiotics in children," *New England Journal of Medicine*, 2000, *342* (4), 225–231.

Nowalk, Mary Patricia, Chyongchiou Jeng Lin, Mahlon Raymund, Jamie Bialor, and Richard K Zimmerman, "Impact of hospital policies on health care workers influenza vaccination rates," *American journal of infection control*, 2013, *41* (8), 697–701.

Osterholm, Michael T, Nicholas S Kelley, Alfred Sommer, and Edward A Belongia, "Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis," *The Lancet infectious diseases*, 2012, *12* (1), 36–44.

Prosser, Lisa A, Megan A OBrien, Noelle-Angelique M Molinari, Katherine H Hohman, Kristin L Nichol, Mark L Messonnier, and Tracy A Lieu, "Non-traditional settings for influenza vaccination of adults," *Pharmacoeconomics*, 2008, *26* (2), 163–178.

Simonsen, Lone, Robert J Taylor, Cecile Viboud, Mark A Miller, and Lisa A Jackson, "Mortality benefits of influenza vaccination in elderly people: an ongoing controversy," *The Lancet infectious diseases*, 2007, *7* (10), 658–666.

Stearns, Jenna and Corey White, "Can Paid Sick Leave Mandates Reduce Leave Taking?," 2016.

Stewart, Alexandra M, "Mandatory vaccination of health care workers," *New England Journal of Medicine*, 2009, *361* (21), 2015–2017.

Stiglitz, Joseph E, *Economics of the Public Sector.*, W.W. Norton and Company, New York, 1988.

Thomas, Roger E, Tom Jefferson, and Toby J Lasserson, "Influenza vaccination for healthcare workers who care for people aged 60 or older living in long-term care institutions," *The Cochrane Library*, 2016.

Thompson, MG, DK Shay, H Zhou, CB Bridges, PY Cheng, E Burns, JS Bresee, NJ Cox et al., "Estimates of deaths associated with seasonal influenza-United States, 1976-2007.," *Morbidity and Mortality Weekly Report*, 2010, *59* (33), 1057–1062.

Ward, Courtney J, "Influenza vaccination campaigns: is an ounce of prevention worth a pound of cure?," *American Economic Journal: Applied Economics*, 2014, *6* (1), 38–72.

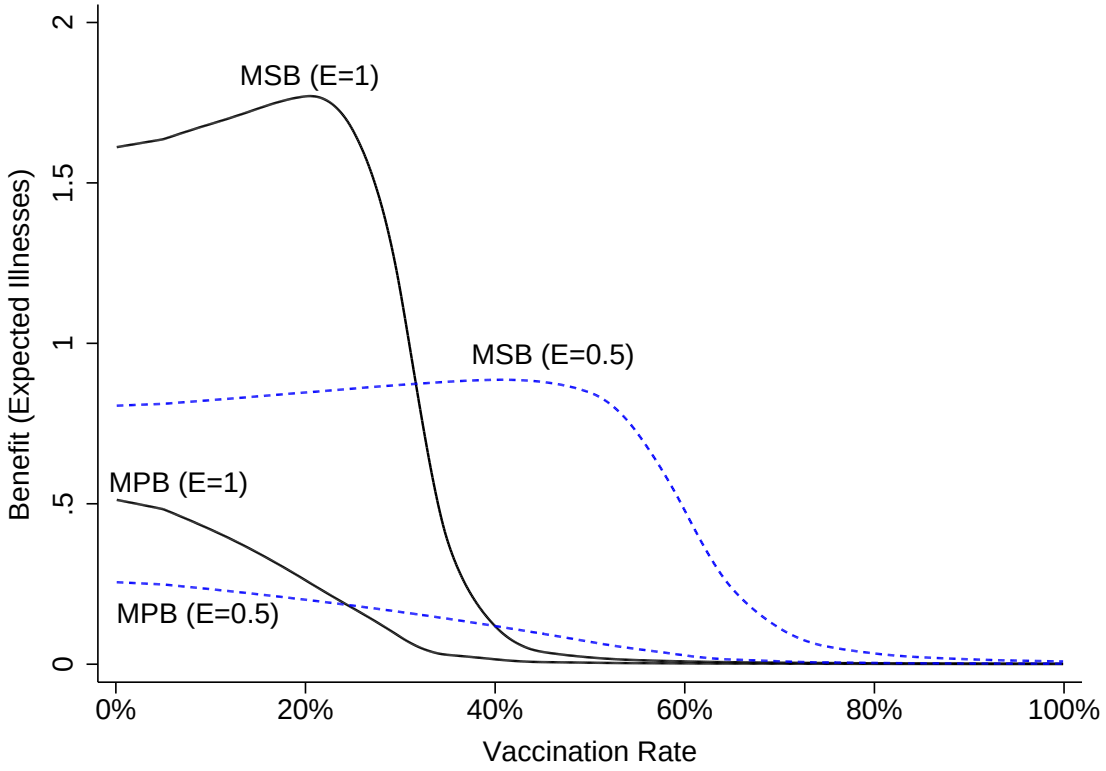
Wooldridge, Jeffrey M, "Distribution-free estimation of some nonlinear panel data models," *Journal of Econometrics*, 1999, *90* (1), 77–97.

Zhou, Hong, William W Thompson, Cecile G Viboud, Corinne M Ringholz, Po-Yung Cheng, Claudia Steiner, Glen R Abedi, Larry J Anderson, Lynnette Brammer, and David K Shay, "Hospitalizations associated with influenza and respiratory syncytial virus in the United States, 1993–2008," *Clinical infectious diseases*, 2012, *54* (10), 1427–1436.

Zivin, Joshua Graff and Matthew Neidell, "Environment, Health, and Human Capital," *Journal of Economic Literature*, 2013, *51* (3), 689–730.

Figures & Tables

Figure 1: Marginal Benefits of Vaccination



This figure reproduces the results of Boulier et al. (2007), describing the theoretical marginal social and marginal private benefits of influenza vaccination for two levels of vaccine effectiveness (100% and 50%). Even in well-matched seasons, estimates of influenza vaccine effectiveness are closer to the 50% figure.

