Paid Sick Leave and Absenteeism: The First Evidence from the U.S.

Thomas Ahn
Department of Economics
University of Kentucky

Aaron Yelowitz*
Department of Economics
University of Kentucky

August 12, 2016

Abstract: Using health shocks arising from varying severity of influenza across geography and time, with a balanced sample of workers from cross-sections of the National Health Interview Survey, we estimate the causal effects of paid sick leave (PSL) insurance on absenteeism in the United States. Data from Google Flu Trends and measures of moderate and severe sick days reveal that a substantive portion of the estimated impact of PSL on absenteeism (~1.2 days per year) is attributable to moral hazard. Estimates with workers of different occupations and firm types help to inform potential public health benefits and impacts to firm profitability.

Keywords: Paid Sick Leave, Absenteeism, Presenteeism, Moral Hazard, Google Flu Trends
JEL Classification: H51, I18, J22

---

* Corresponding author.

Contact information for Ahn: Department of Economics, University of Kentucky, Gatton School of Business and Economics, 550 South Limestone Street, Lexington, KY 40506; (telephone) (859) 257-5975; thomas.ahn@uky.edu (email).

Contact information for Yelowitz: Department of Economics, University of Kentucky, Gatton School of Business and Economics, 550 South Limestone Street, Lexington, KY 40506; (859) 257-7634 (telephone); aaron@uky.edu (email); www.yelowitz.com (website).

We thank Glenn Blomquist, Frank Scott, Bill Hoyt, John Garen, Nicolas Ziebarth and Tim Harris for helpful comments. We also thank Matthis Schendstok for research assistance. Any errors are, of course, our own.

I. Introduction

Paid sick leave (PSL) in the United States has recently received a flurry of attention at the federal, state, and local levels. President Obama has called for a federal bill that "gives every worker in America the opportunity to earn seven days of paid sick leave."[1] Several large cities and states have recently mandated that firms offer PSL to employees. For example, California's *Healthy Workplaces, Healthy Families Act of 2014* enables workers to accrue one hour of PSL for every 30 hours worked from July 2015 onward; it is expected to affect more than 6.5 million employees with no paid sick days, or roughly 40 percent of state's labor force.[2] PSL campaigns were active in 22 states in 2014 (National Partnership for Women & Families, 2015), and when put to the ballot, have been approved by voters by wide margins.[3] The Council of Economic Advisers has recently published a report advocating for increased access paid sick and family leave, arguing that it would improve retention and raise productivity (CEA, 2014).

The reasons for the current policy interest garnered by government sick leave mandates are many. Worker sickness has negative implications for not only the productivity of the employee him/herself, but may also create negative externalities for coworkers and/or customers that the sick employee interacts with. The largely voluntary nature of paid sick leave offering by firms may lead to sorting of workers who are most willing to use/abuse PSL towards generous

---

[1] Remarks by President Obama in State of the Union Address, January 20, 2015. See https://www.whitehouse.gov/the-press-office/2015/01/20/remarks-president-state-union-address-january-20-2015 .

[2] Press Release "Governor Brown Signs Legislation to Provide Millions of Californians with Paid Sick Leave." See http://gov.ca.gov/news.php?id=18690 .

[3] The first PSL mandate was passed in San Francisco, CA in 2006. Other large cities with the mandate include New York, NY, Washington, DC, Seattle, WA, Portland, OR, and San Diego, CA. Both Connecticut and Massachusetts have passed state-level mandates. See http://www.nationalpartnership.org/research-library/campaigns/psd/state-and-local-action-paid-sick-days.pdf .

firms, possibly leading to a race-to-the-bottom as firms withdraw PSL as a benefit. The cost of providing PSL is perceived to be fairly inexpensive, since many workers do not use the full allotment of paid sick days offered to them; the U.S. Bureau of Labor Statistics reports that average cost for sick leave per employee was $0.23/hour (BLS, 2013). It is also a benefit that is not widely provided in the U.S. economy, especially for low wage workers. Roughly 80 percent of low wage workers in the U.S. are not guaranteed paid time off due to illness. Indeed, the U.S lags behind many other nations in mandating that employers offer PSL (Heymann et al., 2009).

Much of our current understanding of the impact of PSL is based on studies from Northern Europe. These studies find that more generous benefits lead to a modest increase in work absenteeism (Johansson and Palme, 1996, 2002; Olsson, 2009; Ziebarth and Karlsson, 2010; Puhani and Sonderhof, 2010). As useful as the insights from these studies are, there may be difficulties in translating the results wholesale to the American experience. In contrast to the U.S., PSL in Europe is provided to all workers as national insurance. As such, the representative recipient of PSL benefit in the U.S. looks starkly different from the average worker in Europe. In addition, the large differences in design of PSL in the U.S. compared to Europe imply that the worker's utility optimization may operate at different margins. For example, German firms must offer six weeks of PSL with 100 percent replacement rate. Beyond the first six weeks, employees may receive 80 percent of their full salary up to 78 weeks (Ziebarth and Karlsson, 2014). In contrast, less than 1 percent of workers with PSL in the U.S. receive 6 weeks of paid sick days (U.S. Bureau of Labor Statistics, 2013). In addition, other labor market institutions – such as the structure of health insurance or worker protections against getting fired – may interact with PSL that lead to different impacts on absenteeism in the U.S.

This massive difference in generosity of the systems implies that the impact of PSL in the U.S. may be very different. On the one hand, if there is a large pent-up demand, even a small increase in PSL may have tremendous impact on workers, potentially leading to more pronounced changes in absenteeism behavior (whether for legitimate illness or as de facto vacation days). On the other hand, if most workers view PSL as insurance to be used only in emergencies, absence behavior for the majority of workers should change very little, and those who do need to use PSL will be limited to a small number of sick days due to the relative stinginess of PSL in the U.S. Then, the average change in absenteeism may be much smaller than what is found in European studies.

While these institutional differences call for a study of absenteeism response to PSL in the U.S., the literature has been largely silent until quite recently due to two large problems that we surmount in this study.[4] First, a data set must contain information on both whether a firm offered PSL and how many days the employee was absent from work. The most frequently used labor market datasets by economists in the U.S. fail to ask at least one of these questions. However, the National Health Interview Survey (NHIS) asks about both.[5] The NHIS is a large cross-sectional survey of workers that asks a battery of questions regarding health and employment status. Although these questions have been asked in the NHIS for many years, this study is the first to analyze them. Additionally, the NHIS asks respondents how many days were spent severely sick (i.e. bed-ridden), allowing for a distinction between debilitating/severe sick

[4] There are certainly studies on causes or consequences of absenteeism (or presenteeism), but such studies typically do not focus on PSL. See, for example, Cawley, Rizzo and Hass (2007, 2008) and Pauly, et al. (2008).
[5] Widely-used microdata, like the Panel Study of Income Dynamics, National Longitudinal Study of Youth, Current Population Survey, American Community Survey, and Survey of Income and Program Participation fail to ask sufficient questions to address this issue.

days and moderate/discretionary sick days. Moral hazard due to PSL should manifest itself in increased moderate days but not severe days. Although we cannot follow workers across time to capture employment dynamics (particularly if workers change jobs in response to whether a firm offers fringe benefits), the detailed questions and large sample sizes makes the NHIS useful to examine PSL in the U.S.

Second, a study of PSL in the U.S. faces identification difficulties. The European studies are usually analyzed in a difference-in-differences framework arising from a legislative change at a point in time. The lack of legislative mandates in the U.S. (until very recently) have made such identification strategies impossible.[6] In this study, we take a multi-pronged approach to identifying the moral hazard effects of PSL. Using rich information on the respondent's health status, employment, demographics, and firm characteristics, we create a balanced sample of administrative workers – such as secretaries and administrative assistants – who are offered or not offered PSL largely based their industry draw. While we expect that balancing the sample this way will go a long way toward mitigating the selection problem, it is still possible that workers will select into industries that tend to offer PSL which may also provide other job amenities or fringe benefits that affect absenteeism.

To further mitigate the selection problem, we take advantage of a truly unanticipatable, exogenous health shock to workers: the varying severity of influenza across geography and time. Even if workers select into jobs based on PSL, whether they will be compelled to use it due to the flu in a particular year will be truly random under reasonable assumptions. Workers facing a

[6] With recently enacted PSL legislation, empirical work on PSL in the U.S. has experienced something of a renaissance. See Ahn and Yelowitz (2015), Pichler (2014), Pichler and Ziebarth (2015), and Ziebarth and Pichler (2016).

severe flu season will be more likely to exhaust their PSL allowance, and their replacement rate will exogenously drop from 100 percent to zero. The difference in absenteeism behavior of workers with or without PSL in the "control region" (the severe flu outbreak region/years, where the replacement is low for both groups due to exhausting PSL benefits), compared to the difference in absenteeism behavior of workers in the "treatment region" (the mild flu outbreak region/years) identifies the change in the worker's absence-taking behavior in response to change in the replacement rate.

