# Measuring the Social and Externality Benefits of Influenza Vaccination [*]

Corey White

California Polytechnic State University, San Luis Obispo

cwhite46@calpoly.edu

This Version: March 28, 2018

## 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 annual variation in vaccine quality, I estimate 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 1,134 fewer deaths and 8.9 million fewer work hours lost due to illness each year. The mortality reductions are concentrated among individuals 75 and older, but over half of the effect is attributable to 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 extremely high externality benefits: vaccination mandates for health care workers. I find that mandates lead to reductions in hospital diagnoses for influenza in affected counties, consistent with substantial externality impacts. For both the general population and the population of health care workers, the estimates suggest that programs increasing vaccine take-up are likely to be cost-effective under reasonable assumptions about the costs.

JEL Codes: I12, I18, D62, H23

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

# 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 maintaining immunity 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 large. I analyze the impacts of county-level 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 efficacy of the vaccine. Vaccine efficacy 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, 2017).

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

I find that higher 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 1,134 fewer deaths per year in expectation. The mortality benefits primarily accrue to individuals 75 and older, but 66% of the benefits are attributable to the vaccination of people under 75, suggesting substantial externality benefits. I also find that influenza vaccination significantly reduces the probability of illness-related work absences and hours lost due to illness. The estimates indicate that a one percentage point increase in the U.S. vaccination rate would result in approximately 9.2 million fewer work hours lost due to illness annually, in expectation. 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.

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 at least $74 in benefits 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 significant for the 75 and older population. On the dimension of illness absences, I find that each vaccination confers benefits of approximately $57.

Because the first component of the analysis exploits existing state-level 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 large potential externalities 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., long-term care facilities). I find that these mandates increase hospital worker vaccination rates by 10.6 percentage points, reduce the number of influenza diagnoses for inpatient visits by 21.4%, and reduce the number of influenza diagnoses for outpatient emergency department visits by 9.4% during seasons with an effective vaccine. For inpatient visits, the impact is twice as

2

large for influenza diagnoses that were not present at the time of admission (i.e., hospital-acquired infection). I estimate the marginal benefit of HCW vaccination in terms of health care cost savings to be $143 per vaccination.

An exercise comparing the two components of the analysis suggests that health care worker vaccination is approximately seven times more effective at reducing the spread of influenza in comparison to vaccination in the general population.[3] 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 that would increase vaccination at reasonable cost in either population are likely to be cost-effective.

The primary contribution of this paper is to provide causal estimates of the social and externality benefits of influenza vaccination. While a large medical literature evaluates the benefits of influenza vaccination, much of the existing evidence on these benefits is derived from RCTs in which vaccination is randomized across individuals within a group, leaving no method for capturing externality effects.[4] 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 effects of vaccination on severe and economically important 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

---

[3]This figure should be taken with caution as it is not possible to use the same identification strategy and outcomes for both components of the analysis, and thus the calculation relies on a number of fairly strong assumptions.

[4]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).

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) is a notable exception. Ward (2014) uses exogenous variation in vaccine efficacy to evaluate the impacts of a regional influenza vaccination campaign in Ontario, Canada. The author finds that the program increased vaccination rates for non-elderly adults by approximately 10.8 percentage points (the post-treatment vaccination rate 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), but 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 efficacy. My strategy has the advantage of exploiting variation in vaccination rates and outcomes across 51 states and 22 influenza seasons. As such, the average impacts that I estimate are not unduly influenced by the experience in any one region or time period. Compared to many other studies of the social benefits of influenza vaccination, the estimates presented here are smaller in magnitude. That being said, the magnitude of the estimates presented here are plausible and still suggest that the social and externality benefits of vaccination are substantial.

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 small number of studies that assess the impacts of vaccination requirements primarily for employees of long-term care facilities. One recent paper critiques the findings of four recent studies by noting the implausibility of estimates based on non-specific outcomes (De Serres et al., 2017). Another meta-analysis rates the overall quality of evidence on the subject as either "low" or "very low" (Thomas et al., 2016). This study has the advantages of using a highly specific outcome (hospital diagnoses for influenza), extremely large scale (over 75% of hospitals in California were subject to the mandates at the end of the sample), and the ability to identify impacts in settings other than long-term care facilities.

4

While the specific contributions of this paper are described above, this paper also contributes to an empirical literature within economics that seeks to identify the economic impacts of influenza infection more generally. Much of this literature has focused on the effects of in-utero exposure to influenza on human capital development. Within this literature, most research has examined *pandemic* influenza (Almond and Mazumder, 2005; Almond, 2006; Kelly, 2011; Lin and Liu, 2014; Brown and Thomas, 2016), though exposure to *seasonal* influenza (the focus of this study) has been found to negatively impact health at birth and later-life outcomes as well (Currie and Schwandt, 2013; Schwandt, 2017). There are fewer studies that focus on the more contemporaneous impacts of influenza in the adult population, with notable exceptions studying vaccination (Ward, 2014) and other factors that can influence the spread of influenza (Adda, 2016; Stoecker et al., 2016).

This paper also contributes to a recent set of empirically-focused papers within economics that study various questions related to vaccination. For example, Carpenter and Lawler (2017) find that state requirements that children receive a tetanus, diphtheria and pertussis (Tdap) booster vaccine prior to middle school entry result in higher vaccination rates for Tdap, and lower rates of pertussis among the treated age group (private benefits) and other age groups (externality benefits). They also find evidence of behavioral spillovers in that the Tdap requirement increased vaccination for other diseases. Lawler (2017) finds that both non-binding recommendations and mandates for Hepatitis A vaccination are effective at increasing vaccination rates and decreasing morbidity for the disease in the aggregate (social benefits). Oster (2018) examines the vaccination-disease relationship in the opposite direction, and finds that pertussis outbreaks in a county decrease the share of unvaccinated children entering kindergarten in subsequent years.

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 prominent in environmental economics, where many papers have sought to measure the impacts of pollution on a variety of outcomes (see Graff Zivin and Neidell (2013) for a review). 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 (exceptions above) 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 (Part

5

I) describes the analysis of aggregate vaccination rates in the general population, and Section 4 (Part II) describes the analysis of health care worker mandates in California. Finally, Section 5 offers a discussion 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 benefit curves 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 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. I focus on age as the primary dimension of heterogeneity. This discussion reflects the findings 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 in 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 the 65 and older population 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 on 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 vaccine efficacy – the extent to which vaccination protects against laboratory-confirmed influenza – is determined by several factors. Vaccine match is an especially important factor, but it is important to note that even when the vaccine is perfectly matched it is not 100% effective. Vaccine efficacy also varies with age; diminished immune response among the elderly means that they are less able to create the antibodies needed to gain immunity. Estimates of vaccine efficacy 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 efficacy in the elderly population are more contentious, 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 efficacy 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 efficacy; studies of vaccine efficacy 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 convenes a meeting in order to make recommendations on the composition of the following season's 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.[5] 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

---

[5]The quadrivalent vaccine was introduced in 2012.

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 becomes a dominant strain in a given season. Vaccine mis-matches are unpredictable prior to the start of influenza season. When a mis-match occurs, the non-matched virus strain is typically included in the following year's vaccine. This means that matched and mis-matched strains do not represent a different set of virus strains, since mis-matched strain become a matched strains in following seasons. It is also important to note that not only is the vaccine formulated well before influenza season begins, but individuals typically have no information on 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 and marginal externality benefits of influenza vaccination. Before moving on to the estimation, it is useful to consider a simple economic framework of externalities in the specific case of influenza. In this framework, there is a marginal private benefit of vaccination (MPB) and a marginal social benefit of vaccination (MSB). The MSB is assumed to be 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 is 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 benefit curves in the specific context of influenza provides additional insight.

Boulier et al. (2007) combine basic externality theory with a workhorse model of disease dynamics (the susceptible-infected-removed model) and parameterize the model to the case of influenza in order to derive theoretical predictions for the shape of the marginal benefit curves. Figure 1 produces a version of their result, allowing the MPB and MSB to depend on vaccine efficacy. I have plotted each assuming either 100% or 50% efficacy (denoted $E$). Recall that even when the vaccine is well-matched, estimates of vaccine efficacy 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. 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 external 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 ("herd immunity"). 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 approximately 60%.

I caution that this model depends on a number of parameter choices that are difficult to estimate accurately, but considering the general shape implied by the model helps to guide the interpretation of the results to follow.[6] Importantly, the model predicts that the externality makes up the majority of the social benefits of vaccination. The model also predicts relatively constant marginal social benefits of vaccination below the threshold, implying 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 of course 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.

Discussion of this framework also presents the opportunity to discuss potential heterogeneous impacts of vaccination (though heterogeneity is not explicitly built into the model). It is worth considering how two groups in particular may differ from the remainder of the population: the elderly population and health care workers. For the elderly, the cost of infection is high and vaccine efficacy is relatively low. These factors combined imply that the elderly benefit substantially from the vaccination of others and that a particularly large portion of the benefits to the elderly will operate through an externality. Health care workers (HCWs) are particularly interesting for two reasons. First, HCWs come in relatively frequent contact with infected individuals and thus vaccination is more likely to prevent infection in this group. Second, HCWs come in relatively frequent contact with individuals who have a high cost of infection (e.g., individuals with a compromised immune system), and thus the vaccination of HCWs may reduce the spread of infection precisely to those who would suffer

---

[6]Important parameters include vaccine efficacy and the "contact number", which is the number of additional infections that result from a single infection when the entire population is susceptible.

the most severe consequences. These potential heterogeneous impacts motivate the focus on these groups in the empirical analyses to follow.

# 3    Part I: Aggregate Vaccination Rates

## 3.1    Data

This analysis requires data on mortality by cause of death, illness-related work absences, influenza vaccination rates, the timing and magnitude of influenza activity, and the vaccine match rate. The unit of analysis is the state-year-month and the data cover the years 1994-2016. Because the analysis centers around influenza seasons, years will be redefined to represent "flu-years", defined as running from July through June so that each flu-year represents a distinct influenza season. Note that I often use "flu-years" and "influenza seasons" interchangeably. The data coverage is ultimately July 1994 through June 2016. In most specifications, the two flu-years affected by the 2009 H1N1 pandemic are omitted (2008/09 and 2009/10) so that the estimates represent seasonal influenza. Summary statistics for all data described here are provided in Table 1.

### 3.1.1    Outcomes: Mortality and Work Absences

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 includes state identifiers beyond 2005. It is important to note the use of multiple causes of death in classifying mortality as influenza-related. 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, deaths are classified 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 used in the analysis of mortality is deaths with any diagnosis for pneumonia/influenza (PI). Because 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

10

infection, these non-respiratory/circulatory deaths are used as a falsification test.[7] Finally, estimates are also presented that use only the primary cause of death instead of multiple causes for each of 34 mutually exclusive cause-of-death categories.

Data on illness absences and hours absent due to illness are derived from the Current Population Survey (CPS) basic monthly files. Similar to Stearns and White (2018), 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. Second, individuals who are employed and at work during the reference week report both their usual hours worked 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 usual. 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). For each worker reporting an absence, the survey also asks the number of hours missed. The mean number of hours missed due to illness (at the state-year-month level) is the main outcome of interest for work absences. In addition to illness-related absences, absences for other reasons are analyzed as falsification tests.

All measures of absence can be constructed only for individuals who work at least 35 hours per week and thus represent only full-time workers. A standard set of individual covariates are regressed out at the individual level prior to collapsing the residuals to the state-year-month level; covariates are 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).

### 3.1.2 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.[8] The BRFSS asks whether each participant has received an influenza vaccination within the past 12 months.[9] Due to the phrasing of the question, there are several months wherein the season to which the vaccine applies is ambiguous. I drop

---

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

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

[9]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.

all survey responses taken from September through December since these are the months in which the vast majority of influenza vaccinations are received and as such are the months with substantial ambiguity as to which season the reported vaccination applies. 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 was not collected in the survey years 1994, 1996, 1998 or 2000. To fill in these gaps, vaccination rates are linearly interpolated in missing flu-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.[10]

### 3.1.3 Match Rates

Data on the vaccine match are derived from annual influenza season summaries, which consist of data compiled from the CDC's virologic surveillance system.[11] This system consists of laboratories located throughout the country that test respiratory specimens for the presence of any influenza virus and characterize viruses according to the exact strain. The 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, 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.[12]

---

[10]Furthermore, the estimates are not sensitive to the exclusion of these years.

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

[12]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 (along with all input data) is available upon request.

### 3.1.4 Influenza Activity & Other Data

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.

Population data is required to construct mortality rates. Population by state, year and age are derived from the U.S. Census Bureau. Controls for temperature, humidity, and precipitation are included and derived from the Global Summary of the Day files.

## 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:

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

$Y_{smy}$ measures the outcome in state $s$, flu-year $y$, and month $m$. The outcomes are either the PI mortality rate (per 100,000) or the average number of work hours lost due to illness. $V_{sy}$ is the vaccination rate for state $s$ and flu-year $y$, and $\delta_{my}$ are month-year fixed effects. $W_{smy}$ represents flexible controls for temperature, humidity and precipitation.[13] The benefit of relating state-level vaccination rates to outcomes at the state-level is that all within-state externalities are captured. The issue of course is that 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 get vaccinated. Indeed the estimate of $\beta$ from the naive regression described above is 0.11 (s.e. = 0.031); taken causally, this would imply that higher vaccination rates lead to more influenza-related mortality. Controlling for state-specific factors may ameliorate this endogeneity issue: the inclusion of state-by-month fixed effects in the above equation produces a coefficient estimate of -0.03 (s.e. = 0.014). That being said, it is reasonable to argue that if fixed differences in vaccination rates across states are endogenous, then changes in vaccination rates are likely to be endogenous as well. This would be the case if, for example, states with rapidly increasing vaccination rates are more (or less) likely to pass effective health policy, or have individuals that are more (or less) likely

---

[13]Temperature is expressed as the number of days in one of seven 10-degree mean temperature bins from <20F to >80F (60-70 omitted). Humidity is expressed as the number of days in three 5g/kg specific humidity bins from <5 to >20 (<5 omitted). Precipitation is expressed as a cubic in total monthly precipitation.

to adopt other beneficial health behaviors. Such possibilities are far from implausible, and as such the strategy described below does not rely on potentially endogenous fixed differences or changes in vaccination rates for its identifying variation.