Figure 2: Geography of Vaccination and Mortality Rates

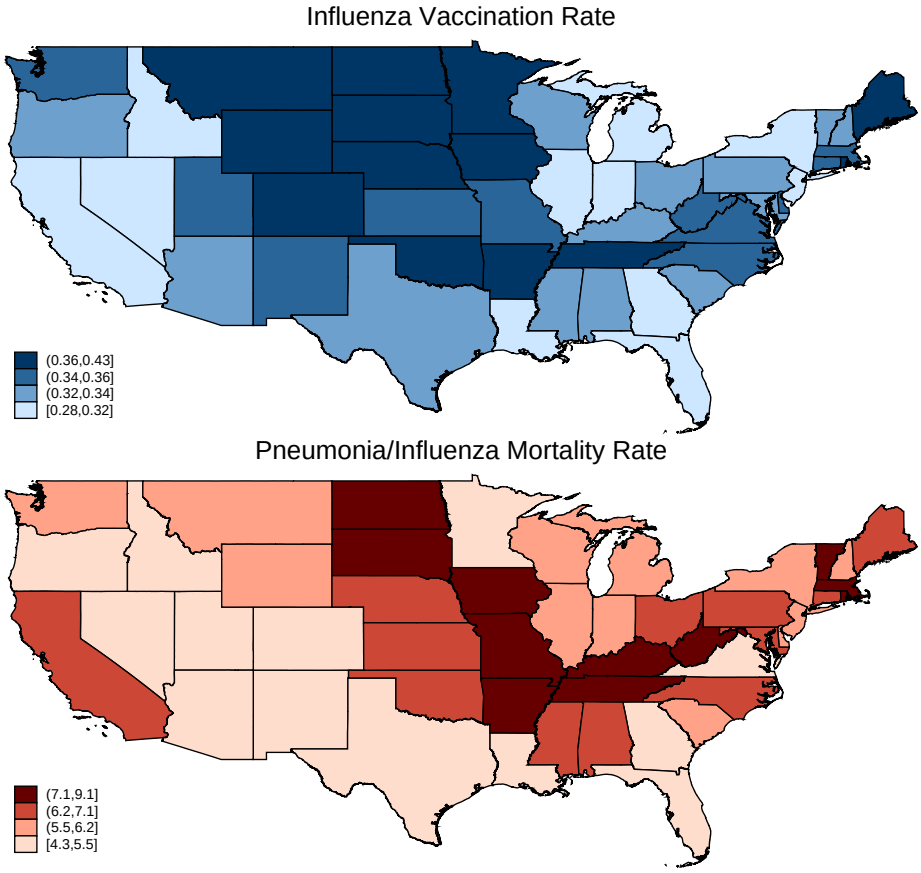
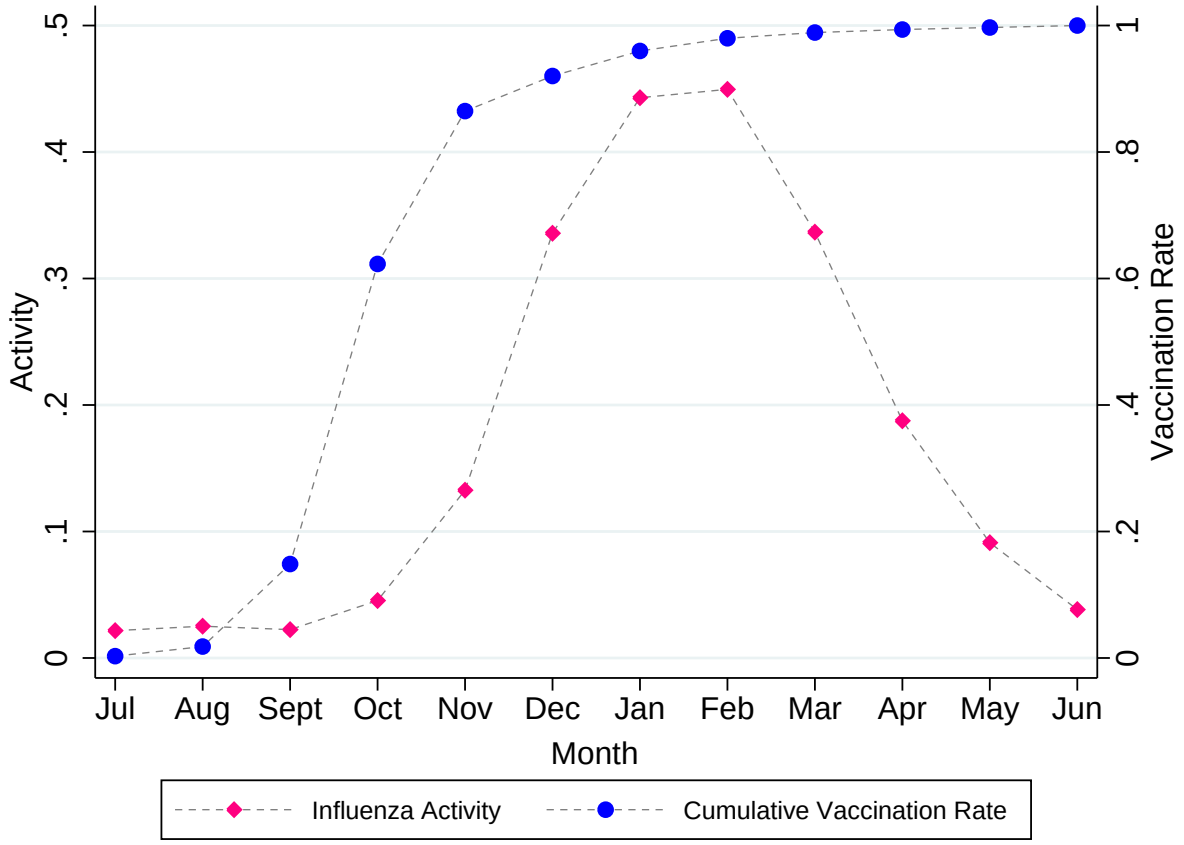
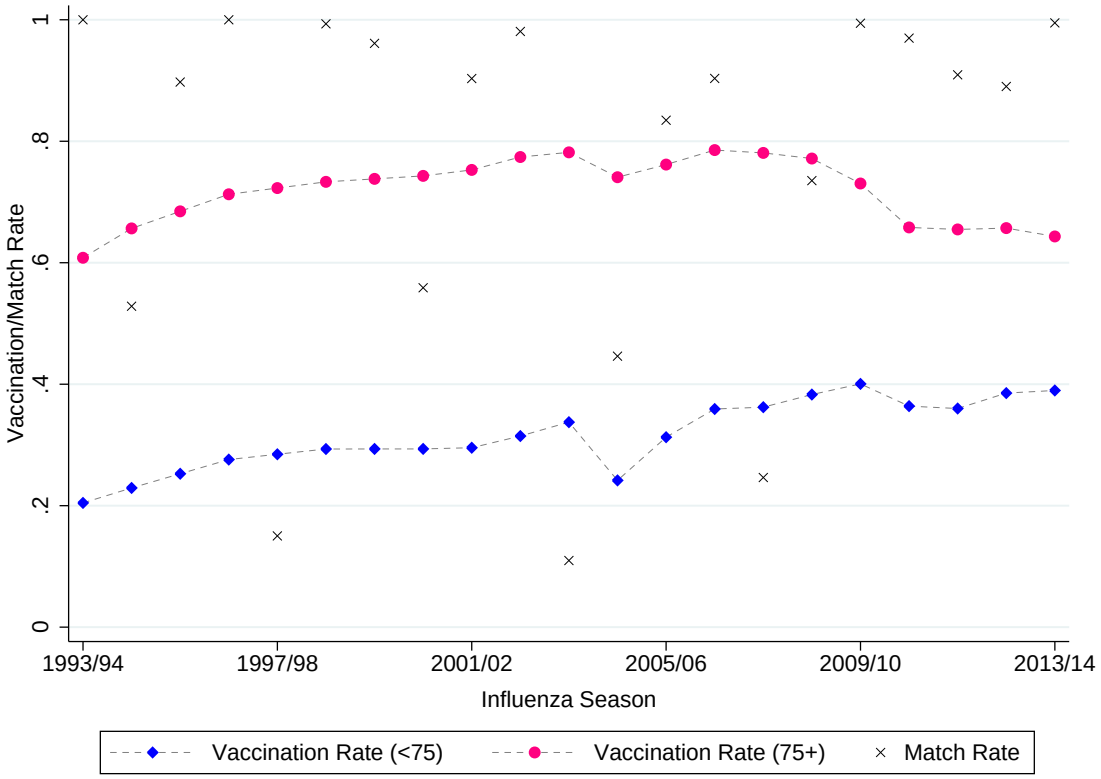