We confirm that absenteeism is responsive to PSL, with approximately 1.2 additional days per year for a balanced sample of administrative workers. Using flu outbreaks, the causal impact is estimated to be approximately 0.9 days per year for these same workers. A substantive portion of the PSL response is due to moderate sick days; absenteeism due to severe illness is unresponsive. With an average work year of roughly 250 working days, this PSL impact is quite modest and much smaller than estimates found in some European studies. That there is a difference is not surprising, considering the large differences in benefit generosities. However, that the estimated magnitudes are no more than 40 to 50 percent of the impact found in Europe invites a deeper analysis of *why* we see such differences.

In addition, we extend our analysis using flu outbreaks to different samples of workers: all workers, workers in output-based compensation jobs, and workers in vulnerable occupations. The "all workers" sample qualitatively replicates our balanced sample of administrative workers, showing that the PSL impact is economy-wide and generalizable. Output-based compensation workers, those that receive both tips or commissions and base pay, face a lower "effective" replacement rate and are shown to be less responsive to PSL, buttressing our main result that at least a part of PSL can be attributed to moral hazard. Results with workers in vulnerable occupations (jobs with close physical contact to other workers or customers), point to potential public health benefits from the increased likelihood of workers to take sick leave. We find such workers are mostly unresponsive to PSL perhaps due to very high initial rates of PSL; employees without PSL may be in firms operating in unusual circumstances.

Finally, we indirectly examine the impact on firm profitability. One would expect moral hazard in sick days would impact profits more so for small firms than large ones, because differently sized firms will have differing capacities to respond to absences. This differential impact appears to be recognized in legislation; for example, Connecticut's statute only applies to firms with 50 or more workers.[7] Large firms can more easily find other employees to fill in for short-term absences. Small firms, on the other hand, may face severe problems if even a single employee is not available. We find significant amounts of moral hazard in large firms but not small firms.

The remainder of the paper is arranged as follows. Section II describes the current state of the literature, differences between the European and U.S. experience in PSL, and challenges of estimating the impact of PSL on worker absenteeism in the U.S. Section III presents the data and identification strategy. Section IV shows the results. Section V concludes.

II. Background

A small but vibrant strand of the economic literature has examined the impact of PSL, generally finding that more generous benefits lead to a relatively modest increase in work absenteeism. To date, academic studies of PSL have been almost exclusively focused on

[7] See https://www.ctdol.state.ct.us/wgwkstnd/NoticeSickLeavePoster2014%20.pdf .

Northern Europe for two important reasons. First, mandated PSL is a much more widespread in Europe than in the United States. Germany, Austria, Sweden, and Norway (where most of the academic studies have been based) all have a long history of very generous national sickness insurance programs. Because of the perceived costs or benefits of PSL, these countries instituted large changes to structure or generosity of their programs. These structural breaks served as natural experiments for economists to analyze the impact of PSL, usually in a difference-in-differences framework. The studies almost universally find that more generous sick leave policy results in higher absenteeism. Elasticity of absenteeism with respect to the replacement rate (percentage of daily wage paid to the employees when they miss work due to sickness) is estimated to be about 0.6.[8] Some studies conclude that reducing the net benefit of taking absence induces workers to return to work quicker (Johansson and Palme 2005; Markussen, Mykletun, and Røed 2012).

Second, because of the universal nature of national sickness insurance in these countries, the data requirements to conduct an analysis are straightforward. Survey respondents do not have to be queried about whether their workplace offers PSL. There are no sample selection issues to worry about, at least with respect to PSL, as workers and firms do not match with each other based on this fringe benefit. Because changes in the sick leave law impact all workers at the same time, the control group is simply everyone in the data in the period prior to the law change, and the treatment group is everyone in the data after the change.

Stark differences between the U.S. and Northern Europe warrant a separate analysis of PSL in the U.S. In contrast to the more ethnically and socio-economically homogenous

[8] For example, a decline in replacement rate from 100 to 80 percent led to a 12 percent decline in the number of absent days taken (Ziebarth and Karlsson, 2010).

populations in Northern Europe, the average U.S. worker earns less, is more likely to be black or Hispanic, and takes less than half the number of sick days.[9] Voluntary PSL provision in the U.S. requires researchers to think carefully about non-random worker-firm matching. The PSL policy debate in the U.S. is fundamentally different, as it centers on imposing a floor on the number of days that firms must supply to all workers. Many firms already offer PSL benefits that match or exceed proposed PSL requirements. For workers that do gain this new benefit, the financial gains may be quite modest, with a sharp increase in the replacement rate from zero to 100 percent for only a very short duration. Summary statistics from the National Compensation Survey, presented in Table 1, illustrates the U.S. PSL landscape in 2013 (U.S. Bureau of Labor Statistics, 2013). About two-thirds of workers in the U.S. are offered PSL; of those workers offered PSL, the majority receives less than 10 paid days off per year. A larger portion of workers are offered paid holidays or vacations, and these other benefits often go along with PSL. Larger firms are more likely to be generous in the provision of PSL; 85 percent of employees at very large firms receive PSL, compared with 52 percent in smaller firms. When PSL is offered, it is overwhelmingly offered in a fixed number of days per year, often with "use it or lose it" restrictions on the number of days that can be carried over from one year to the next.

Figure 1 and Table 2, derived from the NHIS, confirm that U.S. workers take far fewer sick days compared to their European counterparts. The average administrative worker takes less than 5 sick days a year, with the number of moderate and severe sick days evenly split. This

[9] For example, manufacturing workers in Sweden from Johansson and Palme (2005) earned an hourly wage of approximately $22 (in 2015 dollars) and took 10 to 12 sick days per year (Wage/demographic data extracted from Statistics Sweden: http://www.scb.se/statistik/AA/OV0904/2004A01/OV0904_2004A01_BR_SV_A01SA0401.pdf). Manufacturing workers in the U.S. in 2015 earn 10 percent less and take fewer than 5 sick days (Wage/demographic data extracted from Bureau of Labor Statistics: http://www.bls.gov/news.release/empsit.t24.htm).

magnitude is very similar to the findings for U.S. workers in Ziebarth and Karlsson (2010), where U.S. workers had the fewest absence days out of workers in 17 countries. In addition, the NHIS reveals the top 20 percent of workers account for 80 percent of all sick days taken. Most employees do not come close to exhausting their allotment of sick days. These stark differences in the U.S. relative to Europe motivate the need to examine the potential impact of mandated (or expanded) PSL in the U.S. more systematically. However, the empirical issues are formidable in the U.S. The next section describes these issues in detail and outlines our solutions.

III. Data and Identification Strategy

A. Sick Leave Data

Because many U.S. states and cities are only now considering mandating PSL, there are few natural experiments similar to the European cases that we can exploit for identification purposes.[10] Estimating the impact of *mandated* PSL with individual-level data in several states may be feasible in the near future. Thus, any current dataset used in the absenteeism analysis must contain both the number of sick days the employee took and whether the firm offered PSL.[11] Part of the reason why relatively little U.S. research has been done on the economic effects of sick leave is due to lack of such data.[12] In the subsequent analysis, we rely on the

---

[10] Ahn and Yelowitz (2015) examine labor market effects of Connecticut's statewide mandate, but the American Community Survey data does not have information on absenteeism.

[11] Ideally, one would also want to know replacement rates for sick leave, accrued days, use-it-or-lose-it provisions from one year to the next, and substitutability with vacation days. We are not aware of any large scale dataset with such detailed information.

[12] The General Social Survey (GSS), National Survey of America's Families (NSAF). American Time Use Survey (ATUS), National Longitudinal Study of Adolescent to Adult Health (Add Health), and A Three-City Study ask about paid sick leave. The PSID's "Transition to Adulthood" subsample also asks about sick leave. The ATUS asks about a very compressed time schedule, making it impossible to gauge an individual's use of paid sick leave throughout the

NHIS, which is conducted by the National Center for Health Statistics. The NHIS obtains information about the amount and distribution of illness, its effects in terms of disability and chronic impairments, and the kinds of health services people receive. It provides a continuous sampling and interviewing of the civilian, noninstitutionalized population of the United States.

Starting in 1997, the NHIS was redesigned to include a basic module, a periodic module, and a topical module. The basic module corresponds to the NHIS core questionnaire and is made up of the family core, the sample adult core, and the sample child core questions. We use data from the 2005 to 2013 NHIS sample adult files, which ask, on an ongoing basis, sample adult workers both about PSL and absences from work due to illness. In addition, it asks about "bed days", which can be thought of as severe sick days.

The number of sick days taken captures the worker's behavior, impacted partly by the presence of the sick leave benefit, where the worker may truly be incapacitated or engaging in moral hazard behavior. The number of bed days, in contrast, most likely captures the worker's involuntary (or true) absence due to sickness.[13] In essence, the difference between the former and the latter captures "moderate" sick days where the worker could be induced to return to work if necessary. Returning to Table 2, one observes that the typical administrative worker is absent from work for 4.2 days per year, and reported sick days are nearly evenly divided between severe and moderate days.[14] We remain agnostic about whether workers taking these moderate days off

---

year. In recent work, Susser and Ziebarth (2016) examine the sick leave landscape in the U.S. using the ATUS. Miller (1981) also uses a small survey of workers, the Quality of Employment Survey, 1972-73, which also asks about paid-sick-leave provision by firms.