The strategy I use is to interact differences in vaccination rates with year-to-year variation in vaccine efficacy to generate plausibly exogenous variation in *effective* vaccination, while controlling for the actual vaccination rate. This strategy is illustrated in Figure 2. The top panel plots vaccination rates for two groups of states over time: a group of low-vaccination states (the five states with the lowest mean vaccination rate over time) and a group of high-vaccination states (the five states with the highest mean vaccination rate). The figure also plots the match rate in each influenza season. Note that vaccination rates in both groups evolve relatively smoothly over time, and that the gap in vaccination rates evolves smoothly over time as well. Further note that there is no visual evidence indicating that vaccination rates are systematically different during high or low match seasons. The bottom panel plots the *effective* vaccination rate over time (i.e., the product of the actual vaccination rate and the match rate). During seasons in which the match is close to one, the gap in effective vaccination between low- and high-vaccination states is preserved. During seasons in which the match is poor, however, there is little difference between these states in effective vaccination. The identification strategy compares the difference the outcome between low- and high-vaccination states in flu-years with a highly effective vaccine against the same difference in flu-years with a relatively ineffective vaccine. Equation (2) describes this difference-in-differences approach:

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

In this equation, $M_y$ is the match-rate, measured nationally for each flu-year. A match rate of zero implies that the vaccine is minimally effective whereas a match rate of one implies maximum efficacy.[14] $\gamma_2$ represents the potentially endogenous component of the relationship between vaccination rates and the outcome of interest: this measures the relationship between vaccination rates and the outcome in seasons in which the vaccine is minimally effective (i.e., zero match rate). $\gamma_1$ is the object of interest, and represents the differential effect of an increased vaccination rate between flu-years when the vaccine is at maximum versus minimum efficacy. Intuitively, $\gamma_1$ picks up the impact of *effective* vaccination (i.e., the causal effect of vaccination), but not the component of the relationship between vaccination rates

---

[14]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 efficacy does not imply that the vaccine is perfectly effective. In fact, maximum efficacy is approximately 60%.

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. One potential concern is whether it is possible for individuals to respond to the match rate in terms of vaccination behavior (e.g., choose not to receive a vaccine if the match is poor). Such behavior would introduce bias if there are differential responses 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. Also supporting this notion is the fact that I find no evidence of serial correlation in the match rate, and no evidence of a trend in the match rate over time.[15] Furthermore, evidence described below suggests that there is limited scope for individuals to respond to the vaccine match.

From 2007 onward, the BRFSS has asked respondents not only whether they received a vaccination, but the month in which they were vaccinated. Figure A1 plots the average cumulative vaccination rate (across flu-years) by month for the flu-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 administered 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 limited scope for responding to the match rate at all, much less differentially across states. That being said, while there is little scope for responding to match rates, it is not impossible.

Table 2 provides more direct evidence on the question of whether individuals are responding to match rates in terms of vaccination behavior. Column one reports estimates from a regression of the vaccination rate in a given state and influenza season on the match rate. The estimate is small and statistically indistinguishable from zero at conventional levels. The estimate implies that vaccination rates are 0.31 percentage points higher in 100% match seasons compared to 0% match seasons. The 95% confidence interval rules out an effect size larger than 0.69 percentage points. In a more direct test of the identifying assumption, Column 2 presents estimates of a test for differential responses to vaccine match by interacting the match rate with the mean vaccination rate (over time) for each state; the results indicate no evidence of a differential response among states that tend to have higher or lower vaccination rates. To the extent that any concerns remain over the possibility of responses

---

[15]Both serial correlation and trends are tested at the flu-year level (the level at which the match rate is defined). A regression of $M_y$ on $M_{y-1}$ yields a coefficient estimate of -0.051 (s.e. = 0.263), and a regression of $M_y$ on a time trend (flu-year) yields a coefficient estimate of 0.004 (s.e. = 0.010).

to the match rate in terms of vaccination behavior, additional specifications are estimated wherein only *pre-existing* variation in vaccination rates is employed. More specifically, the vaccination rate and each of its interactions are instrumented using vaccination rates from prior seasons. The estimates are insensitive to the choice of prior season, including the average vaccination rate over the prior three seasons, or a time-invariant vaccination rate defined as the average vaccination rate in the first three seasons of the sample.

Further supporting a causal interpretation of the estimates is their robustness to a variety of specifications and falsification tests, described in detail in Section 3.3. One falsification test uses the idea that influenza vaccination will only have a causal effect on mortality and work absences during periods in which influenza is circulating. This idea is explicitly built into the main estimation strategy (described below), in which within-year variation in the timing and magnitude of influenza activity is exploited as a third source of variation. This can be done either by estimating Equation (2) separately for periods of high and low influenza activity, or more formally in a triple-difference approach. While both versions are presented, the advantage of the triple-difference approach is that it allows for the use of a continuous measure of influenza activity, which more precisely focuses on the periods in which the largest impacts would be expected relative to a more arbitrary classification of months into "high" and "low" influenza activity.

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

Influenza activity, $A_{my}$, can be defined in different ways. One option is to take more of an "intent-to-treat" approach and define $A_{my}$ as an indicator representing months in which influenza activity tends to be the highest (i.e., December-March). Another approach is to more precisely focus on periods in which the largest impacts are expected, and define $A_{my}$ as a measure of influenza activity. Although these two methods obtain similar results in terms of magnitude, the latter is preferred for precision. The influenza activity measure is an index that is scaled to equal one during the month in the sample with maximum influenza activity, and the measure is common to all states.[16] Note that the main effect for match ($M_y$), the main effect for activity ($A_{my}$), and their interaction ($M_y * A_{my}$) are implicitly included in the

---

[16]Even though the measure of influenza activity is at the national level, there is still the concern that the measure is endogenous to the extent that it is determined by local conditions. This is addressed in two alternative specifications in Section 3.3.4: (1) by defining $A_{my}$ as an indicator for typically high influenza months, and (2) by using regional variation in influenza activity for seasons in which this is available (1998-) and defining $A_{smy}$ as influenza activity in all census divisions other than the one that contains state $s$.

month-by-year fixed effects. In Equation (3), $\phi_1$ measures the difference in the causal effect of vaccination ($\gamma_1$ from Equation (2)) between periods of maximum influenza activity and periods of zero activity. The coefficient $\phi_2$ represents a falsification test: the causal effect of vaccination during periods in which essentially zero influenza is circulating (and as such $\phi_2$ should equal zero).

Equation (3) includes both month-by-year fixed effects and state-by-month fixed effects. The state-by-month fixed effects not only account for any fixed state-specific factors, but allow seasonality in the outcome to vary by state. Note that the identification strategy does not rely on the inclusion state fixed effects, but inclusion of state-by-month fixed effects is preferred for precision. All models also include flexible controls for temperature, precipitation, and humidity (Barreca and Shimshack, 2012). No state-level covariates that vary at the annual level are included; 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. These results are presented alongside a range of other specifications in Table A1. In all specifications, standard errors are clustered at the state level.

## 3.3  Results

In interpreting the estimates, note that the three regressors (*Vaccination*, *Match*, and *Activity*) are all continuous 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.

### 3.3.1  Main Results

Table 3 provides estimates of the difference-in-differences equation described in Equation (2). These estimates are presented for both the all-age PI mortality rate per 100,000 population (Panel A) and average work hours lost due to illness absence (Panel B). Column 1 includes all months, Columns 2-3 compare influenza season months (December-March) to non-season months (April-November), and Columns 4-5 compare months with high influenza activity months (activity index $\geq 0.5$) and low activity months (activity index $< 0.5$).

First, let us consider the coefficient estimates for "Vacc", which represent the relationship between vaccination rates and the outcomes when the match rate is zero. These estimates are generally positive, and typically larger during influenza season months or periods of high influenza activity. This suggests that states with higher influenza-related mortality or more influenza-related work absences tend to have higher vaccination rates. In other words, these estimates pick up the potentially endogenous component of the relationship between

17

vaccination and the outcomes.

The coefficient estimates for the interaction term are the objects of interest. These estimates represent the differential effect of higher vaccination rates between flu-years in which the match is equal to one relative to flu-years in which the match is equal to zero. For PI mortality, the all-month estimate in Column 1 is negative and significant; the point estimate of -0.047 implies that a one percentage point increase in the vaccination rate decreases PI mortality by 0.047 per 100,000 in flu-years when the vaccine is perfectly matched relative to flu-years when the match is zero. Magnitudes will be discussed in greater detail in the discussion of Table 4 below. For work absences, the all-month estimate in Column 1 is also negative, but it is not statistically different from zero. The point estimate of -0.0014 implies that a one percentage point increase in the vaccination rate decreases the average number of hours lost due to illness for full-time workers by 0.0014 (an approximate 0.5% decrease). The lack of precision and relatively small magnitude of these estimates in part motivates analyzing months in which the benefits are expected to be largest.

For both PI mortality and work absences, the general patterns in comparing high influenza months to low influenza months are similar: the estimates are large in magnitude and statistically significant during high influenza months (influenza season or high activity months) and relatively small in magnitude during low influenza months (non-season or low activity months). Compared to the all-month estimates in Column 1, Column 2 shows that the estimates are larger during influenza season, when influenza activity is typically higher. Taking this comparison further, Column 4 shows that the estimates are even larger during months with high measured influenza activity. The estimates for low activity months (Columns 3 and 5) are smaller in magnitude compared to the all-month estimates; these are negative and significant in some cases, which is not unexpected since the estimates represent months with relatively little influenza activity rather than zero activity.

The triple-difference model formalizes the comparison between Columns 4 and 5 of Table 3. This formalization has the advantages of (1) utilizing the entire distribution of influenza activity rather than an arbitrary classification of months in to "High" and "Low" activity, and (2) providing a falsification check in the form of a coefficient estimate that represents the causal effect of vaccination during periods of *zero* measured influenza activity. The triple difference estimates for both PI mortality and work absences are reported in Table 4. Column 2 in this table represents the main specification, but the table presents two additional specifications as well: Column 1 represents a more parsimonious specification that omits the state-by-month fixed effects and Column 3 builds on the main specification by instrumenting for the vaccination rate (and all interactions) using the average vaccination rate over three seasons prior (and all interactions). Additional specifications that vary the

18

fixed effects and use of the instrument are presented in Table A1, and alternative defini-
tions of the instrument are explored in Table A2.[17] Across all specifications in Table 4, and
Tables A1 and A2, the estimates are of comparable magnitude.

Next consider the magnitude of the estimates in Column 2 of Table 4. Note that all
subsequent analyses and calculations are based off of this specification. For both PI mortality
and hours absent for illness, the coefficient on the $Vacc \times Match \times Activity$ interaction is
negative and significant, while the coefficient on the $Vacc \times Match$ interaction is small and
indistinguishable from zero. The implication is that influenza vaccination rates decrease both
PI mortality and hours absent during periods of high influenza activity, but have no impact
on the outcomes when measured influenza activity is zero. The point estimate of -0.223
on PI mortality implies that a one percentage point increase in the influenza vaccination
rate will decrease the PI mortality rate by 0.223 per 100,000 individuals during months with
maximum influenza activity relative to months with no influenza activity, and during seasons
in which the vaccine is perfectly matched relative to seasons in which the vaccine is poorly
matched. Similarly, the coefficient estimate of -0.0106 on hours absent implies that a one
percentage point increase in the influenza vaccination rate will decrease the average work
hours lost per week for full-time workers by 0.0106 under the same conditions.

Because these interpretations are somewhat nonintuitive, I also report an estimate of the
"Expected Annual Benefit" of vaccination. To calculate the expected annual benefit, I first
calculate expected monthly benefits and then sum across months (described in table notes).
These estimates imply that in a population of 100,000 individuals, a one percentage point
increase in the vaccination rate (i.e., 1,000 additional vaccines) would decrease PI mortality
by 0.351 and decrease work hours lost to illness among full-time workers by 2,762 hours in
expectation. Put differently, the estimates imply that 2,849 vaccinations are required to save
one life and 2.53 vaccinations are required to save one 8-hour work day. Finally, suppose the
estimates are scaled to the size of the U.S. population: this implies that a one percentage
point increase in the vaccination rate for one year would result in 1,134 fewer deaths and
8.9 million fewer work hours lost in expectation. For mortality, this is a substantial, albeit
plausible number given that estimates of average annual deaths due to influenza lie in the
range of approximately 20,000-40,000.

Before moving on, it is useful to explore alternative definitions of the outcomes for the
purposes of (1) ensuring that the estimated benefits are concentrated in the outcomes that are

---

[17]The most important specification check in Table A1 is a model that includes state-by-year fixed effects;
the similarity of the estimates implies that the addition of any state-level annually varying covariates would
not substantially influence the estimates. The most important specification check in Table A2 is a model
that uses only variation in base period vaccination rates (i.e., the average vaccination rate over the first three
flu-years of the sample).

expected to be most heavily influenced, (2) testing whether the current definitions capture the full extent of influenza-related deaths, and (3) providing falsification checks using outcomes that should not be affected by influenza vaccination. First, we examine mortality by 34 standard cause of death categories in Figure 3. The cause-of-death categorization here uses the underlying cause of death only so that each category is mutually exclusive (as opposed to using multiple causes of death to define PI deaths as in the main analysis). As such, it is possible that deaths in non-PI categories have a secondary diagnosis for PI and so categories other than PI should not be considered as falsification tests. It is reassuring nonetheless that the largest and most highly significant coefficient estimate is for the PI category. The first 10 categories (in black, including PI) represent deaths due to respiratory and circulatory diseases and are thus relatively more likely to be influenced by influenza infection in comparison to the following 24 categories. The largest declines for non-PI deaths are in the Ischemic Heart Disease (e.g., heart attack) and Cerebrovascular Disease (e.g., stroke) categories; both of these conditions represent potential downstream consequences of influenza infection. Also note that the estimates are presented in levels, so that in relative terms the size of the estimate for the PI category is especially large in comparison to these other large categories.

Table A3 presents alternate definitions of mortality based on multiple causes of death as well as alternative definitions for work absences. In both cases, the main estimates are provided for reference in Column 1. For mortality, Column 2 represents deaths with any respiratory or circulatory cause, and Column 3 represents all-cause death. Each represents successively more broad definitions of mortality in comparison to PI deaths. The estimates grow when these more broad definitions are used, suggesting that using only the PI definition likely represents a lower bound on the total mortality benefits. That being said, there is substantial loss of precision using these broad categories, and as such the precision affording by the more specific PI definition is preferred for the main analysis. Column 4 represents deaths with no respiratory or circulatory diagnosis, and can be considered a falsification test; the estimate is small and indistinguishable from zero. Alternative definitions for work absences are provided as well. In Column 2, the outcome is the proportion of workers absent due to illness rather than the average number of hours absent due to illness. This estimate is negative and significant, and of similar magnitude.[18] Columns 3-4 examine both average hours absent and the proportion absent for reasons other than illness as falsification checks; these estimates are small (relative to the mean) and insignificant.