Figure 3: Vaccination and Influenza Timing



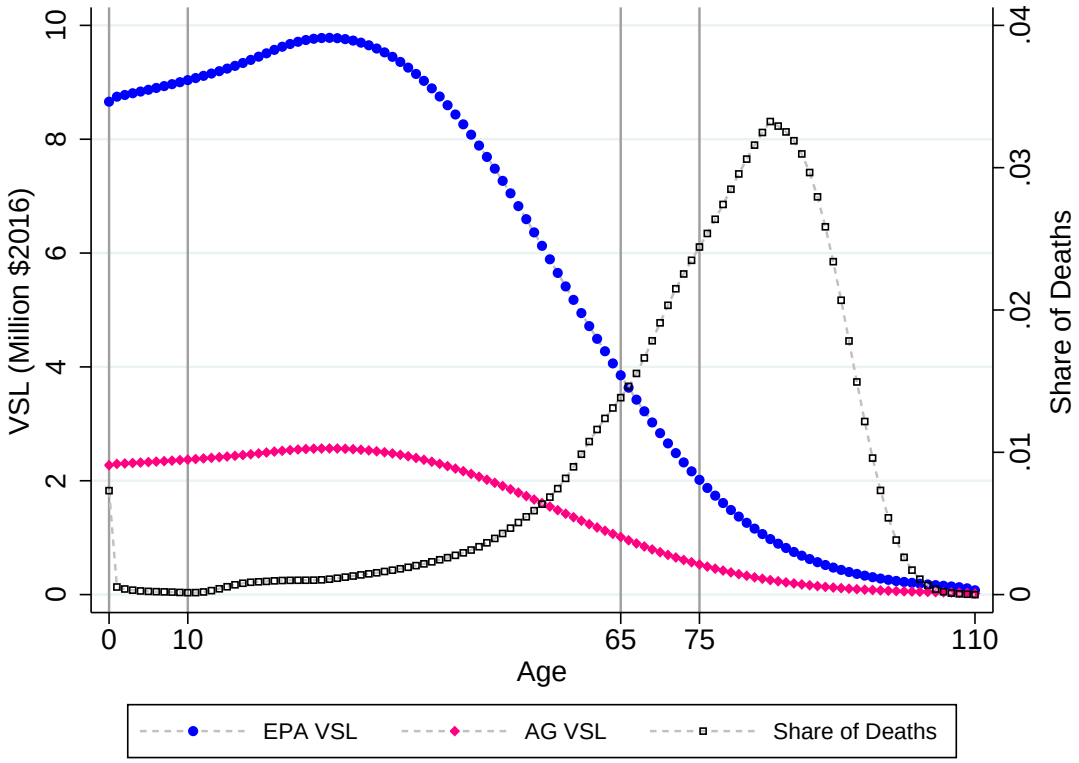
Note – This plot displays average monthly influenza activity and the average cumulative vaccination rate across years. Data on the timing of vaccination is available beginning in 2007. The year of the H1N1 influenza pandemic (2009) was excluded from the averages represented in this figure as it was a highly abnormal year in terms of the timing of both influenza activity and vaccination.

Figure 4: Actual and Effective Vaccination Rates



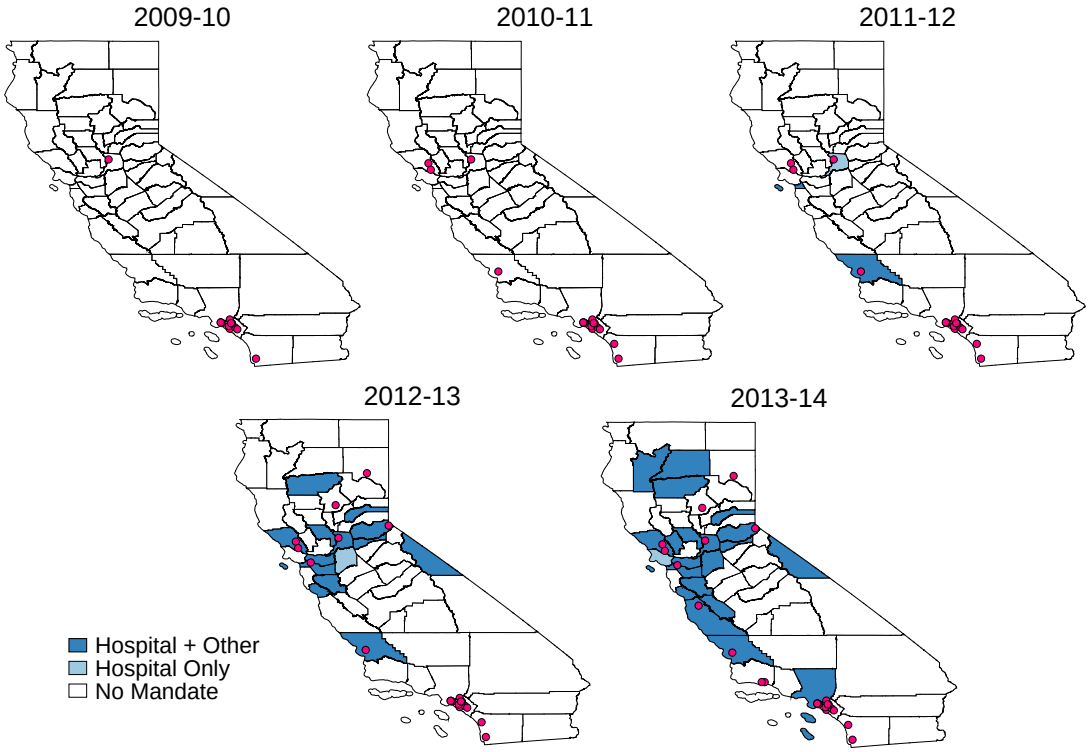
Note – There was a vaccine shortage in the 2004/05 season. As a result of this shortage, vaccines were allocated to "high-risk" groups including the elderly; as such, members of the non-elderly adult population (18-64) were disproportionately affected.

Figure 5: Age-Adjusted Value of a Statistical Life



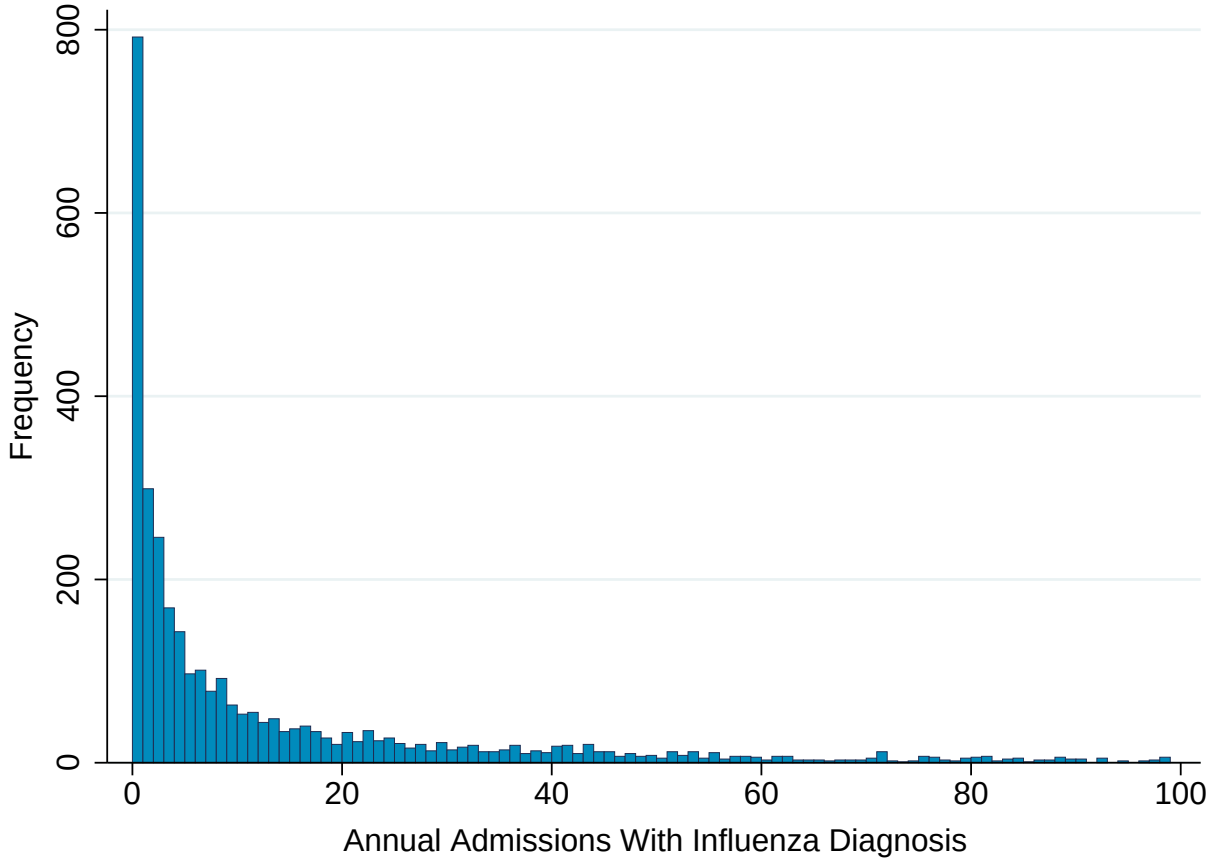
Note – The single-age VSL figures are derived from Murphy and Topel (2006), but adjusted to reflect the either the EPA's prime-age VSL figure of $8.8 million or the estimate from Ashenfelter and Greenstone (2004) of $2.31 million (denoted "AG"), and converted to 2016 dollars. To construct a VSL figure for each age group, I take a weighted average of these single-age VSL figures, where the weight is the share of deaths represented by each age. The implicit assumption is that deaths avoided because of influenza vaccination would have had the same age distribution as all-cause deaths within each age group.