[13] This is, of course, assuming that the worker responds truthfully to the survey. Our subsequent results make us confident that the responses are a valid measure of severe illness.

[14] The NHIS asks about sick days from work and bed-ridden days. Bed-ridden days is asked of both workers and non-workers, and includes both work days and non-work days (i.e. weekends). Thus, the sum of moderate days (created from the difference between sick days at work and bed-

is welfare enhancing. While from an employer's naïve perspective, any productivity from a sick worker may be preferable to none (provided that the employer must pay full wages anyway), there are at least two reasons why the worker staying home to recuperate may benefit the firm. First, resting at home may accelerate the worker's return to full productivity. Second, staying at home may prevent transmission of sickness to other employees and/or customers.[15]

B. Flu Data

While we attempt to mitigate the selection problem in the NHIS data by identifying a subsample of workers that are most likely to be marginal to firm profitability (and thus less likely to drive firm-level HR decisions to offer PSL) and conditioning on a set of very rich employee demographic, behavioral, and health information as well as firm characteristics, we cannot completely discount the possibility that our sample of workers match with a firm due, in part, to PSL. To disentangle this remaining selection issue, we augment the NHIS with data from Google Flu Trends. We focus on influenza because of the widely held belief in the literature that most sick leave days are taken for acute, short-term illness such as the flu or the common cold, where symptoms that would incapacitate the worker usually lasts less than one week (Ziebarth, 2013; Johansson and Palme, 2005). Illnesses of longer duration would not be relevant for PSL in the U.S.

---

ridden days, and truncated at zero) and severe days does not equal sick days from work. However, Panel 2 shows that among those reporting bed-ridden days, they are more than twice as likely to report severe days in a work-week increment (5 days, 10 days, etc.) as a weekly increment (7 days, 14, days, etc.), suggesting that many workers perceive this as a question about the severity of work-related illness. As Panel 3 shows, virtually all individuals report fewer severe days than sick days from work. For example, just 4 percent of workers who took 3 or more sick days from work report more bed days than sick days.

[15] See Susser and Ziebarth (2016) for a deeper analysis of this issue, commonly referred to in the literature as "presenteeism."

Google Flu Trends measures flu prevalence using search data for about 40 flu-related queries.[16] Researchers at Google have shown that the measures in Google Flu Trends track well with the data on flu severity (number of outpatient visits and/or hospitalizations from a sample of hospitals across the U.S.) released by the Center for Disease Control and Prevention (CDC) (Ginsberg, et al., 2009). However, research by Cook, et al. (2011) shows instances where Google Flu Trends measure of flu prevalence diverges sharply from data collected by the CDC.[17] In particular, Google Flu Trends under-counted the severity of the H1N1 flu outbreak of 2009, in comparison to CDC data on hospitalizations. This divergence assists us in our identification, because the CDC data captures more severe cases of the flu that require hospital visits, while Google Flu Trends captures milder cases where the afflicted worker still feels well-enough to search the internet for information about the flu, or post status-updates indicating that he or she has the flu. Focusing on hospital visits may miss mild flu cases where the workers could return to work.[18] We use the CDC data as a robustness check to confirm that workers do not alter their behavior in response to PSL when the flu is severe enough to merit a visit to the hospital.

C. <u>Identification Approach 1: Balanced Sample of Administrative Workers</u>

The voluntary nature of PSL offers also means that we have to account for sample selection issues. In particular, if PSL is offered as a part of a more comprehensive benefits package, employees who highly value these benefits (who may differ in age, experience,

[16] Flu Search Activity, Google Flu Trends. See https://www.google.com/publicdata/explore?ds=z3bsqef7ki44ac_alse.

[17] The CDC data captures the degree of flu prevalence by counting the number of hospital visits and viral surveillance outcomes from hospital laboratory specimens for "influence-like-illnesses (ILI)." ILI is defined as "fever (temperature of 100°F [37.8°C] or greater) and a cough and/or a sore throat in the absence of a known cause other than influenza" (CDC).

[18] That is, if a worker feels the need to visit the hospital or doctor's office for an ILI, we do not regard it as moral hazard behavior.

productivity, and health) may match with firms that offer more generous compensation packages. A simple comparison of absenteeism between workers who have PSL and those who do not will be biased, since those who value sick leave most highly (and thus may be most inclined to use it, whether he/she needs to or not) will seek to match with the firms that are the most generous with this benefit.

While the NHIS solves our problem of finding a dataset that identifies who has access to PSL and takes advantage of it, it does not solve our sample selection problem. The first column of summary statistics from Table 3, which separates roughly 66,000 workers from 2005-2013 by whether their employer offered PSL, shows large differences in observable characteristics. Workers with PSL take more sick days than those without, but also have large differences in lifestyle (smoking, exercise, and alcohol use), socioeconomic status (age, gender, marital status, race, education) and workplace characteristics (tenure, firm size, earnings, salaried workers, private sector, and health insurance).

The fact that workplace characteristics vary between workers with and without PSL suggests non-random matching between workers and firms (Jovanovic, 1979; Postel-Vinay and Robin, 2002; Hwang, Mortensen and Reed, 1998; Garen, 1998; and Woodbury, 1983).[19] However, studies by Poterba, Venti and Wise (1995) and Chetty et al. (2014) argue that such

[19] Although some authors have estimated the effects of firm policies with an instrumental variables approach, it is often difficult to find compelling instruments. Evans, Farrelly and Montgomery (1999) estimate the impact of workplace smoking bans on the likelihood of smoking, using firm size as an instrumental variable. In our context, examining absenteeism, firm size is unlikely to be a valid instrument because it likely has a direct effect on absenteeism. For example, Alvarez (2002) examines work absences and finds firm size is one of the job-related characteristics that affect absenteeism.

worker-firm matching is likely unimportant for pension benefits.[20] We expect that if worker-firm matching is unimportant for secondary components of the compensation package (like pensions), such matching will be unaffected by tertiary features (like PSL, vacation days, wellness/exercise programs, employee discount programs, life insurance, and employee assistance programs).[21] Even if such fringe benefits are initially unimportant for the worker-firm match, they may become more important over time through non-random worker attrition. Workers for whom such fringe benefits are quite valuable may be more likely to stay at the firm. In our work, we address this by examining workers with low tenure where non-random attrition should not be problematic, and find substantively similar effects to our main results.

To minimize the selection issue, we create a sample that is balanced on both sides of the treatment variable: "Do you have PSL on this MAIN job or business?" Broadly speaking, we attempt to isolate the sample to an occupation category that does not require a high amount of initial human capital (academic or experience), is relatively homogenous in job description (thus making employees easily substitutable from the firm's perspective), does not lead to large increases in pay or status after years of employment with the firm (thus leading to workers and firms considering implications of a long term, sustained match), and is well-represented across all industry groups. These employees may or may not be offered PSL, but a firm is unlikely to

[20] Poterba, Venti and Wise (1995) argue "The first approach relies on the largely exogenous determination of 401(k) eligibility, given income. Eligibility is determined by employers. If household saving behavior is largely independent of individual characteristics related to the probability of working at a firm with a 401(k) plan, a hypotheses we evaluate based on saving behavior before 401(k)s became available, then a comparison of the financial assets of families with and without 401(k) eligibility can be used to infer the saving effect of these plans." More recently, Chetty, et al. (2014) examine effects of automatic firm contributions in Denmark among those who switched firms. They show their results are not affected by endogenous sorting.

[21] Harris and Yelowitz (2016) examine effects of employer sponsored life insurance.

design its human resources policy with this class of workers as a primary class of employees to satisfy. That is, from the employer's perspective, whether they offer PSL for these employees would be incidental.

We restrict the sample to non-elderly adult workers with exactly one job, who are paid either on a salaried or hourly basis, work in the public or private sector, provide valid, non-missing answers to all relevant questions, and have at least one year of job tenure.[22] To balance the sample the best we can with respect to job characteristics/job amenities (other than PSL), we restrict the sample to workers in a very common occupation: "Office and Administrative Support Operations." This is the largest single work category (approximately 15% of the employed sample) and is prevalent across many industries which offer very different fringe benefit packages to workers.[23] Restricting to administrative occupations further reduces the sample size to 9,632 observations over nine years.

We use the Administrative Workers as the primary sample for analysis. Summary statistics in Table 3 from Administrative Worker sample, compared to the full sample from the NHIS, clearly shows that our sample is much more balanced. While similarities across demographic characteristics such as gender, race, and marital status are encouraging, the most salient feature of our sample is that the number of bed days is virtually identical, whether PSL was offered or not. This is clear indication that workers in this occupation category did not select

[22] Conditioning on at least one year of job tenure allows workers to accumulate sick leave days, and would also allow workers to move beyond probationary periods, which itself affects absenteeism (Ichno and Riphahn, 2005).
[23] In their analysis of health insurance, Einav, Finkelstein, and Cullen (2010) note that as a consequence of Alcoa's business structure, "employees doing the same job in the same location may face different prices for their health insurance benefits due to their business unit affiliations."

into a particular job to take advantage of PSL because they were more or less susceptible to severe illnesses.