---

[18]To compare the proportion-absent estimate with the hours-absent estimate, the proportion-absent estimate is scaled by the average number of hours per absence in the data (10.2), resulting in an estimate of -0.0066 which is comparable to the hours-absent estimate of -0.011.

### 3.3.2 Age Heterogeneity and Decomposing Externality Effects

To this point, all estimates have represented the social benefits of vaccination. We next consider evaluating mortality benefits by age with the ultimate goal of disentangling private and external benefits. Age-specific estimates for mortality are provided for five age groups (infants under 1, 1-9, 10-64, 65-74, and $\geq$75) in Panel A of Table 5. Age-specific mortality rates are calculated using the total state population in the denominator rather than the age-specific population so that the estimates can be interpreted as an accounting of the total benefits. The numbers in brackets represent the percentage of total mortality reductions that accrue to each age group.[19]

The age-specific estimates indicate that the vast majority of the mortality benefits 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 $\geq$ 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 92% of the reduction in mortality due to influenza vaccination is experienced among the 65 and older population, and 85% is experienced in the 75 and older population.

To evaluate the extent to which the mortality benefits of influenza vaccination operate through an externality effect, I use the fact 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.[20] More specifically, the following equation is estimated:

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

In Equation (4), the full set of interactions described in Equation (3) for both people under 75 and people at least 75 are included. As such, $\psi$ represents a combination of direct and externality effects, where the externality effects are limited to capturing the spread of influenza among people within the 75 and older group. The coefficient $\omega$ represents the effect of vaccination among people under 75 on influenza-related mortality for individuals who are at least 75; this represents a pure externality effect. The results of this exercise are presented

---

[19]This is calculated as the corresponding coefficient estimate divided by the all-age estimate in Column 2 of Table 4. I provide an additional set of estimates in Table A4 that uses the age-specific population in constructing mortality rates with very similar results.

[20]Age-specific vaccination rates are constructed using the BRFSS, which is individual-level data with information on each respondent's age.

in Panel B of Table 4. The figures in brackets represent the proportion of the total effect attributable to each age group, and indicate that 66% of the total mortality reductions in the $\geq 75$ age group are attributable to the vaccination of those under 75. In other words, the results suggest that over half of the mortality benefits of influenza vaccination operate through an externality channel. This finding accords with theoretical predictions presented in Figure 1 which indicate the majority of the social benefits of vaccination operate through an externality, and given the relatively low efficacy of influenza vaccination in older individuals, this is an important though not necessarily surprising finding.

### 3.3.3 Monetizing Benefits

Any policy aimed at increasing influenza vaccination take-up should weigh the costs and benefits of doing so; the goal of this section is to provide monetary estimates of the marginal social benefits of vaccination in terms of both mortality and work absences.

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 the VSL 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 EPA standard of \$8.8 million as an upper bound.[21] 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 the age groups presented in Table 5, taking a weighted average of single-year VSL estimates where the weight is the share of deaths from each single-year age group. These VSL estimates, along with estimates of the monetary benefits of vaccination are presented in Table 6.

For each age group, I consider the expected annual benefits of a policy that increases the national influenza vaccination rate by one percentage point. I use estimates of age-specific reductions in mortality from Table 5 to estimate the expected annual number of deaths avoided for each age group, and multiply by the age-adjusted VSL.[22] Benefits per vaccination are determined by dividing the total benefit by 3.23 million (1% of the 2016 U.S. population), as that is the number of additional vaccinations required to achieve the corresponding total benefit. While the reductions in mortality are concentrated among individuals at least 75,

---

[21]Each VSL figure is reported in 2016\$.

[22]This calculation is similar to the "Expected Annual Benefits" calculation in Table 4. Specifically, the expected annual number of age-specific deaths avoided per 100,000 population ($\sum_m \hat{\phi}_1 \times \overline{Match} \times \overline{Activity}_m$) is scaled to the U.S. population by multiplying by a factor of (323,100,000/100,000).

the monetary benefits are more equally distributed given the relatively low VSL for elderly individuals. Considering all ages, I find the benefit of an additional vaccination to be either $146 (AG VSL) or $555 (EPA VSL). That being said, the only age-specific estimates that are statistically significant are the estimates for the at least 75 group. As such, my preferred estimates of the marginal social benefits of vaccination in terms of mortality are limited to this group: $74 using the AG VSL, and $282 using the EPA VSL. As a lower-bound estimate of the marginal externality benefits, these numbers can be multiplied by 0.66 (the estimated proportion of externality benefits from the previous section), resulting in estimates of $49 using the AG VSL and $186 using the EPA VSL.

Because monetary benefits of mortality reductions are inherently controversial among the general public, and because the mortality reductions in this specific case are concentrated among the elderly population who may already be suffering from other illnesses, it is advantageous to also present benefits of vaccination that are subject to less controversy and require fewer assumptions in monetizing. The process of calculating monetary benefits for work absence reductions is somewhat more straightforward: the expected annual number of hours saved for the U.S. population is multiplied by the median hourly wage.[23] The calculation suggests that each vaccine confers benefits equal to $57 in terms of reduced work absences among full-time workers, in expectation.

### 3.3.4 Robustness Checks

This section provides a very brief summary of a number of robustness checks. All results are presented in the Online Appendix, and specifications are discussed more fully in table and figure notes.

First note that a number of these robustness checks that have already been mentioned. Table A1 presents a series of specifications demonstrating that the estimates are insensitive to the choice of fixed effects and the use of the instrumental variables strategy. Table A2 demonstrates that the estimates are insensitive to various definitions of the instrument, including defining the instrument as a time-invariant vaccination rate in a base period (the average over the first three seasons). Table A4 demonstrates that the age-specific estimates are not sensitive to defining mortality rates using the age-specific population.

Potential nonlinearities in the marginal effect of vaccination are explored in Figure A2. This exercise shows little evidence that the marginal benefit of vaccination changes substantially over the observed distribution of vaccination rates (consistent with the theoretical

---

[23]This calculation is also similar to the "Expected Annual Benefits" calculation in Table 4. Specifically, the expected annual number of hours saved in the U.S. is $\sum_m \hat{\phi}_1 \times \overline{Match} \times \overline{Activity}_m \times (30.5/7) \times 126,000,000$, where the final term is the approximate number of full-time workers in the U.S.

model in Figure 1). The possibility of lagged effects on mortality or work absences is examined in Table A5. The reported estimates represent the sum of contemporaneous and lagged impacts (for multiple lag lengths), and the stability of the estimates suggest that the vast majority of the impacts operate contemporaneously.

Table A6 presents several additional checks, all of which demonstrate the insensitivity of the main estimates. Column 1 excludes flu-years that use interpolated data on vaccination. Column 2 includes the two flu-years affected by the H1N1 pandemic. Column 3 uses a limited sample that ends in the 2007/08 flu-year to ensure that the estimates are not substantially influenced by three factors (1) the H1N1 pandemic, (2) the development of the high-dose vaccine in 2009, and (3) the development of the quadrivalent vaccine in 2012. Columns 4-5 use the sample 1998/99-2014/15 in which region-specific information on the match rate and influenza activity are available. Column 4 allows the match rate to be defined as region-specific. Column 5 defines influenza activity to be the average influenza activity in all census divisions *except* that which contains the corresponding state in order to ease concerns over the possibility that nationally-defined influenza activity is endogenous. Finally, Column 6 uses an influenza season indicator in place of influenza activity in the triple difference specification in order to further assuage such concerns.

# 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 among those who are closest to the margin of the decision 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 likely to be particularly large. I examine the effects of mandates requiring health care workers be vaccinated against influenza on the outcomes of hospital patients in counties and hospitals subject to the mandates. To begin, I describe the institutional background.

## 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 required all hospitals to report to the

California Department of Public Health (CDPH) on the percentage of HCWs vaccinated against influenza in each season; this allows for the measurement of hospital-worker vaccination rates. This law also required that all health facilities offer free vaccinations to employees and required that they sign a statement declaring that he or she had declined vaccination if that was the case. 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 means that baseline levels of vaccination are relatively high by the time the first mandates go into effect in 2009.

In May of 2009, the H1N1 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 flu-years affected by H1N1 differently than others 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 mandates. Beginning in the 2011/12 influenza season, counties implemented county-wide vaccination mandates, and in each season since more counties have followed.[24] In the 2015/16 season, over 75% of hospitals were subject to a mandate.

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 most mandates applied much more broadly. Typically, all licensed health care facilities would be subject to these more broad mandates. Figure A3 maps the implementation of both hospital and county-level mandates over time and distinguishes between the type of county-level mandate. Because the vast majority of mandated hospitals are subject to county-level mandates that apply beyond the hospital, these are the policies on which I focus. The main outcomes are measured at the hospital level, but because these mandates apply more broadly, it is not necessarily the case that an observed infection was transmitted within the hospital. It is possible that these mandates affect the transmission of influenza in non-hospital health care settings and in the community at large if HCWs act as important vectors for disease.

## 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, vaccination rates for HCWs and hospital-level outcomes. In the main analysis, the data cover flu-years 2007/08 through 2015/16.

---

[24]Implementation dates for all hospital- and county-level policies are described in Table A7.

### 4.2.1 Mandates & Vaccination Rates

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 that lists health care organizations across the U.S. that mandate influenza vaccination and the dates of implementation.[25] CDPH maintains a list of county-level mandates with implementation dates, but the list is not completely accurate with respect to the implementation dates. Through a process of searching 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 CDPH.[26] Summary statistics in Table 1 indicate the number of hospitals subject to mandates in each flu-year.

As required by California law, all licensed hospitals report information on the vaccination status of their workers to CDPH for each season. 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%. Reporting in the first season in which reporting was required (2008/09) was particularly poor, and so data from this season was omitted. For the remaining flu-years, the main first-stage estimates use only hospitals that have response rates of at least 90% in all flu-years, though the estimates are not sensitive to this restriction.

### 4.2.2 Hospital Patient Outcomes

The primary data source on outcomes are two restricted data files on the universe of inpatient hospital admissions and outpatient ED visits in California between 2005 and 2016, obtained through California's Office of Statewide Health Planning and Development (OSHPD). In the analysis to follow, inpatient admissions and outpatient ED visits are analyzed separately, though I refer to any visit to the hospital (inpatient admission or ED visit) simply as a "visit". Note that ED visits often result in hospital admission; to avoid double-counting, ED visits are dropped if the patient is transferred to another health care facility (i.e., all ED visits are outpatient).

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

---

[25]Source:http://www.immunize.org/honor-roll/influenza-mandates/.

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

The inpatient data include a number additional features that are utilized. One particularly useful feature is that each diagnosis includes an indicator for whether it was present at the time of admission, allowing me to focus specifically on hospital-acquired infection. Certain outcomes are less specific to influenza, but may be significantly affected during periods of very high influenza activity. These include average length of stay, average hospital charges, and the in-hospital death rate. 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 NVSS.

## 4.3 Empirical Framework

I estimate the impacts of vaccination requirements using a standard difference-in-differences (DD) framework that exploits quasi-experimental variation in the timing of mandates. Similar to the aggregate analysis, I also use a triple-difference framework that additionally exploits the timing and magnitude of influenza activity for outcomes that are not necessarily specific to influenza (and thus exhibit variation throughout the year). Because HCW vaccination rates are measured annually, the triple-difference strategy is not necessary or possible in estimates of the first-stage. Furthermore, the primary outcomes of interest are influenza-specific, and because there is no variation in these outcomes in months with no influenza circulating, the triple-difference strategy is not appropriate for these outcomes.[27] Consider the following DD equation to be estimated at the annual level:

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

In Equation (5), $Y_{hy}$ represents either vaccination rates (first stage), or the number of influenza diagnoses (reduced form) at hospital $h$ in flu-year $y$. $\text{Required}_{hy}$ is a variable indicating whether there is a vaccination requirement in effect; $\delta_h$ and $\delta_y$ are hospital and flu-year fixed effects. The coefficient of interest, $\pi$, is identified under the assumption that variation in the timing of the mandates is uncorrelated with other unobserved time-varying determinants of the outcomes. While the identifying assumption is fundamentally un-testable, I provide evidence from indirect tests that support the assumption. Importantly, in the discussion of results I provide an event study version of Equation (5); this exercise indicates that changes in the outcomes coincide with the implementation of the policy, and that the treatment effects are not identified off of differential trends between treatment and control hospitals. Furthermore, the event study provides information on the necessity of county- or

---

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

hospital-specific time trends, which are included in some specifications.

Because influenza diagnoses represent a highly specific outcome, there are many hospital-by-flu-year cells with the outcome equal to zero. Given the count nature of the data, and the over-dispersion indicated in Table 1 (i.e., the variance is greater than the mean), a negative binomial model is used to estimate Equation (5).[28] 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, which varies considerably across the sample. This is done through the use of an exposure variable, which is set to be the mean annual number of all-cause visits in 2005/06-2006/07 (prior to the period of analysis). Note that the first stage is estimated via OLS.

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:

$$Y_{hmy} = \alpha + \theta_1(\text{Required}_{hmy} * \text{Activity}_{my}) + \theta_2\text{Required}_{hmy} + \delta_h + \delta_{my} + \varepsilon_{hmy} \qquad (6)$$

In Equation (6), the policy indicator is interacted with an index of influenza activity, $\text{Activity}_{my}$. 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). The main effect for activity is absorbed by the month-year fixed effects. $\theta_1$ measures the effect of influenza vaccination mandates during a time of peak influenza activity relative to a period with zero influenza activity. Furthermore, $\theta_2$ measures the effect of influenza vaccination mandates during times of very low influenza activity and is expected to be near zero. Because the outcomes of interest here (average charges, average

---

[28]There are several possible count models available, and in the case of panel data requiring fixed effects (as here) the choice is not trivial (See Cameron and Trivedi (2013a,b) for a review of count models in general and specifically for panel data.). 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). 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 the results are not sensitive to the choice of count model.

length of stay, and the in-hospital death rate) are not counts, Equation (6) is estimated via OLS.

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 estimates are conducted at the county level (with county fixed effects), and the outcome is PI deaths per 100,000 population (estimated via OLS). In all models, standard errors are clustered at the county level.