Figure 6: California Mandates



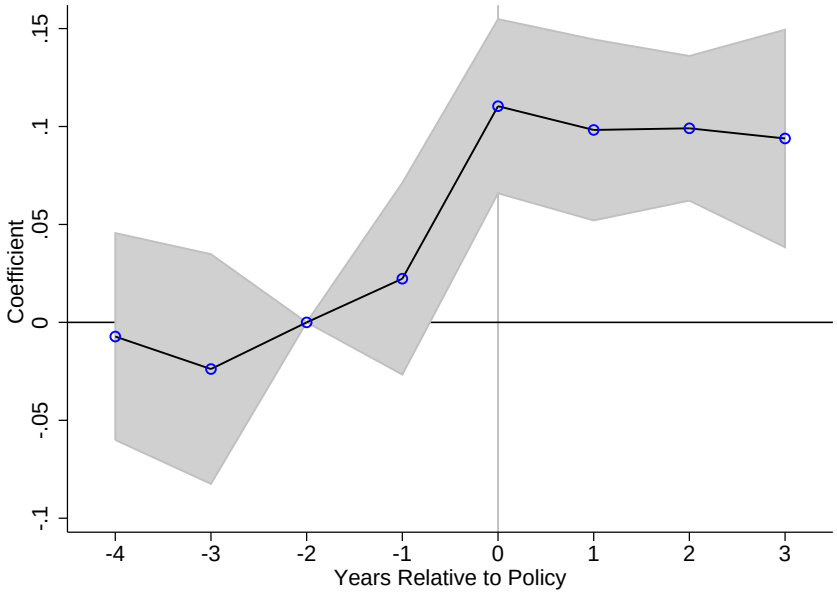
These plots display the roll-out of influenza vaccination mandates. Circles represent policies implemented at the hospital level and shaded regions represent policies implemented at the county level. The lighter shaded regions represent county-level policies that apply only to hospitals, and the darker regions represent county-level policies that apply more broadly. In two cases, county-level policies that applied only to hospitals and were subsequently replaced with policies applying more broadly.

Figure 7: Distribution of Influenza Admissions



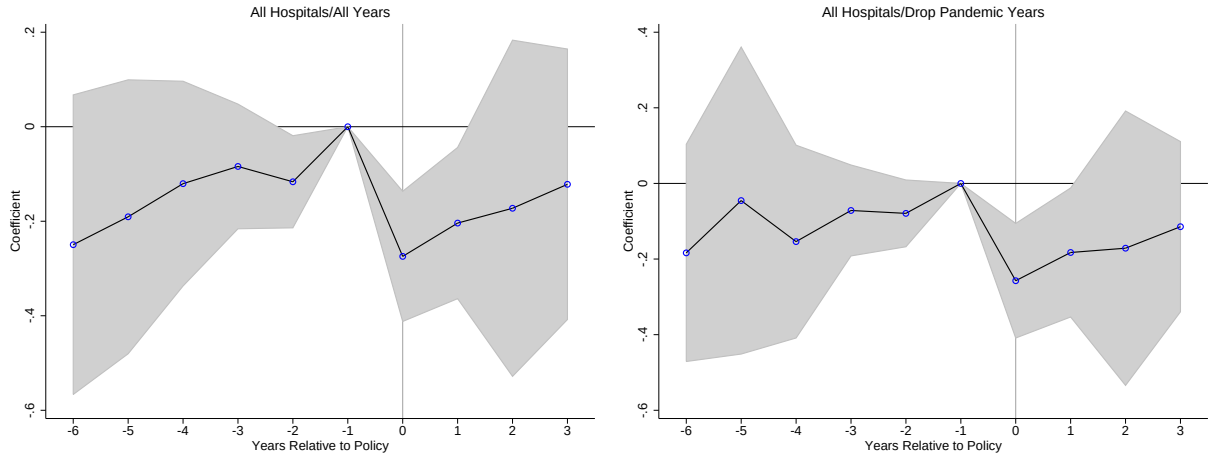
This plot displays the distribution of the number of admissions with an influenza diagnosis at the hospital-by-year level. The large number of hospital-years with zero observations is driven by a relatively small number of very small hospitals. The right tail is longer than implied by this plot, though there are very few observations with more than 100 influenza diagnoses; the maximum value is 403.

Figure 8: First Stage Event Study (HCW Vaccination Rates)