Figures 2a and 2b confirm that PSL policies for Administrative Workers are broadly representative of PSL policies for all other worker categories across all industries. The ranking of PSL availability for all workers is very similar to that of administrative workers. In addition, Figure 3 shows that Administrative Workers compose a sizable fraction of workers across many industries, typically ranging from 10 to 20 percent of all workers. Therefore, an analysis of the balanced category of Administrative workers is actually a substantively broad analysis of all industrial sectors in the economy.

Although our sample of administrative workers is much more balanced across PSL status compared to the full sample, there are still some noteworthy differences. For example, administrative workers with PSL are more likely to be working for public institutions and more likely to be offered employer-sponsored health insurance. These workers are also more highly paid, but this could be attributable to higher average tenure at the job. Although it is unsurprising that workers in firms that offer PSL would also be offered better overall compensation packages, this may lead to some systematic differences in the budget constraint of workers. As most of these differences mildly increase the budget of those employees who have PSL, the estimated impact of sick leave on worker absenteeism is expected to be an upper-bound.[24] We include firm-level characteristics, such as firm size, in an attempt to account for these unobserved

[24] This follows from assuming that the generous benefits and salaried status are treated as non-labor income. Therefore, in a labor-leisure choice model, we expect pure income effects from these compensation characteristics, leading to more leisure time taken (in the form of more sick leave days taken), compared to the case where sick leave is not offered and non-labor income is lower.

differences. Our specifications also include controls for other employer benefits/amenities (ESHI, any health insurance, earnings, class of worker, tenure, hourly vs. salaried), demographics (age, gender, race/ethnicity, marital status, education), health status (chronic conditions, BMI), health habits (smoking, exercise, alcohol), region, year, and the unemployment rate.

We also attempt to spot discernable sick-day taking patterns of workers who have PSL. If PSL is pivotal in worker decisions to call in sick, we may be able to observe workers "bunching" at the PSL limit, where the replacement rate changes from 100 percent to zero. As shown in Table 2, some administrative workers report that they take five sick days, while others report seven days. We suspect both sets of workers may be indicating that they took one week off from work. Our categories attempt to account for this ambiguity by combining days strategically. Table 1 (from the National Compensation Survey) shows considerable variation exists in benefits packages beyond our coarser measure of PSL offered or not (from the NHIS). Ideally, we would like to have fine-grained provision details, which would allow for sharp hypothesis testing on bunching, but unfortunately do not have it in the NHIS.

D. Identification Approach 2: Exploit Regional Flu Shocks with Balanced Sample

Even though our balanced sample approach likely mitigates the selection problem, job matching may still potentially be driven by unobservable worker and firm characteristics. To more completely handle this problem, we augment the NHIS data with data from Google Flu Trends. The central insight is that even if we have not completely controlled for the selection problem of matching with a firm due to PSL, the flu shock is *completely* random. Therefore, the difference in response of workers in control region (high flu prevalence) and treatment region (low flu prevalence) captures the causal response of PSL.

We divide the U.S. into four Census regions (and years) and capture the virulence of the influenza virus to construct a proxy for latent sickness conditions. The details of the construction, as well as the values for the flu index, are provided in the Appendix. Using flu shocks allows us to further explore whether any moral hazard effects we find related to "legitimate" versus "abusive" use of sick leave (as well as bed days). Data constraints from the public-use file of the NHIS restrict us from using finer geographic areas, and it is fair to ask if this introduces significant measurement error by hiding variation in flu rates within a region.[25] To gauge the appropriateness of using Census regions, we use a variant of a test that evaluates whether peer groups are randomly assigned (Sacerdote, 2001). The own characteristic is the dependent variable, and peer characteristic is the independent variable. The insight is that if assignment is random, peer characteristic should not predict own characteristic. We show the opposite: that "assignment" is not random. Using state-level variation in the Google Flu Trends data, we construct two flu exposure rates: an "own" exposure that averages over states in the Census region except the state in question and an "other" exposure that averages over states outside the Census region. If region is an appropriate level of flu catchment, we would expect "own" flu exposure to be more highly correlated with the state's flu exposure. We estimate the parameter on the "own" to be 0.69 (s.e.=0.11), showing that it is highly predictive of the state's flu exposure. The estimate on "other" is one-third in size and statistically insignificant.

[25] Using finer geography as the catchment area creates its own issues. Research has shown a strong spatial pattern to flu transmission, as the prevalence of flu in own and neighboring areas is strongly correlated (Trogdon and Ahn, 2015). Thus using a finer geographic area as an "exogenous" treatment would be problematic, even if the data were available. There is another interesting reason that we may not wish to use too fine a geography. Flu prevalence may be endogenous to PSL policy at the local level. For example, a city-wide PSL mandate may actually lower flu transmission rates (Pichler and Ziebarth, 2015)

The structure of PSL in the U.S., along with the severity or prevalence of the flu, creates an interesting optimization problem for the worker. Unsurprisingly, areas with high prevalence of the flu (across time) will exhibit more sick leave being taken on average. However, in a year where the flu hits particularly hard, workers may quickly run up against the PSL limit and face a replacement rate of zero. Therefore, while the *level* of sick leave taken should be positively correlated with the average level of influenza prevalence, the *rate* of increase in sick days taken should be negatively correlated with prevalence. Using a balanced group of workers and exogenous flu shocks gets us as close as possible to a natural experiment, given the current legislative landscape of the United States.

IV. Econometric Model and Results

The basic model uses our balanced Administrative Worker sample and estimates the impact of being offered PSL on the use of sick days.

$$SickDays_{ijt} = \beta_0 + \beta_1 PSL_{it} + \beta_2 X_{it} + \beta_3 Z_{it} + \beta_4 URATE_{jt} + \delta_j + \delta_t + \varepsilon_{ijt} \quad (1)$$

Where $SickDays_{ijt}$ is the number of sick days that the worker reported taking in the past 12 months. PSL_{it} is an indicator equal to one if the worker had PSL. X_{it} and Z_{it} represent individual and firm level characteristics that can influence a worker's decision to take sick days. δ_j and δ_t are region and time fixed effects, and $URATE_{jt}$ is the regional unemployment rate.[26] In all models, standard errors are clustered for non-nested two-way clustering on region and time (Cameron, Gelbach and Miller, 2011). In this basic specification, with our detailed individual/firm characteristics and restriction of the sample to a homogenous occupation, the

[26] Pichler (2015) examines absenteeism over the business cycle.

assignment of sick leave for an employee will be close-to-random. Therefore, β_1 should be a measure of the causal effect of sick leave, at least for the balanced group considered. If $\beta_1 > 0$, individuals take more sick days in the presence of PSL insurance compared to the absence of such an insurance.

Table 4.A presents these initial results. The first column shows that having PSL results in workers taking an additional 1.2 days off from work due to sickness, compared to workers without PSL. In the second and third columns, when sick days are divided between moderate sick days and severe sick days, we find that PSL only impacts moderate sick days. The interpretation is then that most of the increase in the number of sick days taken due to PSL can be attributed to moral hazard. The fourth column shows a marginally significant 1 percentage point increase in the probability of hospitalization. This may be indicative of the positive correlation between PSL being offered and other benefits (such as more generous health insurance). In addition, the likelihood of extended absences (either 5 or 10 days absent during the year) increases by 1.4 to 3.3 percentage points when PSL is offered, from baseline rates of 7.4 to 15.7 percent. The next two rows show that both workers with children and without children respond similarly to PSL provisions, with the exception of impacts on severe sick days and hospitalizations.

Stratifying workers by tenure (divided at the median of 5 years of tenure for the sample) shows that workers with higher tenure take more sick days compared to low tenure workers when offered PSL. This may be due to accumulated sick days that carry over from one year to the next.[27] Importantly, the effect on moderate sick days is also significant for the low tenure

[27] NCS summary statistics from Table 1 shows that over half of workers who are offered paid sick leave have some type of carryover provision for sick days.

group as well. If workers who remain at a firm for many years are a better match to the firm's fringe benefits, then the interpretation of PSL causing absences would not be justified; rather, the interpretation could be that absence-prone workers remain at firms with PSL. The fact that a similar pattern emerges for low tenure administrative workers suggests that such non-random attrition is not an important concern.