## 4.4  Results

### 4.4.1  First Stage

The main result for the first stage is illustrated as an event study in Figure 4. This figure shows that there is little evidence of differential trends between hospitals that do and do not adopt vaccination mandates prior to implementation. In the first flu-year of implementation, vaccination rates increase sharply and remain relatively flat thereafter. There is some evidence of catch-up in that the effect fades slightly over time. This is likely a result of adopting hospitals reaching a near maximum vaccination rate while non-adopting hospitals still see increases in vaccination rates over time. The result is shown using a single treatment indicator in Table A8, revealing a highly significant coefficient estimate of 10.6 percentage points in the preferred specification.[29]

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 apply far more broadly than to just hospital workers. Because vaccination rates for other HCWs are not observed, results are displayed in the remainder of the paper as reduced-form policy estimates rather than in an IV 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 2014/15 survey 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 and unfortunately this cannot be tested this with the data available.

---

[29]This table also shows that the estimate is robust to using less restrictive sample selection criteria on data quality, or even using all hospitals regardless of data quality.

To get a rough estimate of the number of additional vaccinations received as a result of these mandates, consider a hypothetical state-wide mandate that affected all workers in licensed health care facilities in California. Figures from the Bureau of Labor Statistics indicate there were approximately 1.18 million workers in industries plausibly affected by these mandates in California during 2015.[30] Assuming the mandates affected all HCWs in the same way hospital workers were affected, multiplying by the first-stage estimate of 0.106 implies 125,053 additional vaccinations (approximately 0.3% of the California population).

### 4.4.2 Influenza Diagnoses

Before discussing magnitudes, first consider a number of specifications for the reduced-form estimates presented in Table 7; Panel A represents inpatient admissions and Panel B represents outpatient ED visits. Column 1 represents estimates from Equation (5) with no included trends or sample restrictions. The estimates are in the expected negative direction for both outcomes, but only marginally significant for inpatient admissions and insignificant for ED visits. While there is not enough variation in vaccine efficacy over the sample period to incorporate this into the model as in Part I, there is one flu-year in the sample (2014/15) that had a poor vaccine match and thus an ineffective vaccine. Dropping this flu-year from the sample results in larger coefficient estimates, reported in Column 2. This flu-year is omitted from all following specifications, which is preferable as it increases the ability to identify a statistically meaningful result, but it does change the interpretation of the results: the estimates represent the effect of HCW mandates during a well-matched season as opposed to the effects of the mandates in expectation.

Columns 3 and 4 include county-specific and hospital-specific linear time trends, respectively. In both cases, the magnitude of the estimates grow substantially. The event studies presented in Figure 5 illuminate this result. It is clear that the result is substantially stronger for inpatient admissions compared to ED visits, although the patterns are similar. Diagnoses for influenza are increasing in adopting hospitals relative to non-adopting hospitals in years prior to mandate implementation. In the first year of implementation, however, there is a sharp decrease in the number of visits that persists subsequently. The inclusion of hospital-specific linear time trends appears to fully correct any pre-existing trends that were present in the model that did not include trends. Models that include trends are preferred, but in models for inpatient admissions that include hospital-specific trends, the maximization algorithm failed to converge. The same is true of models that exclude H1N1 pandemic flu-years (a nontrivial portion of the sample), presented in Column 5 of Table 7. In both cases, the

---

[30]Employment in NAICS industries: 656211, 656214, 656216, 656219, 656221, 656222, 656223, 656231, 656232, 656233, 656239.

resulting point estimates from these non-convergent models are quite similar to the estimates for the (convergent) model that includes county-specific time trends, presented in Column 3.[31] For this reason, the estimates in Column 3 are preferred.

These estimates indicate that in seasons with a well-matched vaccine, HCW vaccination mandates are associated with a statistically significant 21.4% decrease in inpatient admissions with an influenza diagnosis and a marginally significant 9.4% decrease in outpatient ED visits with an influenza diagnosis. In interpreting these results, there are at least two differences between inpatient and outpatient visits to consider. First, for inpatient visits it is possible that the individual in question was infected with influenza during their hospital stay. This means that policy-induced changes in diagnoses in an inpatient setting likely result from a combination of hospital-acquired influenza as well as influenza acquired in a non-hospital health care setting or in the community at large. For outpatient ED visits, the hospital-acquired channel is highly unlikely. Second, influenza may be less likely to be correctly diagnosed in an outpatient setting in which there is less time acquire laboratory confirmation (Dugas et al., 2015). Both of these differences would lead to smaller point estimates for outpatient visits.

Consider next a comparison between the effects of vaccination mandates on hospital- and non hospital-acquired influenza. Again, hospital-acquired influenza is only a possibility for inpatient visits. The final two columns of Table 7 examine the distinction between influenza diagnoses that were present on admission (POA) and those that were not (Not-POA). Diagnoses that were present on admission are more likely to represent influenza acquired outside of the hospital, whereas Not-POA diagnoses likely represent hospital-acquired infection. Influenza diagnoses are coded as present on admission for the vast majority of visits, although there is some question as to how accurate POA coding is, especially since hospitals have financial incentives to avoid coding hospital acquired conditions (Goldman et al., 2011).[32] The point estimate for POA influenza diagnoses is very similar to the estimate for all influenza diagnoses presented previously. Focusing on Not-POA diagnoses reveals an effect almost twice as large in relative terms: vaccination mandates lead to an approximate 40.2% reduction in hospital-acquired influenza.

---

[31]Note that there are 58 counties in California and approximately 450 hospitals (depending on the year); this means that estimating models with hospital-specific time trends requires estimating a substantial number of additional parameters.

[32]The Deficit Reduction Act of 2005 stipulated that hospitals would not receive higher payments for certain secondary conditions that were not POA. Influenza is not one of these specific conditions, but it could still be the case that hospitals develop a habit of under-reporting all hospital acquired conditions.

### 4.4.3 Additional Results & Discussion

A brief discussion of several additional results provides additional insight. Age-specific impacts of HCW mandates on both inpatient and outpatient visits are provided in Table A9. The estimates are negative for all age groups. While the differences tend not to be statistically meaningful, the magnitudes in relative terms tend to be largest for children (infants and children 1-9). This is worth noting in at least one respect: since vaccination of children should not have been affected by these policies, the implication is that much of the identified impacts operate through an externality effect.

Estimates for four additional outcomes that exhibit variation throughout the year are presented in Table A10, estimated using Equation (6). The first three outcomes (average length of stay, average charges, and the in-hospital death rate) are hospital-level outcomes constructed using inpatient data. The estimates indicate small reductions in average length of stay (1.3%) and average charges (2.0%), and no impact on in-hospital deaths during months of peak influenza activity. These outcomes are distinct from influenza-specific outcomes in that they may be viewed as measures of how well the hospital is functioning more generally. In other words it may be that these outcomes are due to changes the number influenza infections, but it may also reflect improvements in the health of the hospital staff, improving the quality of their work. Column 4 presents estimates with the PI mortality rate as the outcome, and represents estimates at the county level. While it would be advantageous to examine the same outcome in Parts I and II of this paper, these estimates are underpowered to do so given the non-specific nature of this outcome. It is perhaps worth noting that the point estimate (-0.243), while not statistically significant, is of comparable magnitude to a one percentage point increase in the vaccination rate of the general population estimated in Part I, yet the number of vaccinations associated with a county-level HCW mandate is one third as large at most.

Finally, similar to Part I, I provide an estimate of the monetary benefits of vaccination. Note that the observed decreases in influenza *diagnoses* do not necessarily represent decreases in the actual number of *visits* if a substantial number of the affected visits are for other diseases that influenza infection complicated. To the extent that at least some of the observed reductions in influenza diagnoses represent a change in the number of visits, it is possible to provide a back-of-the-envelope calculation for the monetary benefit of a HCW vaccination in terms of hospital cost savings. Focusing only on visits most likely to be avoided by the mandates, I re-estimate Equation (5) using only visits with a primary diagnosis for influenza (as opposed to any diagnosis) for outpatient visits, and only visits with a primary diagnosis for influenza that was present on admission for inpatient admissions. I consider a hypothetical statewide vaccination mandate for California (relative to no mandate) that

affects all HCWs in the state. A number of assumptions are required to make this calculation, and a conservative approach is taken in applying each. I estimate that, in terms of reduced health care costs, each vaccine confers benefits of \$121.73 through reduced inpatient visits, and \$21.43 through reduced outpatient visits, for a total of \$143.17.[33] An estimate of the marginal social benefit would of course be larger if other outcomes such as mortality or work absences could be considered in this context as well.

# 5  Discussion and Conclusion

In this paper, I estimate the marginal social benefits of influenza vaccination for the general population (henceforth, GP) and for the population of health care workers (HCWs). Because it is not possible to use the same identification strategy and outcomes for both analyses, a direct comparison of the marginal benefits in each population requires fairly strong assumptions. I believe that such a comparison is still quite useful with this caveat in mind.[34] To compare vaccination in these two populations, I calculate the number of GP vaccinations or HCW vaccinations required to achieve a 1% reduction in influenza-induced mortality (for GP) or a 1% reduction in influenza-induced outpatient visits (for HCW) in a population

---

[33]This calculation is described in more detail here. First, the mean annual number of inpatient and outpatient visits are 4,190 and 34,221, respectively (excluding the H1N1 flu-years). The estimated coefficients using only primary diagnoses for inpatient and outpatient visits are -0.259 and -0.105, implying 1,085 fewer inpatient visits and 3,593 fewer outpatient visits each year. The OSHPD data include information on charges for inpatient visits, but not outpatient visits. Average charges for visits with a POA primary diagnosis for influenza are \$48,719. The Nationwide Emergency Department Sample has information on ED outpatient charges for some states, and the average charges for visits with a primary diagnosis for influenza are \$1,472. Since charges do not represent hospital costs, this figure is multiplied by the national average cost-to-charge ratio (0.507) for outpatient visits and a California-specific cost-to-charge ratio (0.288) for inpatient visits. As such, the hypothetical statewide policy achieves annual savings of \$15.22 million through reduced inpatient visits and \$2.68 million through reduced outpatient visits. To arrive at a per-vaccination figure, I divide by the number of additional vaccinations (125,053) received in a hypothetical statewide policy, derived in Section 4.4.1.

[34]This calculation is described in more detail here. For GP, the main estimates indicate that 1,000 additional vaccinations are required to reduce annual influenza mortality by 0.35 per 100,000. To get this in relative terms, I require an estimate of annual influenza mortality, and use the nationwide 23,607 figure estimated by Thompson et al. (2010). Using the 2007 US population (to reflect the period of study in Thompson et al. (2010)), this implies 7.8 influenza deaths per 100,000 population. As such, my estimates indicate that 1,000 additional vaccinations result in a 4.49% decrease in influenza mortality for a population of 100,000; turning this around, 222.7 vaccinations are required to achieve a 1% reduction. For HCW vaccinations, I use only outpatient visits with a primary diagnosis for influenza as this provides a lower bound (the smallest relative impacts) and I consider the same hypothetical statewide policy as in Section 4.4.3, but scaled to a population of 100,000 instead of the 2016 California population (39.3 million). For the entire state, I found that 125,053 vaccinations are required to achieve a 10.5% reduction in outpatient visits, implying that 11,910 vaccinations are required to achieve a 1% reduction. Scaling this down to a population of 100,000 implies that 30.3 HCW vaccinations are required to achieve a 1% reduction in outpatient ED visits.

of 100,000 individuals. I find that 222.7 GP vaccinations are required to achieve a 1% reduction in mortality and 30.3 HCW vaccinations are required to achieve a 1% reduction in outpatient visits. Under the assumption that influenza mortality and influenza outpatient visits are proportional to each other (or to the number of influenza cases), the implication is that HCW vaccinations are 7.3 times more effective at quelling the spread of influenza in comparison to vaccinations in the general population. While this may seem large, it is not unreasonable to argue that HCWs are many times more likely to come in contact with infected individuals and much more likely to come into contact with individuals who would suffer severe consequences from being infected (e.g., a hospital visit).

For both vaccination in the general population and health care workers, I have provided policy-relevant estimates of the marginal social benefits in terms of the outcomes analyzed: these benefits are estimated to be $74 per GP vaccination in terms of reduced mortality (at minimum), $57 per GP vaccination in terms of work hours gained, and $143 per HCW vaccination in terms of reduced health care costs. For mortality, I find that the majority of the social benefits operate through an externality. For HCW vaccination, I do not explicitly estimate the size of the externality, but it is likely that the vast majority of the social benefit operates through the externality given that the largest relative benefits exist in a group whose vaccination status is not affected (children).

How do these benefits compare to the marginal cost of vaccination? Prosser et al. (2008) estimate that the cost of administering a vaccine (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.[35] 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 the estimates presented in this paper 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 vaccination among health care workers.

The analysis of aggregate vaccination rates is relevant to a policy that would increase vaccination 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 ac-

---

[35]Dollar estimates are converted to 2016$.

cessibility to vaccine providers, for instance. The marginal social and externality benefits estimated here suggest that vaccination policy resulting in marginal increases in the vaccination rate above current levels is beneficial so long as the marginal cost curve does not increase steeply at the current level of vaccination. While 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). Furthermore, Bronchetti et al. (2015) show that a relatively small financial incentive can result in large increases in vaccine take-up, suggesting the existence of low-cost policies that would increase vaccination in at least some segments of the population.

For health care workers, the estimates presented here are large in comparison to administration 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 derived from policies that *mandate* influenza vaccination, creating an extremely high cost for those choosing not to vaccinate.[36] It is possible that other incentive-based programs could achieve a more efficient result if there are individuals for whom 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 estimated here are large enough to suggest that any policy increasing vaccination among health care workers would be cost-effective under reasonable assumptions about the costs.

In summary, I estimate that the social benefits of influenza vaccination are substantial and that much of the total benefits operate through externality effects. Determining the socially optimal level of vaccination depends critically on the marginal cost of vaccination – under reasonable assumptions about these marginal costs, the results of this study indicate that policies increasing take-up of influenza vaccination in either the general population or in the population of health care workers are likely to be cost-effective.

---

[36]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.

# References

**Adda, Jérôme**, "Economic activity and the spread of viral diseases: Evidence from high frequency data," *The Quarterly Journal of Economics*, 2016, *131* (2), 891–941.

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

**Almond, Douglas**, "Is the 1918 Influenza pandemic over? Long-term effects of in utero Influenza exposure in the post-1940 US population," *Journal of Political Economy*, 2006, *114* (4), 672–712.

_ **and Bhashkar Mazumder**, "The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data," *American Economic Review*, 2005, pp. 258–262.

**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 I and Jay P Shimshack**, "Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the United States," *American journal of epidemiology*, 2012, *176*, 114–122.