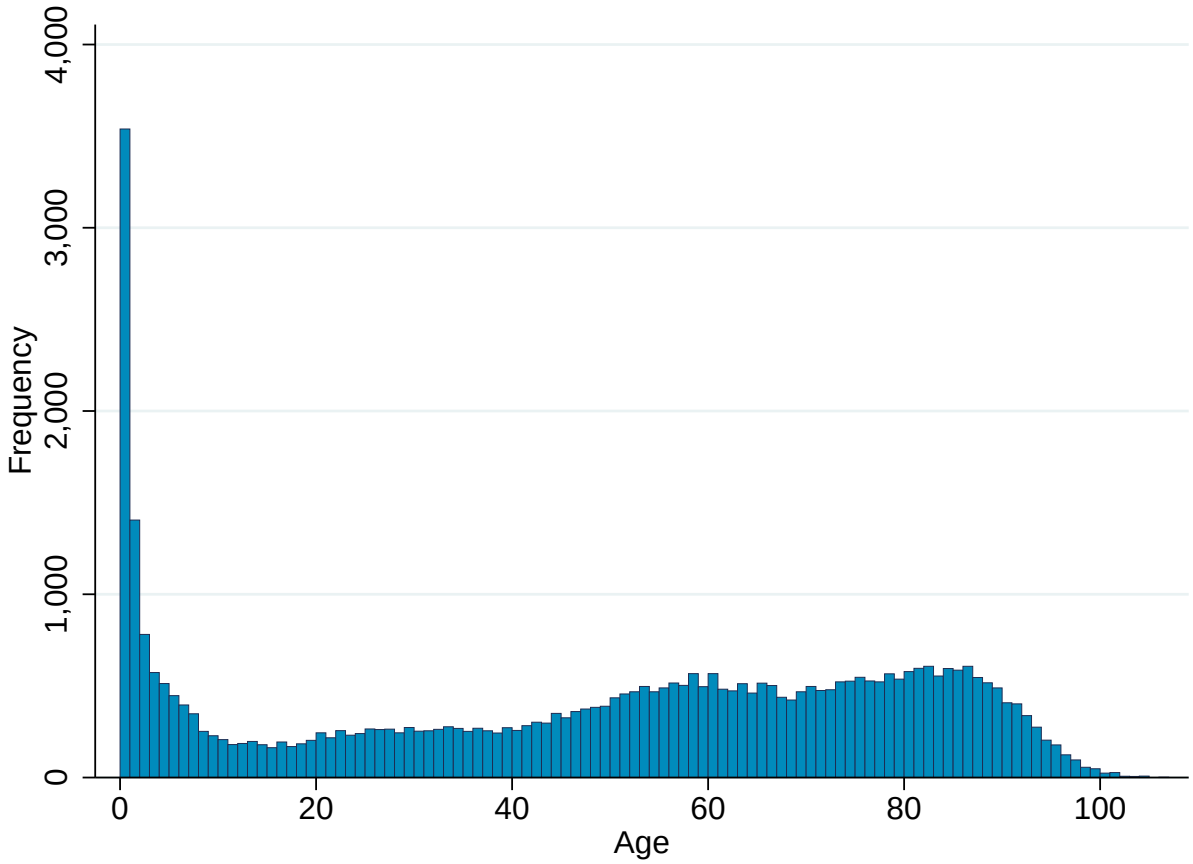
Points on this plot represent the point estimates from an event-study version of Equation (5) with HCW vaccination rates as the outcome. Shaded regions represent 95% confidence intervals around the point estimates. The event-study version is estimated by replacing the policy indicator (Required_{hym}) with a series of variables indicating years relative to the policy: $\sum_{j=-4}^{-3} \gamma_j \text{Required}_{hymj} + \sum_{j=-1}^{3} \gamma_j \text{Required}_{hymj}$. Note that the indicator representing two years prior to the policy is omitted as the reference group.

Figure 9: Reduced Form Event Study (Influenza Diagnoses)



Points on this plot represent the point estimates from an event-study version of Equation (5) with influenza-related admissions as the outcome. Shaded regions represent 95% confidence intervals around the point estimates. The event-study version is estimated by replacing the policy indicator (Required_{hym}) with a series of variables indicating years relative to the policy: $\sum_{j=-4}^{-2} \gamma_j \text{Required}_{hymj} + \sum_{j=0}^{T} \gamma_j \text{Required}_{hymj}$. Note that the indicator representing two years prior to the policy is omitted as the reference group and that the final time period (T) depends on whether hospital-level policies ($T = 4$) or county-level policies ($T = 2$) are being analyzed

Figure 10: Age Distribution of Influenza Admissions



This plot displays the single-year age distribution of influenza admissions. This emphasizes that infants are at the highest risk of hospitalization for influenza. Note that this includes all years of data including the 2009 H1N1 pandemic. The age distribution of the 2009 pandemic was somewhat younger than a typical year, accounting for a larger-than-typical number of admissions in the 20-60 range.

Table 1: Summary Statistics

Variable	Mean	(SE)	# Hospitals Affected	Percentage
National Data				
Vaccination Rate	34.45	(6.51)	-	-
Vaccination Rate ≥ 75	72.06	(7.39)	-	-
Vaccination Rate < 75	31.42	(6.52)	-	-
Match Rate	0.739	(0.287)	-	-
Influenza Activity (1993-2014)	0.188	(0.228)	-	-
Pneumonia/Influenza (PI) Mortality	6.21	(2.14)	-	-
R&C Mortality	36.24	(11.93)	-	-
All-Cause Mortality	71.84	(12.78)	-	-
Non-R&C Mortality	35.60	(9.48)	-	-
Illness Absence	0.025	(0.009)	-	-
Other Absence	0.093	(0.045)	-	-
Hours Worked	37.73	(1.317)	-	-
California Hospital Data				
HCW Vaccination Rate (No Mandate)	0.733	(0.120)	-	-
HCW Vaccination Rate (Mandate)	0.910	(0.061)	-	-
Influenza Activity (2007-2014)	0.254	(0.241)	-	-
Influenza Activity (Excluding H1N1)	0.238	(0.208)	-	-
Influenza Diagnoses (Annual)	16.00	(30.15)	-	-
PI Diagnoses (Annual)	437.22	(429.06)		
Average Length of Stay	5.584	(4.224)	-	-
Average Charges	31,650	(15,821)	-	-
Required 2009-10	-	-	13	0.030
Required 2010-11	-	-	18	0.040
Required 2011-12	-	-	44	0.099
Required 2012-13	-	-	121	0.273
Required 2013-14	-	-	252	0.563

Peak intensity is the average of the annual maximum values of influenza intensity. Peak intensity excluding 2009 is intended to measure the annual peak of *seasonal* influenza activity and is useful for calculating estimated policy effects using the triple difference model.

Table 2: Effects of the Match Rate on Vaccination Rates

	(1)	(2)
Match	-0.461∗	-0.358
	(0.184)	(1.890)
Match × Mean Vacc Rate		-0.003
		(0.055)
N	1,070	1,070

The outcome in these regressions is the vaccination rate. While the regressions are estimated at the state-month-year level (the same level as the analysis to come), neither match rates and vaccination rates vary by month in the data. The interaction with mean vaccination rates is intended to test whether high- and low-vaccination states respond differentially to match rates. Regressions include state fixed effects and a linear time trend. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table 3: Pneumonia/Influenza Mortality (Diff-in-Diff)

	(1)	(2)	(3)
Panel A: All Months			
Vacc × Match	0.002	-0.036**	-0.035**
	(0.023)	(0.010)	(0.013)
Vacc	0.109**	0.007	0.006
	(0.040)	(0.016)	(0.024)
N	12,534	12,534	11,012
Panel B: High Activity Months			
Vacc × Match	-0.135**	-0.161**	-0.139**
	(0.042)	(0.035)	(0.043)
Vacc	0.276**	0.109**	0.071
	(0.059)	(0.036)	(0.044)
N	1,835	1,835	1,376
Panel C: Low Activity Months			
Vacc × Match	0.010	-0.021*	-0.023*
	(0.022)	(0.009)	(0.011)
Vacc	0.091*	-0.006	0.003
	(0.038)	(0.015)	(0.025)
N	10,699	10,699	9,279
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
Mean Dep. Var.	6.21	6.21	6.21

"High Activity" is defined as months where the influenza index is at least 0.5, and "Low Activity" is months where the influenza index is less than 0.5. The "IV" specification indicates that the vaccination rate from three years prior is used as an instrument for the current year's vaccination rate. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table 4: Pneumonia/Influenza Mortality (Triple Difference)

	(1)	(2)	(3)
Vacc × Match × Activity	-0.192**	-0.196**	-0.183**
	(0.054)	(0.047)	(0.062)
Vacc × Match	0.017	-0.009	-0.007
	(0.019)	(0.008)	(0.011)
Vacc × Activity	0.287**	0.220**	0.198**
	(0.061)	(0.050)	(0.062)
Vacc	0.069+	-0.028+	-0.029
	(0.035)	(0.017)	(0.027)
Expected Annual Benefit	-0.302**	-0.309**	-0.288**
	(0.085)	(0.074)	(0.097)
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
Mean Dep. Var.	6.21	6.21	6.21
N	12,534	12,534	11,012