Table 4.B examines "bunching" behavior of workers at common PSL durations – such as one week (measured as 5-7 days, given the difficulties of reporting in the NHIS in Table 2). The 2013 National Compensation Survey (Table 1) showed that conditional on being offered PSL, it was by far the most common for workers in small firms (less than 500 employees) to have 5-9 days of sick leave. Despite data limitations in the NHIS, the analysis finds some evidence of workers changing absence-taking behavior. The "dip" at 0-4 days (4 percentage points) and "bump" at 5-7 days (3 percentage points) for workers in small firms may be showing strategic absence-taking behavior by fully utilizing their PSL benefits and no other days beyond. No such evidence of bunching at one week is found for workers in large firms. However, Table 2 shows that PSL tends to be more generous in large firms, with the plurality of employees receiving two weeks of paid sick leave. There is evidence of extended sick leave spells for workers at large firms, consistent with more generous base level offerings and roll-over provisions. Approximately 11 percent of workers at large firms (500 or more employees) have 15 or more days of paid sick leave. Although these NHIS results are suggestive of bunching, more detailed data with information on PSL provisions would be needed to conclusively show that this behavior is driven by PSL.

Next, we turn to our preferred model which incorporates the flu data from Google Flu Trends to estimate the impact of the "treatment" of receiving more or less negative health shocks in a given region/year for these same administrative workers. The estimating equation becomes:

$$(2) \qquad SickDays_{ijt} = \beta_0 + \beta_1 PSL_{it} LOW_{jt} + \beta_2 PSL_{it} + \beta_3 LOW_{jt} + \beta_4 X_{it} + \beta_5 Z_{it} + \beta_6 URATE_{jt} + \delta_j + \delta_t + \varepsilon_{ijt}$$

Where the variables are defined similarly to before, and LOW_{jt} is an indicator that person *i* resides in a region/year with low prevalence of flu. In equation (2), the coefficient on the interaction term, β_1, represents the impact of paid sick on absenteeism when the replacement rate is high (i.e., the individual has not exhausted the paid days for exogenous reasons related to flu outbreaks), and therefore represents the moral hazard effect. In contrast, the coefficient on β_2 could additionally represent moral hazard (i.e., under the assumption that the sample of administrative workers is balanced) or selective sorting of workers into firms with PSL (i.e., sicker workers choosing firms with more generous PSL). The parameterization of low flu prevalence – for both the Google Flu Trends data and the CDC data – is discussed in detail the appendix; the 36 region-year cells (4 regions, 9 years) are divided at the median for the analysis.

Table 5 shows that workers who live in a high flu prevalence region are more likely to take sick days. Controlling for region and year, being in a high flu region (defined as being above the median for the Google Flu Trend measure), is associated with roughly 0.8 more days of sick days taken.[28] It is worth noting that when we substitute our preferred measure of influenza with data from the CDC, residing in a high flu region is no longer associated with a

[28] We emphasize that this regression is not a tautology. Our regressors for high flu region is from a region-wide measure of flu-related searches from Google, Our dependent variable is the explicit measure of sick days taken from individual worker responses from the NHIS. The large and statistically significant result is central to our use of the Google variable as the treatment indicator.

statistically significant increase in sick days. This lack of correspondence between the two measures can be explained by understanding what types of influenza cases Google Flu Trends and the CDC are measuring. Google Flu Trends captures the overall level of influenza in a region by people's internet search and social media postings. The CDC, on the other hand, captures influenza-associated hospitalizations. While 5 to 20 percent of the population catches the flu in any given year, less than 0.1 percent is hospitalized due to influenza-like symptoms. In addition, the elderly, the very young, and those with pre-existing medical conditions are at higher risk for serious complications from the flu, leading to a higher likelihood of hospitalization.[29] Therefore, the CDC data captures the extreme right-tail of the distribution of influenza severity for a subset of the population that is less likely to be in the labor market.

Having established that workers in high flu regions get sick more often, Table 6 shows our preferred results, which stratifies the administrative sample into high and low prevalence regions. The control group is administrative workers in high flu prevalence regions, and is shown in the first row. For the control group, regardless of whether they have PSL or not, they are more likely to have more bouts of influenza in a given flu season. As such, workers with PSL are more likely to have used up their allotment of sick days and face a replacement rate of zero for marginal absences. When the replacement rate declines, the propensity to take additional sick days also declines. The results indicate the impact of having PSL is substantively smaller (0.73 sick days versus 1.2 for the full sample). In contrast, workers in low flu regions, shown in the second row, display markedly different behavior. Because all workers in these regions are less likely to have received negative health shocks, workers with PSL will be more likely to have full replacement rate sick days available for use. These workers are much more inclined to take sick

[29] Seasonal Influenza Q & A, CDC. See http://www.cdc.gov/flu/about/qa/disease.htm.

days (at an additional 1.76 days), compared to workers in the same region who do not have PSL, and the result is highly significant. Estimating the causal impact of PSL on other measures of absences largely confirm the results we found in Table 4.A. The difference between the two estimates in the first column (1.758-0.728) would be the true estimated causal effect of PSL, free of the selection problem. However, due to the imprecise measure of PSL in the "control" region, we cannot definitively identify the impact.

The third set of results is derived from equation (2) and formalizes our attempt to measure the causal impact of PSL, by combining the sample with the interaction of PSL and low Google flu region. Although the parameter is only significant at the 10 percent level, it is strongly suggestive of the impact of PSL at about 1 day. The next three results in the table buttress our main findings. Instead of dividing the sample into "low" versus "high" by flu prevalence in the median region/year cell, we divide into terciles (dropping the middle one) and again examine the impact of PSL. The qualitative results largely mirror our initial specification, but with somewhat larger effects in the low flu regions and somewhat smaller effects in the high flu regions. This is in-line with our expectations, as we expect the upper and lower tercile to identify "strongly" treated and controlled populations.

For completeness, the final set of results replicate the analysis using CDC measures of flu prevalence. What is most striking is that the impact of PSL seems to be similar across the two populations. This should not be surprising, as the CDC measure captures the extreme right-tail of the flu severity distribution. Worker response to PSL benefits in these regions should *not* differ by rate of hospitalization due to the flu.

Tables 7, 8, and 9 reproduce the findings from Table 6 for the entire sample, vulnerable workers, and sales workers, respectively. Full population results are qualitatively similar but smaller in magnitude compared to estimates using the administrative worker sample. OLS estimates on the full sample for PSL suggest an impact of approximately 0.7 days. Results using the flu region terciles (again, dropping the middle tercile) find a nearly identical effect. Table 7 suggests that the moral hazard effects we derived with our balanced sample of administrative workers is generalizable across the entire population.

Table 8 isolates a sample of workers in “vulnerable” occupations, specifically those in education, health care, and food preparation jobs. Due to close interaction among workers or between customers and workers in these industries, transmission of influenza is more likely and potentially more costly, from a public health perspective. Higher absence in response to PSL being provided in these industries is more likely to be beneficial not only for the worker, but for the firm and the public as well. However, estimates of the causal impact of PSL on days off for this sample is universally small at about 0.2 to 0.5 days and statistically insignificant.

Table 9 estimates the impact of PSL for sales workers. The first row shows the causal impact to be about 0.5 days, and statistically insignificant. Estimates using exogenous flu shocks show the impact to be around 1 day or more, yet all estimates are imprecise.

Finally, Table 10 shows the impact of PSL by firm size. We divide firms into two groups: those with less than 50 employees or at least 50 employees. This line, for example, was used in PSL legislation in Connecticut, which exempted small firms with less than 50 workers. These small firms may find it more difficult to cope with short bouts of worker absences. Estimates show that while employees at both small and large firms increase absence when PSL is offered, workers at large firms are much more responsive (at about 1 day) than their counterparts in smaller firms (at about 0.24 days).

V. Discussion

Using the only publicly available US dataset (NHIS) that asks respondents about the number of moderate and severe sick days taken off from work, whether the firm that employs them offers PSL, and other detailed information about both workers and firms, we find that employees react to PSL by taking roughly 40 percent more sick days, compared to workers that are demographically similar yet work in firms that do not offer PSL. Importantly, PSL only impacts the workers' propensity to take moderate sick days, not severe, bed-ridden sick days.

To further mitigate potential selection issues, we augment this data with Google Flu Trends, allowing us to identify groups of workers where the annual allotment of PSL may be exogenously exhausted, due to unpredictable regional differences in the severity of influenza outbreaks. We find that when workers are more likely to have exhausted their PSL, they are no longer more likely to take sick days, as compared to their counterparts who were never offered PSL. This behavior is consistent with rational workers responding to the replacement rate changing from 100 percent to zero when PSL is exhausted. We replicate this analysis exploiting the influenza natural experiment for all workers to confirm that PSL effect is not just isolated to our balanced sample. These results confirm that our results are robust and generalizable across the entire work force. We then expand our analysis by examining the PSL impact for workers in vulnerable industries, sales workers, and workers in different sized firms. Workers in vulnerable industries, where days off from work due to sickness is most likely to be beneficial for public health, and sales workers are mostly unresponsive to PSL. We also find that workers in larger firms are much more responsive to PSL than their counterparts in small firms.

Our results starkly highlight why further analysis of the impact of PSL in the U.S. is necessary. Quantile regression estimates with German data, when replacement rates change from

100 to 80 percent, show that employees with up to 5.5 absence days reduced absences by about 12 percent. This equates of about 0.66 days. In the U.S. case, our "base" of average absence is smaller, at 4.24 days. If PSL response by U.S. workers were identical to German workers, for a 20 percent change in replacement rate, we should also observe a 12 percent reduction, which equates to 0.5 days. Since the replacement rate change in the U.S. is actually 100 percent, if we linearly extrapolate, worker response should be 2.5 days. With simple estimates from our balanced sample at 1.2 days and flu shock estimates at about 0.9 days, U.S. worker response to PSL is much more muted compared to what was observed in the European studies.