**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.

**Bronchetti, Erin Todd, David B Huffman, and Ellen Magenheim**, "Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination," *Journal of Economic Behavior & Organization*, 2015, *116*, 270–291.

**Brown, Ryan and Duncan Thomas**, "On the long term effects of the 1918 U.S. influenza pandemic," *Working Paper*, 2016.

**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.

**Carpenter, Christopher S and Emily C Lawler**, "Direct and Spillover Effects of Middle School Vaccination Requirements," Technical Report, National Bureau of Economic Research 2017.

**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).

**Currie, Janet and Hannes Schwandt**, "Within-mother analysis of seasonal patterns in health at birth," *Proceedings of the National Academy of Sciences*, 2013, *110* (30), 12265–12270.

**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.

**Dugas, Andrea F, Alexandra Valsamakis, Mihir R Atreya, Komal Thind, Peter Alarcon Manchego, Annum Faisal, Charlotte A Gaydos, and Richard E Rothman**, "Clinical diagnosis of influenza in the ED," *The American journal of emergency medicine*, 2015, *33* (6), 770–775.

**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.

**Goldman, L Elizabeth, Philip W Chu, Dennis Osmond, and Andrew Bindman**, "The accuracy of present-on-admission reporting in administrative data," *Health services research*, 2011, *46* (6pt1), 1946–1962.

**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.

**Kelly, Elaine**, "The scourge of asian flu in utero exposure to pandemic influenza and the development of a cohort of british children," *Journal of Human Resources*, 2011, *46* (4), 669–694.

**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.

**Lawler, Emily C**, "Effectiveness of vaccination recommendations versus mandates: Evidence from the hepatitis A vaccine," *Journal of Health Economics*, 2017, *52*, 45–62.

**Lin, Ming-Jen and Elaine M Liu**, "Does in utero exposure to Illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment," *Journal of Health Economics*, 2014, *37*, 152–163.

**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," *Journal of Public Economic Theory*, 2017, *19* (3), 603–619.

**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.

**Oster, Emily**, "Does disease cause vaccination? Disease outbreaks and vaccination response," *Journal of health economics*, 2018, *57*, 90–101.

**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.

**Schwandt, Hannes**, "The lasting legacy of seasonal influenza: In-utero exposure and labor market outcomes," 2017.

**Serres, Gaston De, Danuta M Skowronski, Brian J Ward, Michael Gardam, Camille Lemieux, Annalee Yassi, David M Patrick, Mel Krajden, Mark Loeb, Peter Collignon et al.**, "Influenza vaccination of healthcare workers: critical analysis of the evidence for patient benefit underpinning policies of enforcement," *PloS one*, 2017, *12* (1), e0163586.

**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?," *Labour Economics*, 2018, *51*, 227–246.

**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.

**Stoecker, Charles, Nicholas J Sanders, and Alan Barreca**, "Success Is something to sneeze at: Influenza mortality in cities that participate in the Super Bowl," *American Journal of Health Economics*, 2016, *2* (1), 125–143.

**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
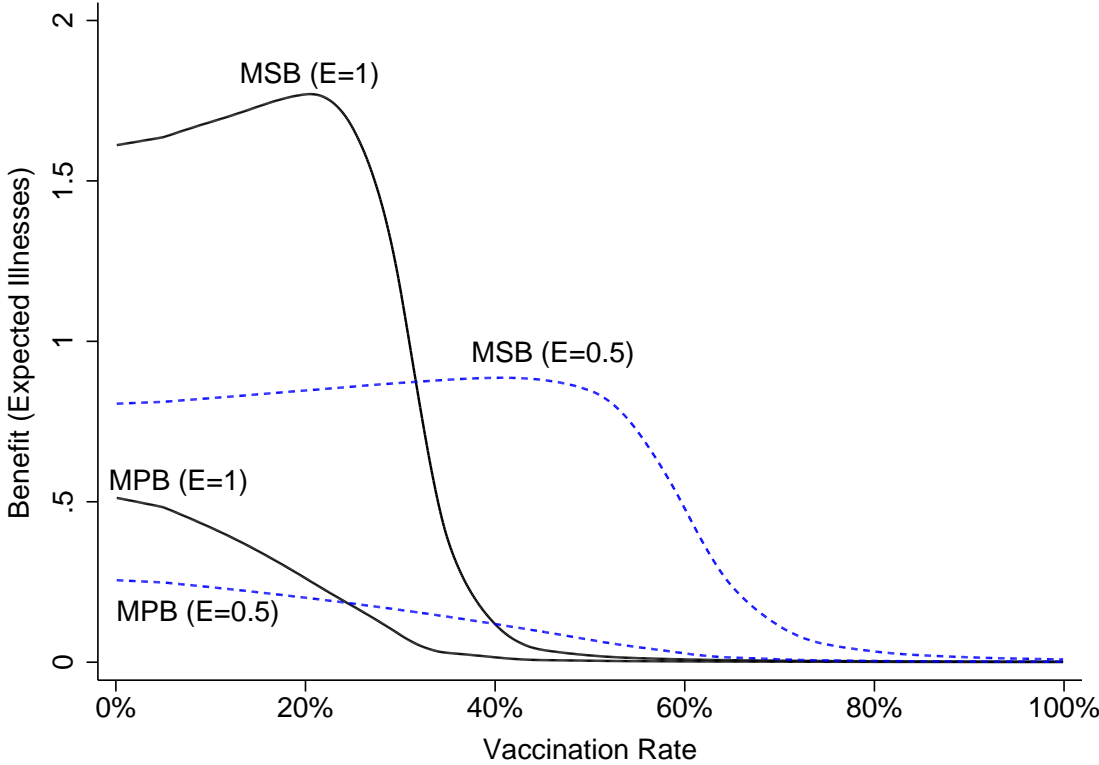


This figure presents a version of a model derived by Boulier et al. (2007) that describes the theoretical marginal social benefit (MSB) and marginal private benefit (MPB) curves for the case of influenza vaccination. These benefits are allowed to depend on vaccine efficacy. Here, two levels of vaccine efficacy are presented: 100% (E=1) and 50% (E=0.5).

Figure 2: Actual and Effective Vaccination Rates
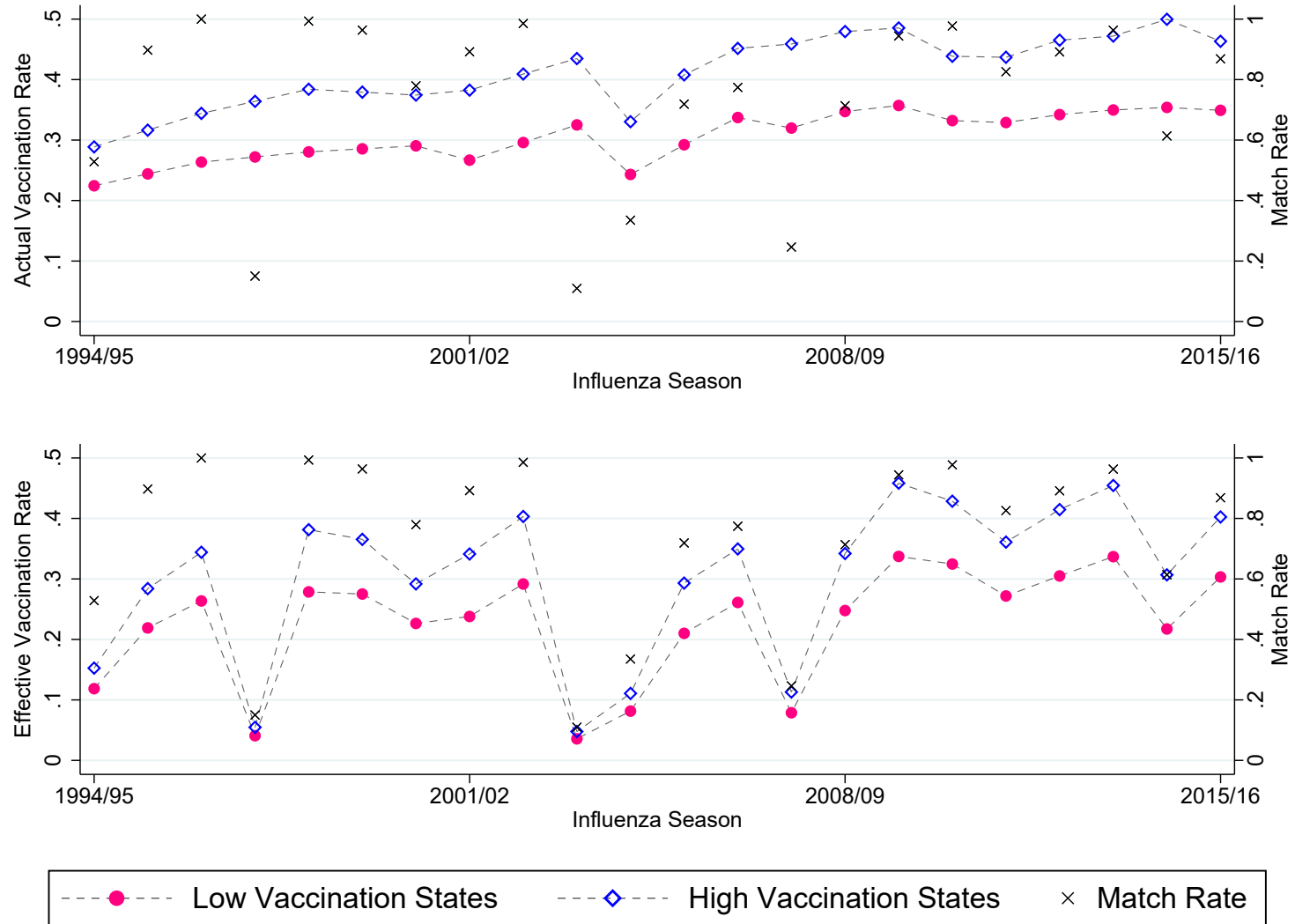
Note – There was a vaccine shortage in the 2004/05 season, accounting for the dip in vaccination rates during that season.

Figure 3: DDD Effect by Cause of Death
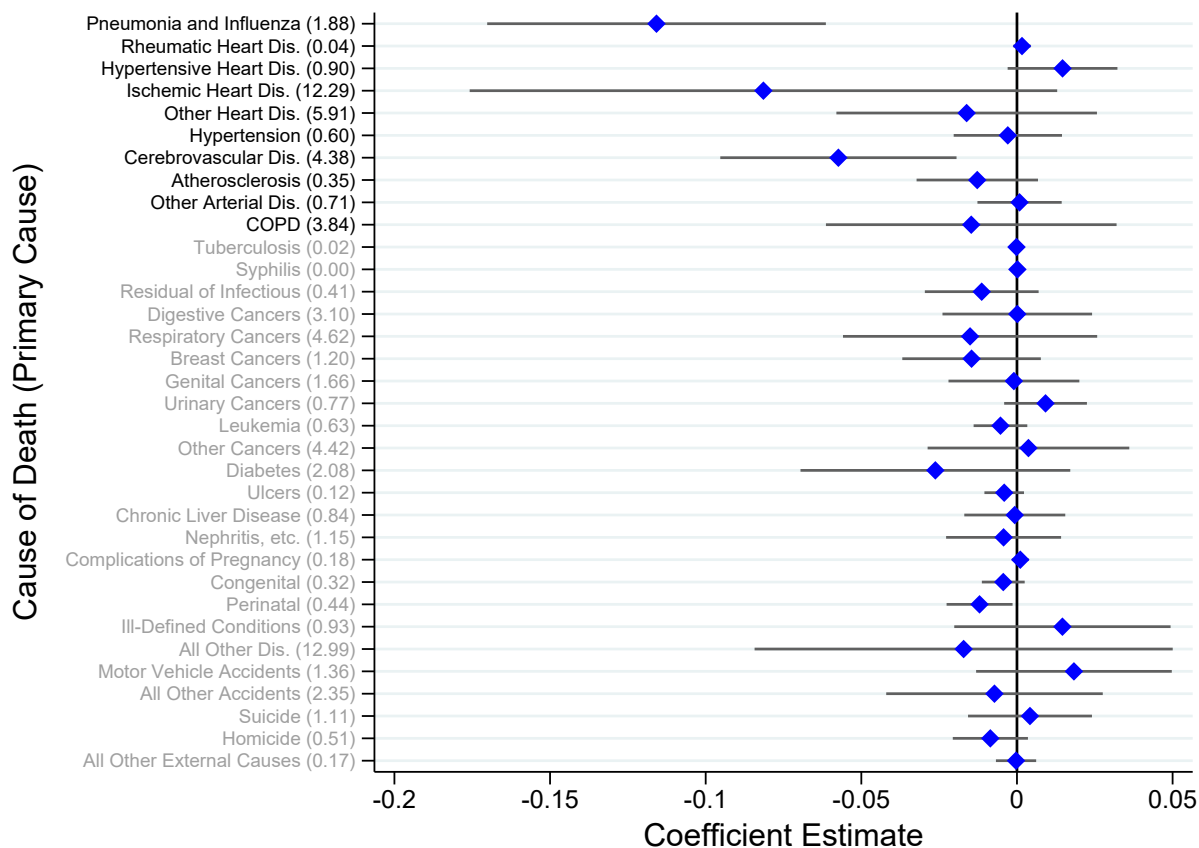
Note – Estimates represent the coefficient estimates (in levels) on the triple interaction in Equation (3) where the outcome is the mortality rate in various cause of death categories. Bars represent 95% confidence intervals. The mean monthly death rate (per 100,000 population) for each category is shown in parentheses. Cause of death is categorized by *primary* cause of death, so that each category is mutually exclusive. Cause of death categories are based on the 34-cause recode used by the NCHS for the period 1971-1998; deaths in the 1999- period were mapped from the updated 39-cause recode. Note that the mean PI death rate (1.88) is approximately one third of the death rate used in the measure for the main analysis that is based on multiple causes of death; this implies that approximately two-thirds of deaths categorized as PI in the main analysis are categorized as such based on a secondary diagnosis. The nine categories following PI (in black) are for respiratory and circulatory diseases and are plausibly related to influenza infection (more often having a secondary diagnosis for PI); the remaining categories are less likely to be related to influenza infection.