The "Expected Annual Benefit" is the coefficient on the triple interaction scaled by a factor of $\overline{Match} \times \sum_m \overline{Activity}_m$ (approximately 1.57) and is intended to measure the annual reduction in the mortality rate that would be expected to result from a one percentage point increase in the vaccination rate. The "IV" specification indicates that the vaccination rate from three years prior is used as an instrument for the current year's vaccination rate. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table 5: Mortality by Cause

	(1)	(2)	(3)
R&C Mortality (Mean = 36.24)			
D-D-D Effect	-0.173	-0.242*	-0.293**
	(0.131)	(0.102)	(0.096)
All-Cause Mortality (Mean = 71.84)			
D-D-D Effect	-0.071	-0.219	-0.269*
	(0.215)	(0.146)	(0.133)
Non R&C (Mean = 35.60)			
D-D-D Effect	0.102	0.022	0.024
	(0.125)	(0.086)	(0.081)
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

"R&C" refers to respiratory and circulatory deaths. Deaths are classified as R&C if any of the diagnosis codes are for a respiratory or circulatory illness (ICD9: 390-419; ICD10: "I" codes and "J" codes); note that pneumonia/influenza deaths are included among R&C deaths. Non-R&C is presented as a falsification test. All coefficients represent estimates of the triple-interaction (Vacc × Match × Activity) from the triple-difference specification. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table 6: Pneumonia/Influenza Mortality by Age

	(1)	(2)	(3)
Age Under 1			
D-D-D Effect	-0.003+	-0.002	-0.002
	(0.002)	(0.002)	(0.002)
	[2%]	[1%]	[1%]
Age 1-9			
D-D-D Effect	-0.001	-0.000	0.000
	(0.001)	(0.001)	(0.001)
	[1%]	[0%]	[0%]
Age 10-64			
D-D-D Effect	-0.006	-0.016+	-0.009
	(0.011)	(0.009)	(0.010)
	[5%]	[8%]	[5%]
Age Over 64			
D-D-D Effect	-0.182**	-0.178**	-0.173**
	(0.050)	(0.043)	(0.056)
	[93%]	[91%]	[94%]
Age Over 74			
D-D-D Effect	-0.178**	-0.165**	-0.175**
	(0.048)	(0.040)	(0.052)
	[90%]	[84%]	[95%]
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

Age-specific mortality rates are calculated as the number of deaths per 100,000 *total* individuals in the population (i.e., the denominator is not age-specific). As such, these estimates represent an accounting of the total mortality benefits of increased vaccination – the sum of the mutually exclusive age categories equals the total effect. The percentage of total benefits is reported in brackets (the age-specific coefficient here divided by the all-age coefficient in Table 4. I report additional estimates in Table A2 in which the denominator is age-specific. All coefficients represent estimates of the triple-interaction (Vacc × Match × Activity) from the triple-difference specification. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table 7: Pneumonia/Influenza Mortality – Externality Effects

	Age $\geq$ 75 Mortality		
	(1)	(2)	(3)
D-D-D Effect ($\geq$ 75 Vaccination)	-0.021	-0.041	-0.110*
	(0.035)	(0.028)	(0.049)
	[15%]	[27%]	[65%]
D-D-D Effect ($<$ 75 Vaccination)	-0.116*	-0.110*	-0.060
	(0.054)	(0.044)	(0.077)
	[85%]	[72%]	[35%]
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

Coefficients represent estimates of the triple-interaction terms (Vacc × Match × Activity) from a triple-difference specification; the specification represented by these estimates includes the full set of variables and interactions for two age-specific vaccination rates ($\geq$ 75 and <75). The estimate corresponding to the $\geq$ 75 group represents a combination of direct and partial externality effects from influenza vaccination; the coefficient corresponding to the $<$ 75 group is intended to measure a pure externality effect. The figure in brackets represents the percentage of the total effect that each individual effect accounts for. ** p<0.01; * p<0.05; + p<0.1

Table 8: Monetized Benefits of Mortality Reductions

		Scaled Nationally		
	Age-Adjusted VSL	Number of Deaths	Monetized Value	Value Per Vaccination
EPA VSL ($8.8 million 2016$)				
Age Under 10	$8,705,051	13.7	$119,259,198	$37.4
Age 10-64	$6,470,611	80.8	$522,825,368	$163.9
Age 65-74	$2,866,840	62.4	$178,890,816	$56.1
Age Over 74	$975,689	829.7	$809,529,163	$253.8
Total	-	986.6	$1,630,504,545	$511.1
AG VSL ($2.3 million 2016$)				
Age Under 10	$2,285,076	13.7	$31,305,542	$9.8
Age 10-64	$1,698,535	80.8	$137,241,675	$43.0
Age 65-74	$752,545	62.4	$46,958,842	$14.7
Age Over 74	$256,118	829.7	$212,501,490	$66.6
Total	-	986.6	$428,007,549	$134.11

Value of a Statistical Life (VSL) estimates are generated using the EPA's figure of $8.8 million or the estimate from Ashenfelter and Greenstone (2004) of $2.3 million (denoted "AG"), applied to the method of Murphy and Topel (2006) to calculate age-adjusted VSL figures for each age group. Estimates correspond to a one percentage point increase in the vaccination rate, and correspond to the model that includes state-by-month fixed effects but does not use the IV strategy.

Table 9: Labor Market Impacts

	(1)	(2)	(3)
Absent for Illness (Mean = 0.025)			
D-D-D Effect	-0.00064**	-0.00052*	-0.00048+
	(0.00024)	(0.00023)	(0.00025)
Hours Worked (Mean = 35.60)			
D-D-D Effect	0.0681	0.0402	0.0372
	(0.0423)	(0.0322)	(0.0387)
Absent for Other Reason (Mean = 0.093)			
D-D-D Effect	0.0013	0.0002	-0.0002
	(0.0008)	(0.0007)	(0.0009)
Value Per Vaccination	$60.50	$49.17	$45.38
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,636	12,636	11,462

The reported "Benefit Per Vaccination" represents a calculation based off of the effects on illness absences and is intended to measure the marginal social benefit of influenza vaccination on the margin of gains in work hours; this calculation is described in Section 3.3.5. All coefficients represent estimates of the triple-interaction (Vacc × Match × Activity) from the triple-difference specification. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table 10: Effects of HCW Mandates – First Stage and Influenza Diagnoses

	(1)	(2)	(3)
Vaccination Rates (First Stage)			
Required	0.0988**	-	-
	(0.0123)		
Influenza Diagnoses			
Required	-	-0.171**	-0.187*
		(0.0652)	(0.0760)
Drop H1N1 Seasons	-	-	X
N	1,524	3,280	2,460

Note that the mean vaccination rate among hospital workers with no vaccination mandate is approximately 73%. H1N1 seasons are the 2008-09 and 2009-10 seasons. These are dropped because the timing of vaccination relative to the epidemic is not clear. Standard errors are clustered at the county level. The smaller number of observations for the first stage is due to the fact that these data are available beginning in 2008, and that these data are not available for all hospitals, as discussed in Section 4.2. ** p<0.01; * p<0.05; + p<0.1

Table 11: Effects of HCW Mandates by Age – Influenza Diagnoses

	(1)	(2)
Age Under 1		
Required	-0.446**	-0.436**
	(0.120)	(0.140)
N	2,008	1,506
Age 1-10		
Required	-0.282**	-0.216**
	(0.0699)	(0.0784)
N	1,648	1,236
Age 10-64		
Required	-0.102	-0.111
	(0.0830)	(0.101)
N	3,272	2,454
Age Over 64		
Required	-0.104	-0.139+
	(0.0786)	(0.0790)
N	3,120	2,340
Drop H1N1 Seasons	-	X