Of course, linearly extrapolating replacement rates from 20 percent to 100 percent and comparing absenteeism is almost certainly incorrect, but that is our point: the base generosity of benefits, the relative changes in replacement rates, and even the demographic makeup of the populations are so vastly different, that it is difficult to translate results from previous studies to inform current U.S. policy. The question of the impact of PSL is far from settled, at least in the United States.

Our analysis of different subsamples of workers confirms some suspicions and yields some unexpected insights. The unresponsiveness of sales workers is no surprise. As earnings in these industries is largely driven by tips and commissions rather than base salary, the "effective replacement rate" would not account for worker's ability to generate sales and commissions. Therefore, for a talented salesperson, provision of PSL (which replaces base salary) may not substantively change the cost of an absence. For a mediocre one, on the other hand, PSL may equate to a relatively high replacement rate. The high level of variance in the PSL effect for these workers may then reflect the distribution of ability.

The lack of PSL response of workers in vulnerable industries is more of a surprise. Workers in these occupations should understand the public health hazard imposed by coming into work sick. However, Figure 2a reveals that in some industries with vulnerable workers (such as education and healthcare), the offering rate of PSL is already very high. Those employees that work in firms in these industries that do not offer PSL, may be in outlier firms operating in very different circumstances compared to the rest of the sample, making apples-to-apples comparison more difficult.

Estimation of the PSL effect with the sample differentiated by firm size shows that workers at large firms take about three-quarters of a day more in absences compared to workers at small firms. This may be due to a combination of easier substitutability among workers allowing workers to take more days off and larger companies being more generous when the benefit is offered due to this flexibility.

In interpreting these findings, we must be careful in assigning value judgements to the increase in propensity to take days off. Most of the literature on European PSL to date has been agnostic about whether workers are abusing the benefit, because while a change in worker absence is easy to demonstrate, the severity of illness during the absence is essentially unknown. The data simply has not been robust enough to differentiate workers' intentions when it comes to taking days off for short, acute illnesses. The popular press, in contrast, tends to present PSL mandates in sharply differing tones. The "virtuous use" of PSL, helping to protect low-income employees when medical issues arise, is sometimes highlighted, while the potential for utilizing the benefit as substitute vacation days is emphasized elsewhere (Miller, 2009; Needleman, 2012; Noguchi, 2015; Boston Globe, 2015; Davis, 2015).

Whether a worker "should" use PSL is a much more nuanced question. At one end, all would agree that a worker taking a sick day to enjoy the sunny weather or to extend a weekend vacation would constitute abuse of the benefit. At the other, if a worker is sick enough to be a health hazard to himself or herself as well as co-workers and customers, it is in everyone's best interest for the employee to stay home and recuperate.[30] Most worker absences fall somewhere between these two extremes. In the data, we showed that the majority of workers do not exhaust their PSL benefit. The generosity of PSL, in replacement rate as well as duration, may be pivotal in determining how ill an employee has to feel before he or she decides to call in sick.

Our analysis goes beyond simply documenting an increase in absenteeism when PSL is offered. Because of the rich NHIS dataset and the use of exogenous variation in influenza prevalence in different regions, we are able to show that the increase in worker absence when PSL is offered is observed mostly on moderate sick days in low influenza regions. This provides stronger evidence that some of increase in absenteeism may be arising due workers abusing the PSL benefit. Because workers stop taking absences for moderate sickness when they are more likely to have exhausted their PSL days (in the high influenza regions), mandating more generous PSL policy by increasing the number of benefit days may have a sizable impact on the labor market. Workers will increase the number of sick days taken, and marginal workers who did not have the flexibility to stop taking sick days at the PSL cap may be induced to seek employment.

[30] A smaller strand of the literature examines presenteeism – the act of attending work while sick – more explicitly. See, for example, Pichler and Ziebarth (2015), Markussen, Mykletun, and Røed (2012), and Dew, Keefe, and Small (2005).

To more fully understand worker behavior, data at the firm level that identifies the generosity of the PSL and the exact dates at which workers use the benefit would be required. For example, if we observe that workers have a higher propensity to call in sick right before major holidays, Mondays, and Fridays, we could more confidently categorize such absences as abuse. Ultimately, our research shows that workers will respond to PSL by taking more absences, and at least some of these absences will be workers misusing the system. Worker response to PSL is a complex question, and more research with richer data is required to fully parse how much of the change in absenteeism is abuse and how much is workers optimally responding to changing incentives.

VI. References

Ahn, T. and A. Yelowitz, 2015. "The Short-Run Impacts of Connecticut's Paid Sick Leave Legislation," Applied Economics Letters, 22(15): 1267-1272.

Alvarez, B., 2002. "Family Illness, Work Absence and Gender," Mimeo, Universidade de Vigo. Retrieved from http://webx06.webs.uvigo.es/sites/default/files/wp0210.pdf.

Boston Globe, 2015. "Paid Sick Leave: A Good Law and a Good Process." Retrieved from https://www.bostonglobe.com/opinion/editorials/2015/06/23/paid-sick-leave-good-law-and-good-process/kq35QDG8ixIKMgCqIpyl9I/story.html.

Cameron, A.C., J.B. Gelbach, and D.L. Miller, 2011. "Robust Inference with Multiway Clustering," Journal of Business & Economic Statistics, 29(1): 238-249.

Cawley, J., J.A. Rizzo, and K. Haas, 2007. "Occupation-Specific Absenteeism Costs Associated with Obesity and Morbid Obesity," Journal of Occupational and Environmental Medicine, 49(12): 1317-1324.

Cawley, J., J.A. Rizzo, and K. Haas, 2008. "The Association of Diabetes with Job Absenteeism Costs Among Obese and Morbidly Obese Workers," Journal of Occupational and Environmental Medicine, 50(5): 527-534.

Chetty, R., J.N. Friedman, S. Leth-Petersen, T.H. Nielsen, and T. Olsen, 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark," Quarterly Journal of Economics 129(3): 1141–1219.

Cook S., C. Conrad, A.L. Fowlkes, and M.H. Mohebbi, 2011. "Assessing Google Flu Trends Performance in the United States during the 2009 Influenza Virus A (H1N1) Pandemic." PLoS One 6(8): e23610.

Council of Economic Advisers, 2014. "The Economics of Paid and Unpaid Leave." Retrieved from https://www.whitehouse.gov/sites/default/files/docs/leave_report_final.pdf.

Davis, P., 2015. "In spotlight, Md. lawmakers take on paid sick leave." Retrieved from http://www.usatoday.com/story/news/nation/2015/03/15/md-paid-sick-leave-bill/24807669/.

Dew, K., V. Keefe, and K. Small, 2005. "'Choosing' to Work When Sick: Workplace Presenteeism," Social Science and Medicine, 60(10): 2273-2282.

Einav, L., A. Finkelstein, and M.R. Cullen, 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices," Quarterly Journal of Economics, 125(3): 877-921.

Evans, W.N., M.C. Farrelly, and E. Montgomery, 1999. "Do Workplace Smoking Bans Reduce Smoking?" American Economic Review, 89(4): 728-747.

Garen, J., 1988. "Compensating Wage Differentials and the Endogeneity of Job Riskiness," The Review of Economics and Statistics, 70(1): 9-16.

Ginsberg, J., M.H Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski, and L. Brilliant, 2009. "Detecting Influenza Epidemics Using Search Engine Query Data," Nature, 457: 1012-1014.

Harris, T. and A. Yelowitz, 2016. "Nudging Life Insurance Holdings in the Workplace," Forthcoming, Economic Inquiry. Retrieved from http://www.yelowitz.com/Harris_Yelowitz_Nudging_Life_Insurance.pdf .

Heymann, J., H.J. Rho, J. Schmitt, and A. Earle, 2009. "Contagion Nation: A Comparison of Paid Sick Day Policies in 22 Countries," Mimeo, Center for Economic Policy and Research. Retrieved from http://cepr.net/documents/publications/paid-sick-days-2009-05.pdf.

Hwang, H.S., Mortensen, D.T., and W.R. Reed, 1998. "Hedonic Wages and Labor Market Search," Journal of Labor Economics, 16(4): 815-847.

Ichino, A. and R.T. Riphahn, 2005. "The Effect of Employment Protection on Worker Effort: A Comparison of Absenteeism During and After Probation," Journal of the European Economic Association, 3(1): 120-43.

Johansson, P. and M. Palme, 1996. "Do Economic Incentives Affect Work Absence? Empirical Evidence using Swedish Micro Data," Journal of Public Economics, 59(2): 195-218.

Johansson, P. and M. Palme, 2002. "Assessing the Effect of Public Policy on Worker Absenteeism," Journal of Human Resources, 37(2): 381-409.