Figure 4: First Stage Event Study (HCW Vaccination Rates)
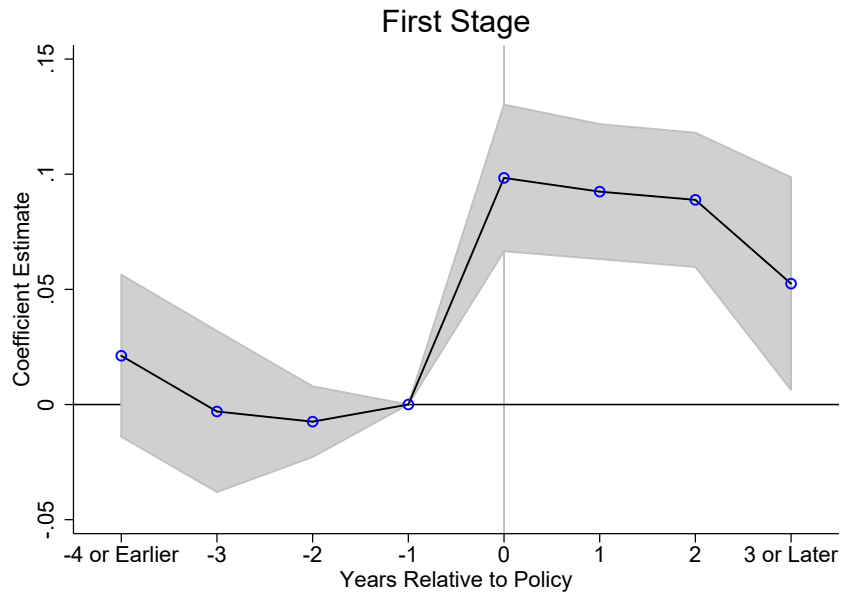


Points on this plot represent the point estimates from an event-study version of Equation (5) with HCW vaccination rates as the outcome. Shaded regions represent 95% confidence intervals. The event-study is estimated by replacing the policy indicator ($\text{Required}_{hy}$) with a series of variables indicating years relative to the policy: $\sum_{j=-4}^{-2} \gamma_j \text{Required}_{hyj} + \sum_{j=0}^{3} \gamma_j \text{Required}_{hyj}$. The indicator representing one year prior to the policy is omitted as the reference group. "-4 or Earlier" represents four or more years prior to policy implementation; "3 or Later" represents three or more years after policy implementation. In both cases, these are aggregated because any estimates beyond this window would be identified off of a very small set of hospitals.

Figure 5: Reduced Form Event Study (Influenza Diagnoses)
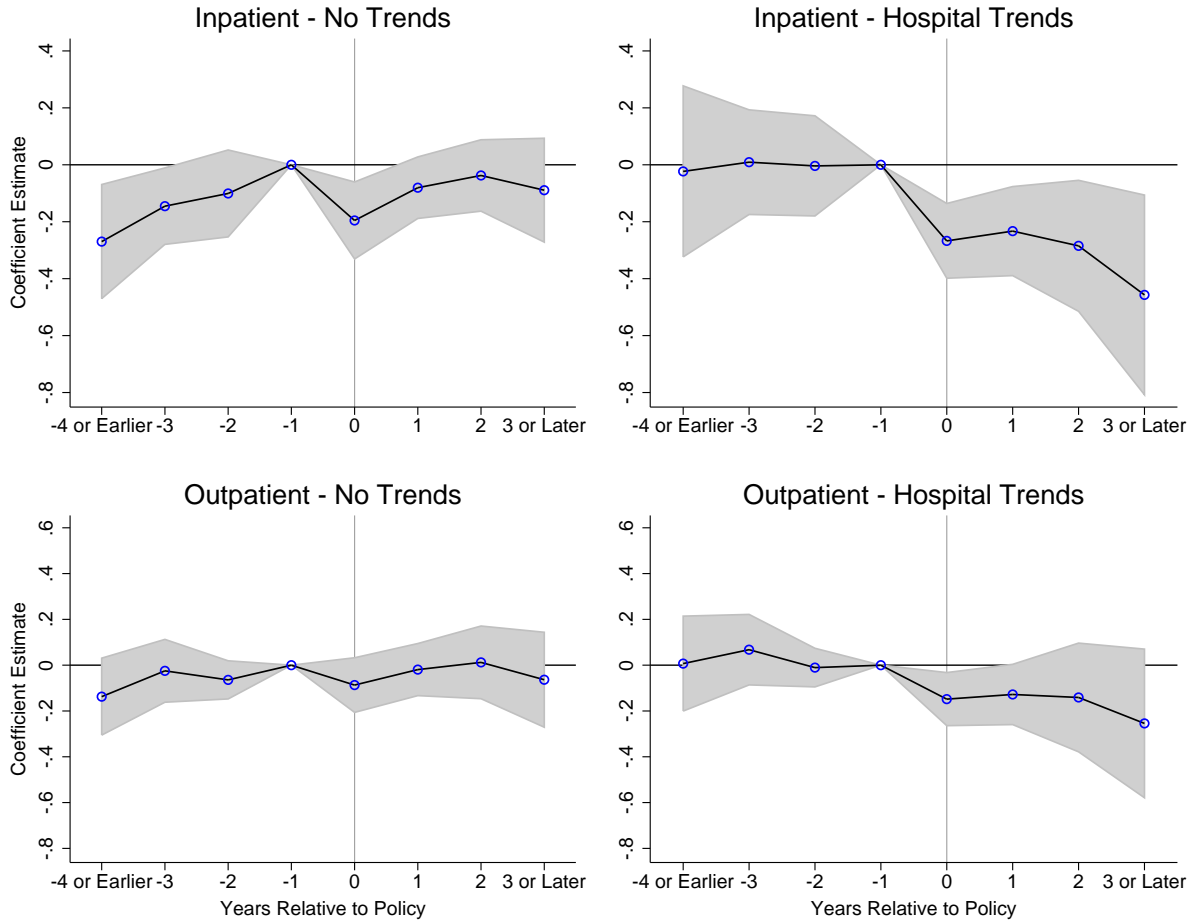


Points on the plots represent the point estimates from event-study versions of Equation (5) with influenza-related inpatient admissions or outpatient ED visits as the outcome. The plot labelled "Hospital Trends" includes hospital-specific linear time trends. Shaded regions represent 95% confidence intervals. The event-study is estimated by replacing the policy indicator ($\text{Required}_{hy}$) with a series of variables indicating years relative to the policy: $\sum_{j=-4}^{-2} \gamma_j \text{Required}_{hyj} + \sum_{j=0}^{3} \gamma_j \text{Required}_{hyj}$. The indicator representing one year prior to the policy is omitted as the reference group. "-4 or Earlier" represents four or more years prior to policy implementation; "3 or Later" represents three or more years after policy implementation. In both cases, these are aggregated because any estimates beyond this window would be identified off of a very small set of hospitals.

Table 1: Summary Statistics

| National Data | Mean | SD | |
|---|---|---|---|
| Vaccination Rate | 0.351 | (0.061) | - |
| Vaccination Rate $\geq 75$ | 0.712 | (0.077) | - |
| Vaccination Rate $< 75$ | 0.322 | (0.062) | - |
| Match Rate | 0.726 | (0.288) | - |
| Influenza Activity (1993-2015) | 0.185 | (0.214) | - |
| PI Mortality Rate | 6.13 | (2.12) | - |
| R&C Mortality Rate | 35.43 | (11.33) | - |
| Non-R&C Mortality Rate | 36.71 | (9.42) | - |
| All-Cause Mortality Rate | 72.14 | (12.57) | - |
| Hours Absent (Illness) | 0.245 | (0.119) | - |
| % Absent (Illness) | 0.025 | (0.009) | - |
| Hours Absent (Other) | 0.928 | (0.523) | - |
| % Absent (Other) | 0.093 | (0.045) | - |
| California Hospital Data | Mean | SD | # Hospitals Affected |
| HCW Vaccination Rate (All Hospitals/Years) | 0.801 | 0.142 | - |
| HCW Vaccination Rate (No Mandate) | 0.744 | (0.141) | - |
| HCW Vaccination Rate (Mandate) | 0.902 | (0.069) | - |
| # PI Diagnoses (Inpatient) | 21.81 | (36.68) | |
| # PI Diagnoses (Inpatient POA) | 21.28 | (35.67) | |
| # PI Diagnoses (Inpatient Not POA) | 0.51 | (1.57) | |
| # PI Diagnoses (Outpatient ED) | 135.3 | (170.6) | |
| Average Length of Stay | 5.43 | (4.04) | - |
| Average Charges | 42,294 | (23,057) | - |
| Required 2009-10 | - | - | 13 |
| Required 2010-11 | - | - | 18 |
| Required 2011-12 | - | - | 45 |
| Required 2012-13 | - | - | 116 |
| Required 2013-14 | - | - | 251 |
| Required 2014-15 | - | - | 335 |
| Required 2015-16 | - | - | 341 |

Zero hospitals had vaccination mandates prior to the 2009-10 season. In the 2015-16 influenza season, the total number of hospitals in California was 450 (so that 75.8% were subject to a vaccination mandate). The total number of hospitals fluctuated between 444 (2011) and 456 (2007) over the sample period.

Table 2: Effect of the Match Rate on Vaccination Rates

|                        | (1)     | (2)      |
|------------------------|---------|----------|
| Match                  | 0.310   | -0.163   |
|                        | (0.195) | (1.923)  |
| Match × Mean Vacc.     | -       | 0.014    |
|                        |         | (0.055)  |
| N                      | 969     | 969      |

The outcome in these regressions is the vaccination rate (which varies at the state-by-flu-year level), and the regressor is the match rate (which varies only at the flu-year level). The regressions are estimated at the state-by-flu-year level. 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 in parentheses are clustered at the state level.

Table 3: Mortality and Absences – Diff-in-Diff

Panel A: Pneumonia/Influenza (PI) Mortality Rate (per 100,000)

|  | All Months | Flu Season | Non Season | High Flu Months | Low Flu Months |
|---|---|---|---|---|---|
| Vacc × Match | -0.047 | -0.088 | -0.027 | -0.175 | -0.024 |
|  | (0.011) | (0.020) | (0.009) | (0.037) | (0.009) |
| Vacc | 0.004 | 0.029 | -0.006 | 0.105 | -0.014 |
|  | (0.015) | (0.021) | (0.014) | (0.036) | (0.014) |
| N | 12,240 | 4,080 | 8,160 | 1,530 | 10,710 |

Panel B: Average Work Hours Absent for Illness

|  | All Months | Flu Season | Non Season | High Flu Months | Low Flu Months |
|---|---|---|---|---|---|
| Vacc × Match | -0.0014 | -0.0045 | 0.0001 | -0.0095 | -0.0002 |
|  | (0.0009) | (0.0016) | (0.0009) | (0.0029) | (0.0010) |
| Vacc | 0.0026 | 0.0058 | 0.0012 | 0.0088 | 0.0016 |
|  | (0.0008) | (0.0014) | (0.0007) | (0.0024) | (0.0007) |
| N | 12,240 | 4,080 | 8,160 | 1,530 | 10,710 |

"Flu Season" represents December-March; "Non Season" represents April-November. "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. Standard errors in parentheses are clustered at the state level.

Table 4: Mortality and Absences – Triple Difference

| Panel A: Pneumonia/Influenza (PI) Mortality Rate (per 100,000) | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Vacc × Match × Activity | -0.184 | -0.223 | -0.199 |
| | (0.059) | (0.052) | (0.061) |
| Vacc × Match | 0.027 | -0.009 | -0.011 |
| | (0.017) | (0.008) | (0.010) |
| Expected Annual Benefit (Deaths per 100,000 population) | -0.289 | -0.351 | -0.313 |
| N | 12,240 | 12,240 | 10,404 |

| Panel B: Average Work Hours Absent for Illness | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Vacc × Match × Activity | -0.0098 | -0.0106 | -0.0099 |
| | (0.0042) | (0.0042) | (0.0039) |
| Vacc × Match | -0.0001 | 0.0004 | 0.0003 |
| | (0.0011) | (0.0010) | (0.0011) |
| Expected Annual Benefit (Hours per 100,000 population) | -2,563 | -2,762 | -2,578 |
| N | 12,240 | 12,240 | 10,404 |

| | | | |
|---|---|---|---|
| Month-Year Fixed Effects | X | X | X |
| Weather Controls | X | X | X |
| State-Month Fixed Effects | - | X | X |
| IV | - | - | X |

All regressions also include the $Vacc \times Activity$ interaction and the main effect for $Vacc$; not included are the $Match \times Activity$ interaction and the main effects for $Match$ and $Activity$ as these are absorbed by the month-year fixed effects. The "Expected Annual Benefit" for mortality is equal to $\sum_m \hat{\phi}_1 \times \overline{Match} \times \overline{Activity}_m$, where $\phi_1$ is the coefficient on the triple interaction. This measures the expected annual reduction in mortality that would be expected to result from a one percentage point increase in the vaccination rate for a population of 100,000 (i.e., 1,000 additional vaccinations). For hours absent, the "Expected Annual Benefit" is equal to $\sum_m \hat{\phi}_1 \times \overline{Match} \times \overline{Activity}_m \times (30.5/7) \times (126/323) \times 100,000$. (30.5/7) represents the number of weeks per month, since the coefficient measures the change in mean weekly hours lost, and (126/323) represents the ratio of full time workers to the population in the U.S. Finally, the calculation for the "Expected Annual Benefit" for absences requires multiplying by 100,000 since the outcome is mean hours lost rather than a rate per 100,000 population (as in the mortality specification). This measures the expected annual reduction in hours lost among full time workers that would be expected to result from a one percentage point increase in the vaccination rate for a population of 100,000 (i.e., 1,000 additional vaccinations). The "IV" specification indicates that the average vaccination rate over the three years prior is used as an instrument for the current year's vaccination rate. Standard errors in parentheses are clustered at the state level.

Table 5: Mortality by Age & Decomposing Externality

Panel A: All-Age Vaccination Rates & Age-Specific Mortality

| | Under 1 | Age 1-9 | Age 10-64 | Age 65-74 | Age 75+ |
|---|---|---|---|---|---|
| D-D-D Effect | -0.002 | -0.000 | -0.017 | -0.015 | -0.189 |
| | (0.002) | (0.001) | (0.011) | (0.014) | (0.043) |
| | [1%] | [0%] | [8%] | [7%] | [85%] |
| N | 12,240 | 12,240 | 12,240 | 12,240 | 12,240 |

Panel B: Age-Specific Vaccination Rates & Age-Specific Mortality

| | Under 1 | Age 1-9 | Age 10-64 | Age 65-74 | Age 75+ |
|---|---|---|---|---|---|
| D-D-D Effect (75+) | - | - | - | - | -0.058 |
| | | | | | (0.028) |
| | | | | | [34%] |
| D-D-D Effect (<75) | - | - | - | - | -0.113 |
| | | | | | (0.045) |
| | | | | | [66%] |
| N | | | | | 12,240 |

Age-specific mortality rates are calculated as the number of deaths per 100,000 all-age 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 A4 in which the denominator is age-specific. In Panel A, coefficient estimates represent estimates of the triple-interaction from Equation (3). In Panel B, coefficient estimates represent estimates of the triple-interactions from Equation (4). The percentage of total mortality reductions attributable to vaccination in each age group is reported in brackets (the corresponding coefficient divided by the sum of the two coefficients). Standard errors in parentheses are clustered at the state level.