H1N1 seasons are the 2008-09 and 2009-10 seasons. These are dropped because the timing of vaccination relative to the epidemic is not clear. Standard errors are clustered at the county level. Hospitals are excluded from each regression if there is not at least one age-specific admission in all time periods, accounting for the difference in sample sizes. ** p<0.01; * p<0.05; + p<0.1

Table 12: Effects of HCW Mandates – Other Outcomes

	ihs(PI Diagnoses)		ln(Length of Stay)		ln(Charges)		ihs(PI Mortality)	
Required×Activity	0.0340	0.0241	-0.0220+	-0.0321∗	-0.0116+	-0.0142∗	-0.0172	-0.0176
	(0.0498)	(0.0535)	(0.0124)	(0.0133)	(0.0059)	(0.0069)	(0.0590)	(0.0588)
Required	-0.007	-0.003	-0.0126	-0.0108	-0.0039	-0.0047	-0.0424	-0.0410
	(0.0187)	(0.0183)	(0.0167)	(0.0173)	(0.0079)	(0.0079)	(0.0451)	(0.0458)
Drop H1N1 Seasons	-	X	-	X	-	X	-	X
County-Level	-	-	-	-	-	-	X	X
N	38,304	28,728	34,080	25,560	38,296	28,720	6,874	5,499

Note that the inverse hyperbolic sine is used in place of the log specification for PI diagnoses and PI mortality due to a small number of zeroes in each measure. H1N1 seasons are the 2008-09 and 2009-10 seasons. These are dropped because the timing of vaccination relative to the epidemic is not clear. The smaller number of observations for the estimates of charges results from the fact that charges are not consistently reported by all hospitals; hospitals that do not report charges for at least 95% of admissions are dropped. Furthermore, length of stay is unreported for approximately 1% of admissions; averages cannot be calculated for these outcomes when all hospital-by-month outcomes are missing – this typically only occurs when there is a single observation in that cell. Standard errors are clustered at the county level. ** p<0.01; * p<0.05; + p<0.1

Table 13: Effects of HCW Mandates – Mechanisms

	(1)	(2)
Influenza Diagnoses (Present on Admission)		
Required	-0.170**	-0.186*
	(0.0647)	(0.0754)
Influenza Diagnoses (Not Present on Admission)		
Required	-0.358*	-0.346+
	(0.149)	(0.193)
Drop H1N1 Seasons	-	X
N	3,280	2,460

H1N1 seasons are the 2008-09 and 2009-10 seasons. These are dropped because the timing of vaccination relative to the epidemic is not clear. Standard errors are clustered at the county level. ** p<0.01; * p<0.05; + p<0.1

Table A1: PI Mortality – Exclude Interpolated Vaccination Rates

	(1)	(2)	(3)
Vacc × Match × Activity	-0.220**	-0.219**	-0.195**
	(0.059)	(0.048)	(0.068)
Vacc × Match	0.024	-0.013	-0.011
	(0.022)	(0.008)	(0.011)
Vacc × Activity	0.293**	0.234**	0.202**
	(0.063)	(0.050)	(0.064)
Vacc	0.078*	-0.026	-0.032
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
Mean Dep. Var.	6.21	6.21	6.21
N	10,398	10,398	9,788

These estimates reproduce the estimates reported in Table 4, but drop the years in which interpolated vaccination rates were employed. The "IV" specification indicates that the vaccination rate from three years prior is used as an instrument for the current year's vaccination rate. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A2: PI Mortality by Age – Age Specific Rates (Triple Diff)

	(1)	(2)	(3)
Age Under 1			
	-0.205+	-0.176	-0.149
	(0.116)	(0.118)	(0.135)
Age 1-9			
	-0.005	-0.002	0.004
	(0.011)	(0.011)	(0.014)
Age 10-64			
	-0.015	-0.026*	-0.017
	(0.014)	(0.012)	(0.013)
Age Over 64			
	-1.082**	-1.158**	-0.968*
	(0.350)	(0.309)	(0.417)
Age Over 74			
	-1.744**	-1.886**	-1.781*
	(0.629)	(0.557)	(0.771)
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

The estimates reported here differ from the estimates reported in the main analysis in that the age-specific mortality rates are calculated using age-specific population counts as the denominator. The advantage of this approach is that the estimates take into account differences in the size of the group affected, and thus indicates whether certain groups are affected *relatively* more than others. All coefficients represent estimates of the triple-interaction (Vacc × Match × Activity) from the triple-difference specification. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A3: PI Mortality Specification Checks (Triple Difference)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Vacc × Match × Activity	-0.140+	-0.192**	-0.183*	-0.169**	-0.150*	-0.178**	-0.165**
	(0.079)	(0.051)	(0.071)	(0.048)	(0.069)	(0.047)	(0.062)
Vacc × Match	-0.016	-0.015	-0.010	-	-	-	-
	(0.017)	(0.009)	(0.011)				
Vacc × Activity	0.234**	0.273**	0.260**	0.255**	0.237**	0.202**	0.177**
	(0.085)	(0.059)	(0.077)	(0.057)	(0.076)	(0.048)	(0.060)
Vacc	0.110**	-0.037*	-0.041	-	-	-	-
	(0.039)	(0.018)	(0.028)				
Month-Year Fixed Effects	X	X	X	X	X	X	X
State Fixed Effects	-	X	X	-	-	-	-
State-Year Fixed Effects	-	-	-	X	X	X	X
State-Month Fixed Effects	-	-	-	-	-	X	X
IV (t-3)	X	-	X	-	X	-	X
N	11,012	12,534	11,012	12,534	11,012	12,534	11,012

Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A4: PI Mortality - Regional Triple Diff

	Exclude Pre-1998			Regional Variation in Match		
	(1)	(2)	(3)	(4)	(5)	(6)
Vacc × Match × Activity	-0.202**	-0.199**	-0.188*	-0.214**	-0.200**	-0.185*
	(0.059)	(0.051)	(0.080)	(0.060)	(0.052)	(0.084)
Vacc × Match	0.030	-0.015	-0.009	0.017	-0.015	-0.009
	(0.026)	(0.010)	(0.013)	(0.023)	(0.010)	(0.013)
Vacc × Activity	0.262**	0.214**	0.196**	0.277**	0.220**	0.199*
	(0.063)	(0.055)	(0.074)	(0.065)	(0.055)	(0.079)
Vacc	0.091*	-0.029+	-0.040	0.100**	-0.029*	-0.039
	(0.036)	(0.014)	(0.028)	(0.036)	(0.014)	(0.029)
Match × Activity	-	-	-	8.282*	8.699**	7.736+
				(3.236)	(2.426)	(3.969)
Match	-	-	-	1.394	0.457	0.113
				(1.370)	(0.666)	(0.640)
Expected Annual Benefit	-0.317	-0.313	-0.295	-0.337	-0.314	-0.291
Mean Dep. Var.	6.02	6.02	6.02	6.02	6.02	6.02
Month-Year Fixed Effects	X	X	X	X	X	X
State-Month Fixed Effects	-	X	X	-	X	X
IV	-	-	X	-	-	X
N	9,486	9,486	7,650	9,486	9,486	7,650

The first three columns represent estimates from regressions similar to Table 4, except that they exclude years prior to 1998. This is included for comparison with the regressions reported in the following columns that utilize data on match rates that are region-specific; this data is only available in years prior to 1998. Note that the Match × Activity and Match variables need to be included in these regressions are they are no longer absorbed by the month-year fixed effects. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A5: PI Mortality - Leave One Out Triple Diff