Johansson, P. and M. Palme, 2005. "Moral Hazard and Sickness Insurance," Journal of Public Economics, 89(9-10): 1879-1890.

Jovanovic, B., 1979. "Job Matching and the Theory of Turnover" The Journal of Political Economy, 87(5): 972-990.

Markussen, S., A. Mykletun, and K. Røed, 2012. "The Case for Presenteeism – Evidence from Norway's Sickness Insurance Program," Journal of Public Economics, 96(11–12): 959–972.

Miller, J.W., 2009. "Belgians Take Lots of Sick Leave, And Why Not, They're Depressed," Wall Street Journal. Retrieved from http://www.wsj.com/articles/SB123145414405365887.

Miller, S., 1981. "An Empirical Model of Work Attendance," Review of Economics and Statistics: 77-87.

National Partnership for Women and Families, 2015 "Workers' Access to Paid Sick Days in the States." Retrieved from http://www.nationalpartnership.org/research-library/work-family/psd/workers-access-to-paid-sick-days-in-the-states.pdf.

Needleman, S.E., 2012. "Sick-Time Rules Re-Emerge: More Governments Look to Require Small Businesses to Provide Time-Off Benefits," Wall Street Journal. Retrieved from http://www.wsj.com/articles/SB10001424052970203986604577253550802792104.

Noguchi, Y., 2015. "Obama's Big Bid to Change Sick-Leave Laws May Hinge on Small Business." Retrieved from http://www.npr.org/sections/health-shots/2015/01/22/379102675/obamas-big-bid-to-change-sick-leave-laws-may-hinge-on-small-business.

Olsson, M., 2009. "Employment Protection and Sickness Absence," Labour Economics, 16(2): 208-214.

Pauly, M.V., S. Nicholson, D. Polsky, M.L. Berger, and C. Sharda, 2008. "Valuing Reductions in On-the-Job Illness: 'Presenteeism' from Managerial and Economic Perspectives," Health Economics 17(4): 469-485.

Pichler, S., 2015. "Sickness Absence, Moral Hazard, and the Business Cycle." Health Economics, 24(6): 692-710.

Pichler, S. and N.R. Ziebarth, 2015. "The Pros and Cons of Sick Pay Schemes: Testing for Contagious Presenteeism and Shirking Behavior," Upjohn Institute Working Paper 15-239, Retrieved from: http://research.upjohn.org/cgi/viewcontent.cgi?article=1257&context=up_workingpapers .

Postel-Vinay, F. and J. M. Robin, 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," Econometrica, 70(6): 2295-2350.

Poterba, J.M., S. F. Venti, and D. A. Wise, 1995. "Do 401(k) Contributions Crowd Out Other Personal Saving?" Journal of Public Economics, 58(1): 1-32.

Puhani, P. and K. Sonderhof, 2010. "The Effects of a Sick Pay Reform on Absence and on Health-Related Outcomes," Journal of Health Economics, 29(2): 285-302.

Sacerdote, B., 2001. "Peer Effects With Random Assignment: Results for Dartmouth Roommates," Quarterly Journal of Economics, 116(2): 681-704.

Susser, P. and N. R. Ziebarth, 2016. "Profiling the U.S. Sick Leave Landscape: Presenteeism among Females," Forthcoming, Health Services Research, DOI: http://dx.doi.org/10.1111/1475-6773.12471 .

Trogdon J. and T. Ahn, 2015." Geo-spatial Patterns in Influenza Vaccination: Evidence from Uninsured and Publicly-Insured Children in North Carolina," American Journal of Infection Control, 43(3): 234-240.

U.S. Bureau of Labor Statistics, 2013. "National Compensation Survey: Employee Benefits in the United States, March 2013" Washington, DC: U.S. Department of Labor. Retrieved from http://www.bls.gov/ncs/ebs/benefits/2013/ebbl0052.pdf.

Woodbury, S. A., 1983. "Substitution between Wage and Nonwage Benefits," The American Economic Review, 73(1): 166-182.

Ziebarth, N.R., 2013. "Long-Term Absenteeism and Moral Hazard – Evidence from a Natural Experiment," Labour Economics 24: 277-292.

Ziebarth, N.R. and M. Karlsson, 2010. "A Natural Experiment on Sick Pay Cuts, Sickness Absence, and Labor Costs," Journal of Public Economics, 94(11-12): 1108-1122.

Ziebarth, N.R. and M. Karlsson, 2014. "The Effects of Expanding the Generosity of the Statutory Sickness Insurance System," Journal of Applied Econometrics, 29(2): 208-230.

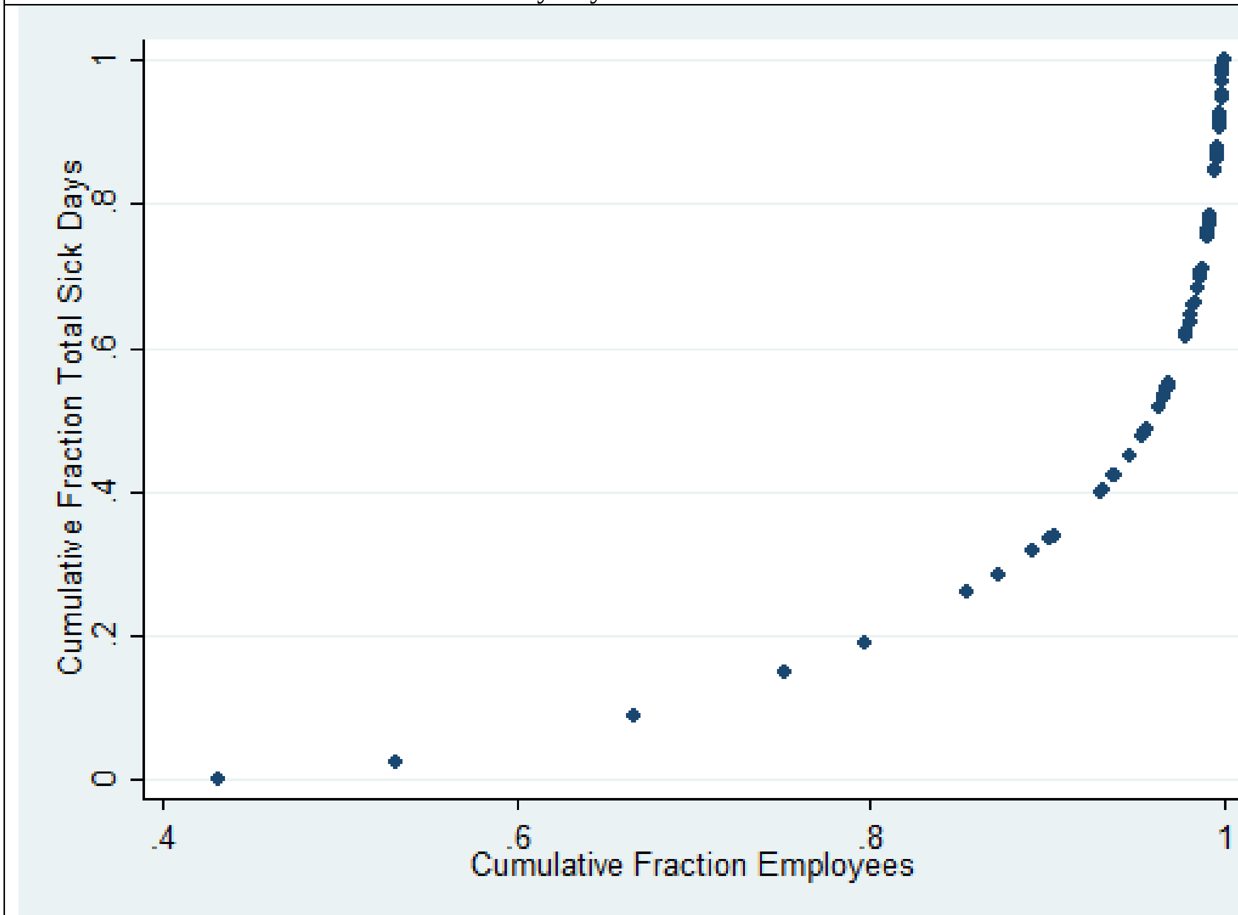
Ziebarth, N.R. and S. Pichler, 2016. "Labor Market Effects of US Sick Pay Mandates," Retrieved from http://www.human.cornell.edu/pam/people/upload/2016-3-23-SickPayMandates.pdf .