Table 6: Monetized Benefits

**Panel A: Mortality (EPA VSL)**

| | Age-Adjusted VSL | Number of Deaths (Scaled Nationally) | Monetized Value (Scaled Nationally) | Value Per Vaccination |
|---|---|---|---|---|
| Age Under 10 | $8,705,051 | 13.1 | $114,036,168 | $35.30 |
| Age 10-64 | $6,470,611 | 86.1 | $557,119,607 | $172.43 |
| Age 65-74 | $2,866,840 | 73.1 | $209,566,004 | $64.86 |
| Age Over 74 | $975,689 | 933.9 | $911,195,957 | $282.02 |
| Total | - | 1,106.2 | $1,791,917,736 | $554.60 |

**Panel B: Mortality (AG VSL)**

| | Age-Adjusted VSL | Number of Deaths (Scaled Nationally) | Monetized Value (Scaled Nationally) | Value Per Vaccination |
|---|---|---|---|---|
| Age Under 10 | $2,285,076 | 13.1 | $29,934,495 | $9.26 |
| Age 10-64 | $1,698,535 | 86.1 | $146,243,863 | $45.26 |
| Age 65-74 | $752,545 | 73.1 | $55,011,039 | $17.03 |
| Age Over 74 | $256,118 | 933.9 | $239,188,600 | $74.00 |
| Total | - | 1,106.2 | $470,377,997 | $145.58 |

**Panel C: Work Absences**

| | Median Hourly Wage | Number of Hours (Scaled Nationally) | Monetized Value (Scaled Nationally) | Value Per Vaccination |
|---|---|---|---|---|
| All Ages | $20.80 | 8,925,508 | $185,650,574 | $57.46 |

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 specification in Column 2 of Table 4. The median hourly wage is calculated as the median weekly wage in 2016 ($832) divided by median hours worked (40) for full-time workers; these figures are derived from the Bureau of Labor Statistics https://www.bls.gov/news.release/wkyeng.t07.htm.

Table 7: Effects of HCW Mandates on Influenza Diagnoses

Panel A: Inpatient Admissions with Influenza Diagnosis

| | Specification Checks | | | | | Present on Admission | |
| | (1) | (2) | (3) | (4) | (5) | POA | Not-POA |
|---|---|---|---|---|---|---|---|
| Required | -0.0708 | -0.1000 | -0.214 | -0.234 | -0.207 | -0.209 | -0.402 |
| | (0.0389) | (0.0459) | (0.0687) | (0.0699) | (0.0674) | (0.0685) | (0.0984) |
| N | 3,609 | 3,208 | 3,208 | 3,208 | 2,406 | 3,208 | 3,208 |
| Converged | Yes | Yes | Yes | No | No | Yes | Yes |

Panel B: Outpatient ED Visits with Influenza Diagnosis

| | Specification Checks | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | | |
|---|---|---|---|---|---|---|---|
| Required | -0.0168 | -0.0243 | -0.0944 | -0.122 | -0.0928 | - | - |
| | (0.0470) | (0.0458) | (0.0544) | (0.0538) | (0.0594) | | |
| N | 2,700 | 2,400 | 2,400 | 2,400 | 1,800 | | |
| Converged | Yes | Yes | Yes | Yes | Yes | | |
| | | | | | | | |
| Exclude 2014-15 | - | X | X | X | X | X | X |
| County Linear Trends | - | - | X | - | X | X | X |
| Hospital Linear Trends | - | - | - | X | - | - | - |
| Exclude H1N1 Years | - | - | - | - | X | - | - |

Reported coefficient estimates are derived from negative binomial regression models, and as such the estimates can be approximately interpreted as percent changes. Regressions are estimated at the hospital-by-year level. The smaller number of observations for outpatient ED visits is due to the smaller number of emergency departments relative to inpatient hospitals. "Converged" indicates whether the maximization algorithm converged; non-convergent specifications are those that require the estimation of many variables (i.e., hospital-specific trends) or have relatively few observations. Standard errors in parentheses are clustered at the county level.

# Online Appendix

Figure A1: Vaccination and Influenza Timing
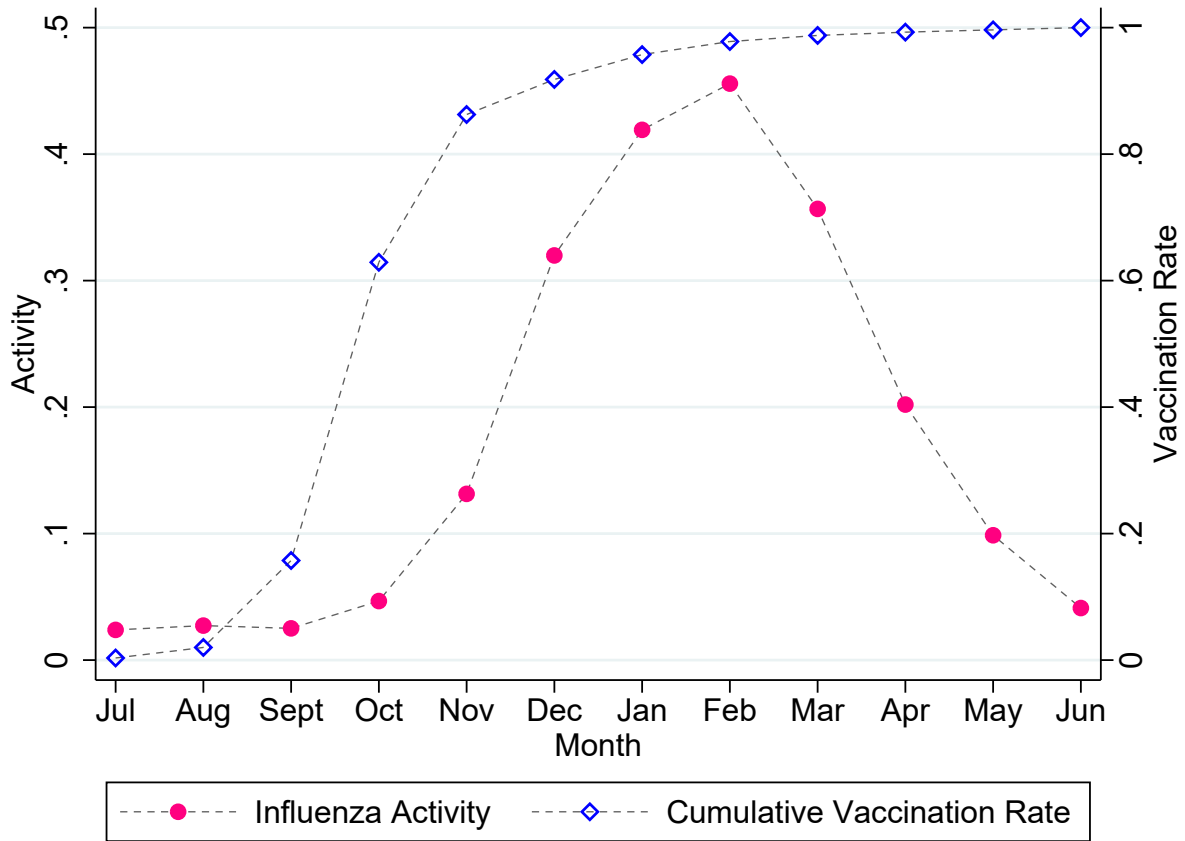
Figure A2: Nonlinearities
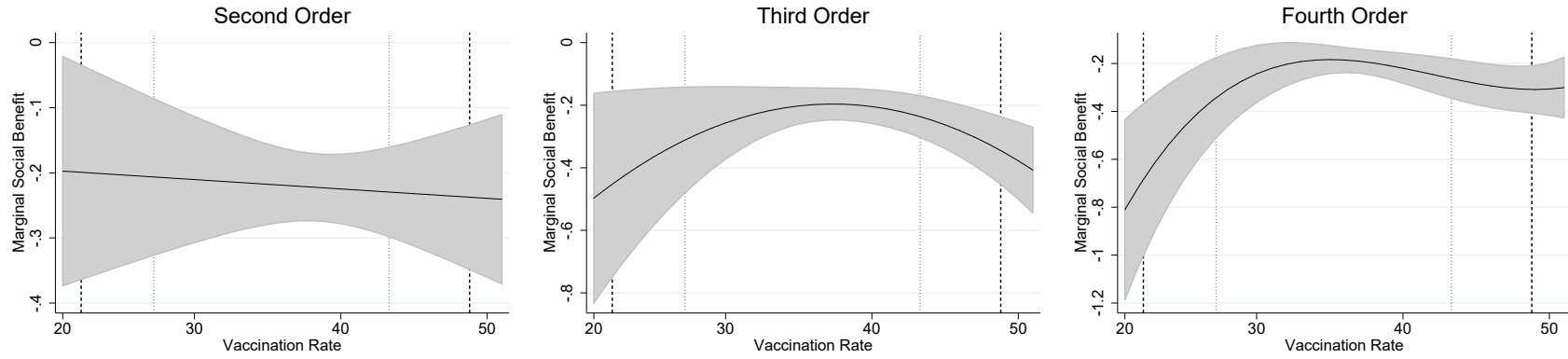
Note – These plots test for nonlinearities in the effects of state-level vaccination rates by allowing for higher order polynomials in the triple interaction specification described in Equation (3). The plots represent the marginal effects of vaccination at various vaccination rates; note that larger negative numbers imply larger social benefits. Dashed lines represent the 1st and 99th percentiles in the distribution of vaccination rates, and dotted lines represent the 10th and 90th percentiles. Shaded regions represent 95% confidence intervals.

Figure A3: California Mandates
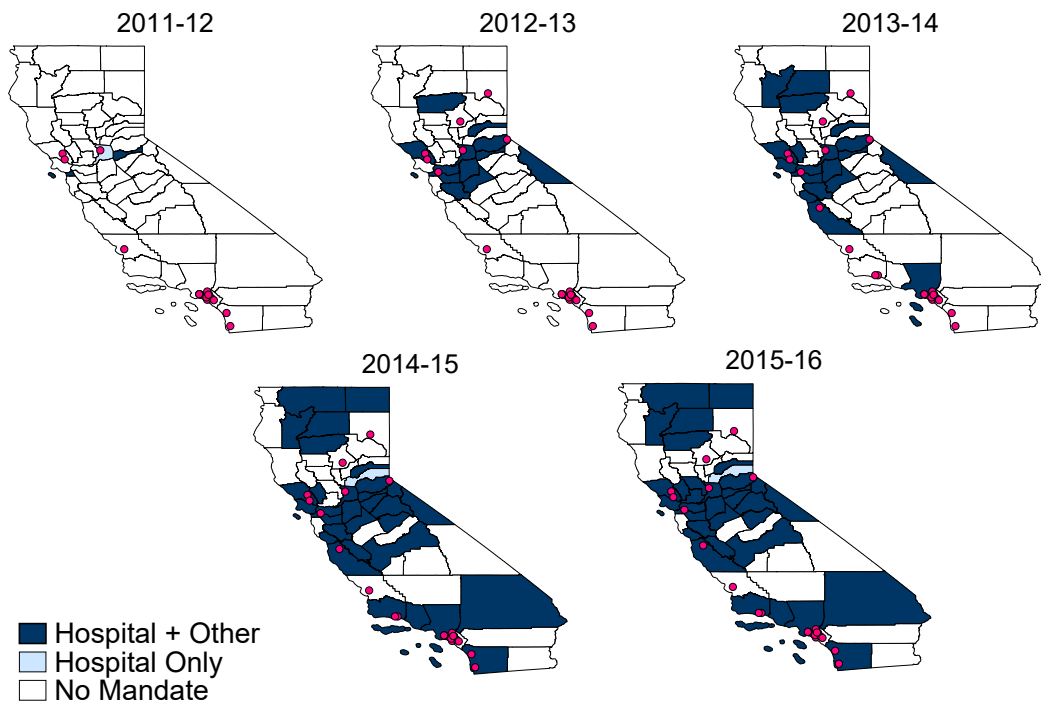


These plots display the roll-out of influenza vaccination mandates. Circles represent policies implemented at the hospital level and shaded regions represent policies implemented at the county level. The lighter shaded regions represent county-level policies that apply only to hospitals, and the darker regions represent county-level policies that apply more broadly.

### Table A1: Mortality and Absences – Specification Checks

**Panel A: PI Mortality**

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| D-D-D Effect | -0.131 | -0.178 | -0.146 | -0.164 | -0.127 | -0.226 | -0.211 |
|  | (0.079) | (0.052) | (0.066) | (0.050) | (0.062) | (0.050) | (0.056) |
| N | 10,404 | 12,240 | 10,404 | 12,240 | 10,404 | 12,240 | 10,404 |

**Panel B: Hours Absent for Illness**

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| D-D-D Effect | -0.011 | -0.010 | -0.009 | -0.010 | -0.010 | -0.011 | -0.011 |
|  | (0.004) | (0.005) | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) |
| N | 10,404 | 12,240 | 10,404 | 12,240 | 10,404 | 12,240 | 10,404 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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 |

Estimates are reported for a variety of specifications distinct from those presented as the main results in Table 4. Standard errors in parentheses are clustered at the state level.