	(1)	(2)	(3)
Vacc × Match × Activity	-0.182**	-0.185**	-0.180**
	(0.060)	(0.050)	(0.064)
Vacc × Match	0.014	-0.013	-0.011
	(0.018)	(0.008)	(0.011)
Vacc × Activity	0.278**	0.217**	0.203**
	(0.065)	(0.052)	(0.064)
Vacc	0.074*	-0.025	-0.026
	(0.034)	(0.017)	(0.027)
Match × Activity	11.07*	14.92**	15.16**
	(4.59)	(3.32)	(3.36)
Activity	-23.85**	-22.58**	-22.55**
	(5.47)	(3.02)	(3.08)
Mean Dep. Var.	6.21	6.21	6.21
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

Influenza activity in these regressions is defined as the average measure of influenza activity across all regions except the region of analysis. The purpose is to ensure that the measure of influenza activity is not influenced by local vaccination behavior. Because activity varies by state in these estimates, the interaction of the match rate with influenza activity (Match × Activity) and the main effect for influenza activity (Activity) are not absorbed by the time period fixed effects. Caution should be exercised in interpreting these estimates, however, as they represent out of sample predictions (i.e., zero vaccination rate). Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A6: PI Mortality - Monthly Triple Diff

	(1)	(2)	(3)
Vacc × Match × High	-0.038+	-0.041∗	-0.044+
	(0.021)	(0.018)	(0.023)
Vacc × Match	0.006	-0.026∗	-0.022+
	(0.020)	(0.010)	(0.012)
Vacc × High	0.082∗∗	0.027	0.014
	(0.026)	(0.016)	(0.021)
Vacc	0.088∗	-0.000	0.004
	(0.033)	(0.016)	(0.026)
Scaled DDD Effect	-0.131	-0.141	-0.151
Mean Dep. Var.	6.21	6.21	6.21
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

"High" Refers to months December through March; these are the months during which influenza circulation is typically highest. The average values of the influenza activity measure during during "High" and "Low" months are 0.38 and 0.09, respectively. Because the main estimates represent the difference in Vacc × Match between periods of maximum influenza activity (Activity=1) and zero influenza activity, the estimates presented here are not directly comparable. "Scaled DDD Effect" is reported for comparison with the main estimates in Table 4, and is equal to the reported DDD coefficient scaled by a factor of 1/(0.38-0.09). Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A7: PI Mortality - Lagged Impacts

	Add One Lag			Add Two Lags		
	(1)	(2)	(3)	(4)	(5)	(6)
DDD Effect	-0.269∗∗	-0.243∗∗	-0.231∗∗	-0.244∗∗	-0.189∗∗	-0.176∗∗
	(0.065)	(0.055)	(0.073)	(0.063)	(0.049)	(0.062)
Month-Year Fixed Effects	X	X	X	X	X	X
State-Month Fixed Effects	-	X	X	-	X	X
IV	-	-	X	-	-	X
N	12,434	12,434	11,012	12,434	12,434	11,012

Estimates test whether the contemporaneous month is sufficient to capture the full extent of influenza-related mortality. "Add One Lag" estimates replicate the estimates in Table 4, but include a one month lag in the interactions that include influenza activity. The reported coefficients are the sum of the contemporaneous and lagged impact. "Add Two Lags" extends this to include an additional month. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A8: PI Mortality by Age – Alternate Categories

	(1)	(2)	(3)
Age Under 10			
D-D-D Effect	-0.0035∗	-0.0027	-0.0016
	(0.0016)	(0.0016)	(0.0014)
	[-17.78]	[-13.65]	[-8.20]
Age 10-64			
D-D-D Effect	-0.006	-0.016+	-0.009
	(0.011)	(0.009)	(0.010)
	[-31.95]	[-80.79]	[-46.39]
Age 65-74			
D-D-D Effect	-0.004	-0.012	0.003
	(0.013)	(0.014)	(0.012)
	[-22.23]	[-62.43]	[14.08]
Age Over 74			
D-D-D Effect	-0.178∗∗	-0.165∗∗	-0.175∗∗
	(0.048)	(0.040)	(0.052)
	[-892.47]	[-829.66]	[-879.29]
Month-Year Fixed Effects	X	X	X
State-Month Fixed Effects	-	X	X
IV	-	-	X
N	12,534	12,534	11,012

This table presents age-specific mortality impacts corresponding to the age groups used in the calculation of monetary benefits. In brackets is the estimated number of deaths avoided due to a national policy increasing the influenza vaccination rate by one percentage points. All coefficients represent estimates of the triple-interaction (Vacc × Match × Activity) from the triple-difference specification. Standard errors are clustered at the state level. ** p<0.01; * p<0.05; + p<0.1

Table A9: Policy Timing

Hospital	Season	County	Season
Children's of Orange	2009 (H1N1)	Sacramento	2011-12
Community Hospital of LB	2009 (H1N1)	San Luis Obispo	2011-12
Hoag Hospitals	2009 (H1N1)	San Francisco	2011-12
Long Beach Memorial	2009 (H1N1)	Alameda	2012-13
Miller Children's	2009 (H1N1)	Amador	2012-13
Orange Coast Memorial	2009 (H1N1)	Contra Costa	2012-13
Pacific Hospital of LB	2009 (H1N1)	El Dorado	2012-13
St. Joseph (Orange)	2009 (H1N1)	Mono	2012-13
St. Jude (Fullerton)	2009 (H1N1)	Nevada	2012-13
UC Davis	2009 (H1N1)	San Joaquin	2012-13
UC Irvine	2009 (H1N1)	Santa Clara	2012-13
UC San Diego	2009 (H1N1)	Santa Cruz	2012-13
Saddleback Memorial	2009 (H1N1)	Sonoma	2012-13
Santa Rosa Memorial	2010-11	Tehama	2012-13
Sierra Vista (SLO)	2010-11	Yolo	2012-13
Tri-City (Oceanside)	2010-11	Los Angeles	2013-14
Petaluma Valley Hospital	2010-11	Marin	2013-14
Oroville Hospital	2012-13	Monterey	2013-14
Banner Lassen Medical Center	2012-13	Napa	2013-14
Barton Memorial	2012-13	San Benito	2013-14
UCSF (Children's - Oakland)	2012-13	Shasta	2013-14
Cottage Hospitals	2013-14	Trinity	2013-14
Salinas Valley Hospital	2013-14	-	-

Note: Hospitals that implemented their mandates in 2009, labelled "2009 (H1N1)", did so in response to the H1N1 pandemic. All other mandates were implemented prior to the beginning of an influenza season.

Table A10: Effects of HCW Mandates – Specification Checks

	(1)	(2)
Add Time Trends		
Required	-0.200∗	-0.200∗
	(0.0929)	(0.0901)
N	3,280	2,460
Fixed Effects Poisson		
Required	-0.156∗∗	-0.168∗∗
	(0.0541)	(0.0578)
N	3,000	2,238
Zero-Inflated Negative Binomial		
Required	-0.199∗∗	-0.189∗
	(0.0630)	(0.0762)
N	3,280	2,460
Drop H1N1 Seasons	-	X

The three panels represent different specification checks. The first panel adds hospital-specific linear time trends. The second panel uses a fixed-effects Poisson estimator. The third panel uses a zero-inflated negative binomial. The zero-inflated model is a two step model that allows different data generating processes to predict first whether a positive count will be observed and second the count (conditional on it being positive). The first step is predicted by only month-year fixed effects while the second uses the same regressors as the main analysis. Standard errors are clustered at the county level except in the case of the fixed effects Poisson, where the standard errors must be clustered at the same level as the fixed effect (i.e., hospital). ** p<0.01; * p<0.05; + p<0.1