## Table A2: Mortality and Absences - IV Specification

**Panel A: PI Mortality**

|  | t-1 | t-1 to t-2 (Avg.) | t-1 to t-3 (Avg.) | t-1 to t-4 (Avg.) | Base Period (3 Year Avg.) |
|---|---|---|---|---|---|
| D-D-D Effect | -0.201 | -0.201 | -0.199 | -0.197 | -0.171 |
|  | (0.059) | (0.057) | (0.061) | (0.065) | (0.061) |
| N | 11,628 | 11,016 | 10,404 | 9,792 | 10,404 |

**Panel B: Hours Absent for Illness**

|  | t-1 | t-1 to t-2 (Avg.) | t-1 to t-3 (Avg.) | t-1 to t-4 (Avg.) | Base Period (3 Year Avg.) |
|---|---|---|---|---|---|
| D-D-D Effect | -0.011 | -0.011 | -0.010 | -0.008 | -0.007 |
|  | (0.005) | (0.005) | (0.004) | (0.005) | (0.005) |
| N | 11,628 | 11,016 | 10,404 | 9,792 | 10,404 |
| IV | X | X | X | X | - |
| Reduced Form | - | - | - | - | X |

These estimates test the sensitivity of the IV estimates to the definition of the instrument for vaccination rates. Columns 1-4 use the vaccination rate in the prior year, the average vaccination rate over the prior two years, the average vaccination rate in the prior three years (the specification presented in Table 4), and the average vaccination rate in the prior four years. Column 5 uses a time-invariant base period vaccination rate (the average vaccination rate over the first three years of the sample). Because the IV strategy requires instrumenting for three interactions plus main effect of vaccination, using a time-invariant instrument in the presence of state-by-month fixed effects means that one of the required instruments drops out. As such, these estimates are reported as reduced-form estimates (i.e., the effect of base period vaccination on the outcome at time $t$), and the magnitudes are not directly comparable. Standard errors in parentheses are clustered at the state level.

Table A3: Mortality by Cause and Absences by Reason

Panel A: PI Mortality

| | PI Mortality | R&C Mortality | All Mortality | Non-R&C Mortality (Falsification) |
|---|---|---|---|---|
| D-D-D Effect | -0.223 | -0.320 | -0.365 | -0.045 |
| | (0.052) | (0.106) | (0.147) | (0.083) |
| N | 12,240 | 12,240 | 12,240 | 12,240 |
| Mean Dep. Var. | 6.07 | 35.43 | 72.14 | 36.71 |

Panel B: Hours Absent for Illness

| | Hours Absent - Illness | % Absent - Illness | Hours Absent - Other (Falsification) | % Absent - Other (Falsification) |
|---|---|---|---|---|
| D-D-D Effect | -0.01061 | -0.00065 | -0.01120 | -0.00051 |
| | (0.00421) | (0.00024) | (0.01256) | (0.00097) |
| N | 12,240 | 12,240 | 12,240 | 12,240 |
| Mean Dep. Var. | 0.245 | 0.025 | 0.928 | 0.093 |

All estimates are from models that use the main triple-difference specification (described in column 2 of Table 4). For both mortality and work absences, the first column duplicates the main estimates as reference. "R&C" refers to respiratory & circulatory mortality, which is a level of aggregation higher than pneumonia and influenza (PI). Deaths with no respiratory or circulatory diagnosis are unlikely to be related to influenza infection (column 4 of Panel A). "% Absent" provides an alternate definition for work absence: the proportion of workers reporting any work absence rather than the number of hours lost. In columns 3-4, estimates are reported for non-illness work absences (e.g., vacation). Standard errors in parentheses are clustered at the state level.

Table A4: Mortality – Age Specific Mortality Rates

|  | Under 1 | Age 1-9 | Age 10-64 | Age 65-74 | Age 75+ |
|---|---|---|---|---|---|
| D-D-D Effect | -0.195 | -0.001 | -0.024 | -0.206 | -2.184 |
|  | (0.137) | (0.011) | (0.015) | (0.216) | (0.620) |
| Mean Dep. Var. | 1.35 | 0.08 | 1.20 | 13.11 | 71.09 |
| N | 12,240 | 12,240 | 12,240 | 12,240 | 12,240 |

in Table 5, age-specific mortality rates are calculated as the number of deaths per 100,000 *total* individuals in the population to facilitate straightforward accounting of the total mortality benefits of vaccination. Here, mortality rates are calculated as the number of deaths per 100,000 population in the relevant age group. These results confirm the main findings that the benefits are concentrated in the 75+ age group. Standard errors in parentheses are clustered at the state level.

Table A5: Mortality and Absences - Lagged Impacts

**Panel A: PI Mortality**

|  | Baseline (One Month) | Two Month | Three Month | Four Month | Five Months |
|---|---|---|---|---|---|
| D-D-D Effect | -0.2209 | -0.2509 | -0.1755 | -0.1995 | -0.1918 |
|  | (0.0513) | (0.0571) | (0.0545) | (0.0612) | (0.0648) |
| N | 12,036 | 12,036 | 12,036 | 12,036 | 12,036 |

**Panel B: Hours Absent for Illness**

|  | Baseline (One Month) | Two Month | Three Month | Four Month | Five Months |
|---|---|---|---|---|---|
| D-D-D Effect | -0.01124 | -0.01185 | -0.00970 | -0.01565 | -0.01432 |
|  | (0.00481) | (0.00504) | (0.00592) | (0.00690) | (0.00748) |
| N | 12,036 | 12,036 | 12,036 | 12,036 | 12,036 |

These estimates test whether the contemporaneous month is sufficient to capture the full extent of influenza-related mortality. The column labelled "Two Months" reports estimates that replicate the main estimates, 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. The column labelled "Three Months" adds one and two month lags in the relevant interactions and the sum of all three coefficients are reported, and so on for columns four and five. Standard errors in parentheses are clustered at the state level.

Table A6: Mortality and Absences – Robustness

**Panel A: PI Mortality**

|  | Exclude Interp. Years | Include H1N1 Years | 1993-2007 Sample | Regional Match (1998-2014) | Leave-One-Out (1998-2014) | Flu Season Triple |
|---|---|---|---|---|---|---|
| D-D-D Effect | -0.225 | -0.216 | -0.174 | -0.200 | -0.202 | -0.061 |
|  | (0.056) | (0.048) | (0.059) | (0.053) | (0.057) | (0.017) |
| Scaled Effect | - | - | - | - | - | -0.188 |
| N | 10,404 | 13,464 | 8,568 | 9,180 | 9,180 | 12,240 |

**Panel B: Hours Absent for Illness**

|  | Exclude Interp. Years | Include H1N1 Years | 1993-2007 Sample | Regional Match (1998-2014) | Leave-One-Out (1998-2014) | Flu Season Triple |
|---|---|---|---|---|---|---|
| D-D-D Effect | -0.00974 | -0.00892 | -0.01423 | -0.00873 | -0.01021 | -0.00470 |
|  | (0.00455) | (0.00433) | (0.00629) | (0.00428) | (0.00406) | (0.00194) |
| Scaled Effect | - | - | - | - | - | -0.01450 |
| N | 10,404 | 13,464 | 8,568 | 9,180 | 9,180 | 12,240 |

All estimates are variants on models that use the main triple-difference specification (described in column 2 of Table 4). Years refer to flu-years; for example, "1993" refers to the 1993-94 flu-year. In column 1, four influenza seasons with interpolated vaccination rates are excluded. Column 2 includes the two flu-years affected by the H1N1 pandemic. Column 3 represents a sample period prior to the H1N1 pandemic and two technological developments in the influenza vaccine (high-dose vaccines and the quadrivalent vaccine). Columns 4 uses regional variation in the match rate (also included in this regression are the main effect for Match and the Match×Activity interaction, which are no longer absorbed by fixed effects). Column 5 uses average influenza activity in all census divisions other than that which includes the corresponding state (also included in this regression are the main effect Activity and the Match×Activity interaction, which are no longer absorbed by fixed effects). Column 6 uses an influenza season indicator in place of influenza activity in the triple difference; for this specification a "Scaled Effect" is provided for comparability with the other estimates in which the point estimate is scaled by a factor of $1/(\bar{A}_{season} - \bar{A}_{off})$, where $\bar{A}_{season}$ and $\bar{A}_{off}$ represent average influenza activity during influenza season and during the off-season, respectively. Standard errors in parentheses are clustered at the state level.

Table A7: HCW Vaccination Policy Timing

| Hospital | Season | County | Season |
|---|---|---|---|
| Children's of Orange | 2009 (H1N1) | Sacramento | 2011-12 |
| Community Hospital of LB | 2009 (H1N1) | San Francisco | 2011-12 |
| Hoag Hospitals | 2009 (H1N1) | Alameda | 2012-13 |
| Long Beach Memorial | 2009 (H1N1) | Amador | 2012-13 |
| Miller Children's | 2009 (H1N1) | Contra Costa | 2012-13 |
| Orange Coast Memorial | 2009 (H1N1) | El Dorado | 2012-13 |
| Pacific Hospital of LB | 2009 (H1N1) | Mono | 2012-13 |
| St. Joseph (Orange) | 2009 (H1N1) | Nevada | 2012-13 |
| St. Jude (Fullerton) | 2009 (H1N1) | San Joaquin | 2012-13 |
| UC Davis | 2009 (H1N1) | Santa Clara | 2012-13 |
| UC Irvine | 2009 (H1N1) | Stanislaus | 2012-13 |
| UC San Diego | 2009 (H1N1) | Sonoma | 2012-13 |
| Saddleback Memorial | 2009 (H1N1) | Tehama | 2012-13 |
| Santa Rosa Memorial | 2010-11 | Santa Cruz | 2013-14 |
| Sierra Vista (SLO) | 2010-11 | Los Angeles | 2013-14 |
| Tri-City (Oceanside) | 2010-11 | Marin | 2013-14 |
| Petaluma Valley Hospital | 2010-11 | Monterey | 2013-14 |
| Oroville Hospital | 2012-13 | Napa | 2013-14 |
| Banner Lassen Medical Center | 2012-13 | Shasta | 2013-14 |
| Barton Memorial | 2012-13 | Trinity | 2013-14 |
| UCSF (Children's - Oakland) | 2012-13 | Alpine | 2014-15 |
| Cottage Hospitals | 2013-14 | Calaveras | 2014-15 |
| Salinas Valley Hospital | 2013-14 | Fresno | 2014-15 |
| - | - | Mariposa | 2014-15 |
| - | - | Modoc | 2014-15 |
| - | - | Placer | 2014-15 |
| - | - | San Benito | 2014-15 |
| - | - | San Bernardino | 2014-15 |
| - | - | Santa Barbara | 2014-15 |
| - | - | Siskiyou | 2014-15 |
| - | - | Tuolomne | 2014-15 |
| - | - | Ventura | 2014-15 |
| - | - | Solano | 2015-16 |
| - | - | Yolo | 2015-16 |
| - | - | Humboldt | 2016-17 |
| - | - | Mendocino | 2016-17 |
| - | - | Butte | 2017-18 |
| - | - | Merced | 2017-18 |
| - | - | San Luis Obispo | 2017-18 |

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 A8:  Effects of HCW Mandates on Hospital Worker Vacc. Rates

|  | (1) | (2) | (3) |
|---|---|---|---|
| Required | 0.106 | 0.0928 | 0.0960 |
|  | (0.0145) | (0.0115) | (0.0119) |
| N | 707 | 1,391 | 2,568 |
| # Hospitals | 101 | 202 | 387 |

The three columns represent different levels of stringency in selecting the sample for the first stage, based on data quality. Hospitals are required to report the percentage of workers vaccinated in each influenza season; poor quality data emerges when hospitals do not collect this information for every worker (the response rate for a particular hospital and influenza season may be less than 100%). Column 1 reports estimates only from hospitals with a response rate of at least 90% in all years (2009/10-2015/16); column 2 reports estimates only from hospitals with a response rate of at least 90% in all but one year (typically the first year); column 3 reports estimates from all hospitals regardless of data quality. Standard errors in parentheses are clustered at the county level.

Table A9: Effects of HCW Mandates by Age

**Influenza Diagnoses – Inpatient Admissions**

|  | Under 1 | 1-9 | 10-64 | 65-74 | 75+ |
|---|---|---|---|---|---|
| Required | -0.249 | -0.242 | -0.171 | -0.118 | -0.121 |
|  | (0.0716) | (0.0898) | (0.0845) | (0.0933) | (0.0655) |
| N | 1,936 | 1,576 | 3,192 | 3,032 | 2,912 |
| Converged | Yes | No | Yes | No | Yes |

**Influenza Diagnoses – Outpatient ED Visits**

|  | Under 1 | 1-9 | 10-64 | 65-74 | 75+ |
|---|---|---|---|---|---|
| Required | -0.109 | -0.217 | -0.0819 | -0.0700 | -0.0704 |
|  | (0.0867) | (0.0854) | (0.0534) | (0.0614) | (0.0980) |
| N | 2,384 | 2,384 | 2,400 | 2,376 | 2,360 |
| Converged | Yes | Yes | Yes | Yes | No |
| Exclude 2014-15 | X | X | X | X | X |
| County Linear Trends | X | X | X | X | X |

Reported coefficient estimates are derived from negative binomial regression models, and as such the estimates can be approximately interpreted as percent changes. Regressions are estimated at the hospital-by-year level. Hospitals with zero age-specific influenza diagnoses in all years are automatically omitted, accounting for the difference in sample size across age groups. "Converged" indicates whether the maximization algorithm converged. Standard errors in parentheses are clustered at the county level.

Table A10: Effects of HCW Mandates – Other Outcomes

| | ln(Avg. Length of Stay) | ln(Avg. Charges) | In-Hospital Death Rate | PI Mortality Rate |
|---|---|---|---|---|
| Required×Activity | -0.0126 | -0.0203 | -0.0000842 | -0.243 |
| | (0.00568) | (0.0105) | (0.000442) | (0.280) |
| Required | -0.00108 | 0.0103 | 0.0000462 | 0.0746 |
| | (0.00326) | (0.00674) | (0.000423) | (0.0947) |
| Exclude 2014-15 | X | X | X | X |
| County Linear Trends | X | X | X | X |
| Hospital-Level | X | X | X | - |
| County-Level | - | - | - | X |
| N | 36,184 | 32,448 | 36,192 | 5,504 |

The estimates presented in columns 1-3 represent average outcomes for inpatient hospital admissions at the hospital-year-month level. The estimate in column 4 uses data on mortality at the county-year-month level. The coefficient estimate for "Required×Activity" represents the impact of HCW vaccination mandates during periods with high influenza activity relative to zero activity, and the estimate for "Required" represents the impact during periods with zero influenza activity. The distributions for length of stay and charges at the micro-level (i.e., before collapsing to the hospital-year-month level) have extremely long tails. To ensure that the 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 (approximately 13% of the hospitals in the sample). This accounts for the smaller number of observations for the estimates of charges. Furthermore, length of stay is unreported for approximately 1% of admissions; averages cannot be calculated for these outcomes when all hospital-year-month outcomes are missing – this typically only occurs when there is a single observation in that cell. All models are estimated via OLS. Standard errors in parentheses are clustered at the county level.