Table 1
Sick Leave Benefits in the U.S., March 2013

	Selected Leave Benefits				Type of Provision for Paid Sick Leave			Paid Sick Days by Length of Service				Carryover Provisions		
	Has Paid Sick Leave?	Has Paid Holidays?	Has Paid Vacations?	Has PSL and Vacation	Fixed days	As needed	Consolidated plan	<5 days	5-9 days	10-14 days	15-29 days	Unlimited Accum.	Limit on Days	No Carryover
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
All Workers	65% (0.6)	76% (0.7)	74% (0.7)	59% (0.6)	72% (0.8)	6% (0.6)	22% (0.7)					23% (0.9)	33% (0.8)	44% (1.0)
1+ years of tenure								18% (0.8)	45% (0.8)	30% (0.8)	6% (0.5)			
5+ years of tenure								17 (0.7)	45 (0.8)	30 (0.8)	7 (0.4)			
10+ years of tenure								17 (0.7)	45 (0.8)	30 (0.8)	7 (0.5)			
Office & Admin	75 (1.1)	87 (0.9)	86 (1.1)	73 (1.2)	72 (1.4)	5 (0.6)	23 (1.3)					19 (1.4)	33 (1.2)	48 (1.4)
Sales & Related	52 (1.2)	67 (0.9)	68 (1.1)	50 (1.2)	80 (1.7)	5 (0.8)	15 (1.3)					10 (0.9)	43 (2.0)	47 (2.1)
1-99 Workers	52 (1.1)	68 (1.1)	69 (1.1)	49 (1.0)	70 (1.6)	9 (1.4)	21 (1.3)	22 (1.7)	55 (1.7)	17 (1.3)	3 (0.3)	11 (1.2)	25 (1.1)	64 (1.4)
100-499 Workers	69 (1.2)	82 (1.0)	81 (1.0)	64 (1.2)	75 (1.7)	3 (0.7)	21 (1.4)	18 (1.3)	51 (1.4)	25 (1.3)	6 (0.6)	18 (1.2)	45 (1.7)	37 (1.8)
500+ Workers	85 (0.9)	82 (0.8)	78 (0.8)	71 (0.9)	72 (1.3)	4 (0.4)	25 (1.3)	12 (1.1)	29 (1.4)	47 (1.6)	11 (1.0)	40 (1.7)	32 (1.2)	28 (1.8)

Source: Various tables in the "National Compensation Survey: Employee Benefits in the United States, March 2013" (U.S. Bureau of Labor Statistics, 2013). Standard errors in parentheses. A consolidated leave plan (column 7) provides a single amount of time-off for workers to use for multiple purposes, such as vacation, illness, or personal business. Plans that allow employees to accumulate unused sick leave from year to year (column 12). The NCS represents 124,992,900 civilian workers and 5,361,947 establishments, based on a sample of 7,633 establishments responding to the survey.

Figure 1
Use of Sick Days By Administrative Workers



Notes: Data Source is 2005-2013 National Health Interview Survey. Sample size is 9,632.

Table 2
Annual Absenteeism Among Administrative Workers

Panel 1: Distribution of absences over previous 12 months (N=9,632)

	Number of sick days from work	Number of moderate days	Number of severe days
Mean	4.24	2.44	2.22
(SD)	(13.36)	(9.72)	(9.01)
50th percentile	1	0	0
75th percentile	3	2	2
90th percentile	8	5	5
95th percentile	15	8	8
99th percentile	60	42	30

Panel 2: Reporting of absences over previous 12 months by time units, conditional on <31 sick days (N=9,422)

	Fraction reporting sick days in this group	Fraction reporting moderate days in this group	Fraction reporting severe days in this group
No absences (0 days)	.441	.620	.573
Work week (5, 10, 15, 20, or 25 days)	.103	.041	.044
Calendar week (7, 14, 21, or 28 days)	.032	.016	.018
Calendar month (30 days)	.009	.001	.003

Panel 3: Severe sick days and work absences (N= 9,632)

	All sick days	0 sick days	1 sick day	2 sick days	3+ sick days
Severe days ≤Sick days	.914	.872	.884	.950	.961

Notes: Data Source is 2005-2013 National Health Interview Survey.

Figure 2a: Distribution of Paid Sick Leave across Industry (All Occupations, N=66,535)

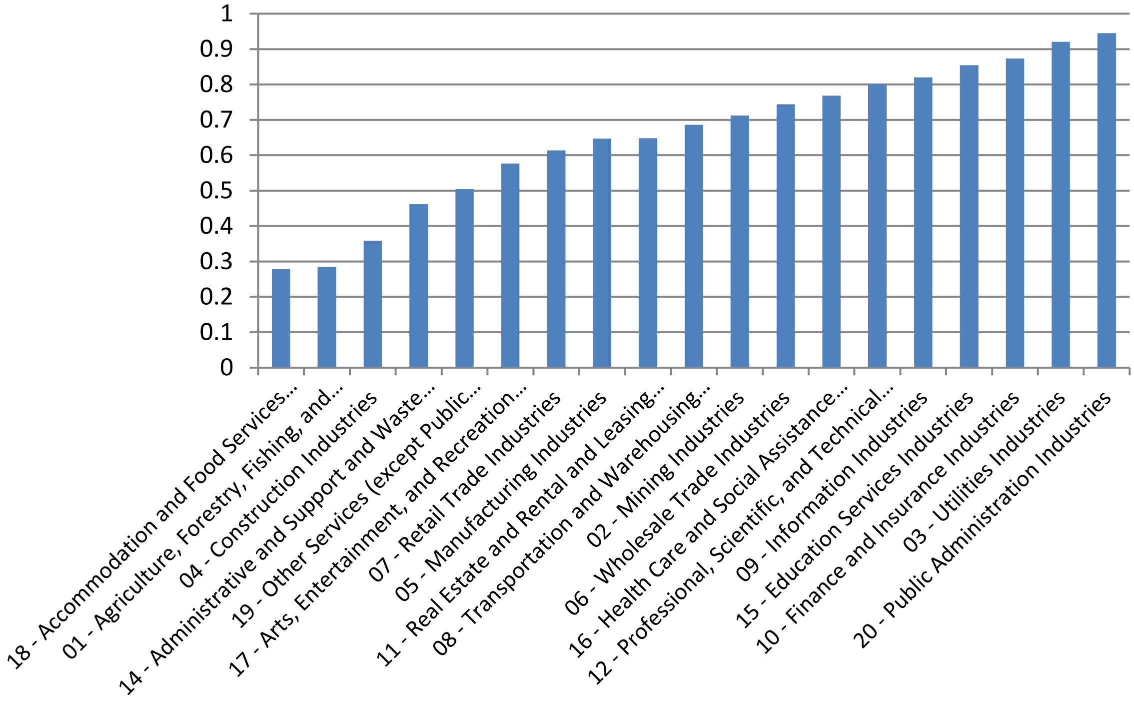
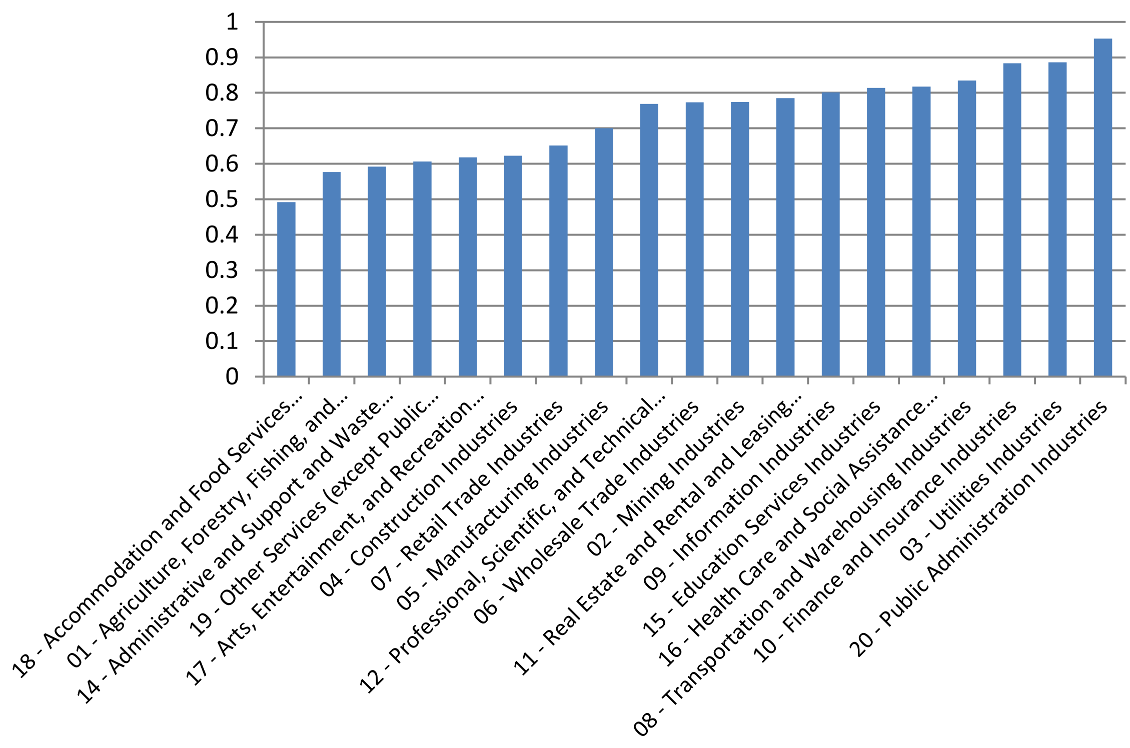
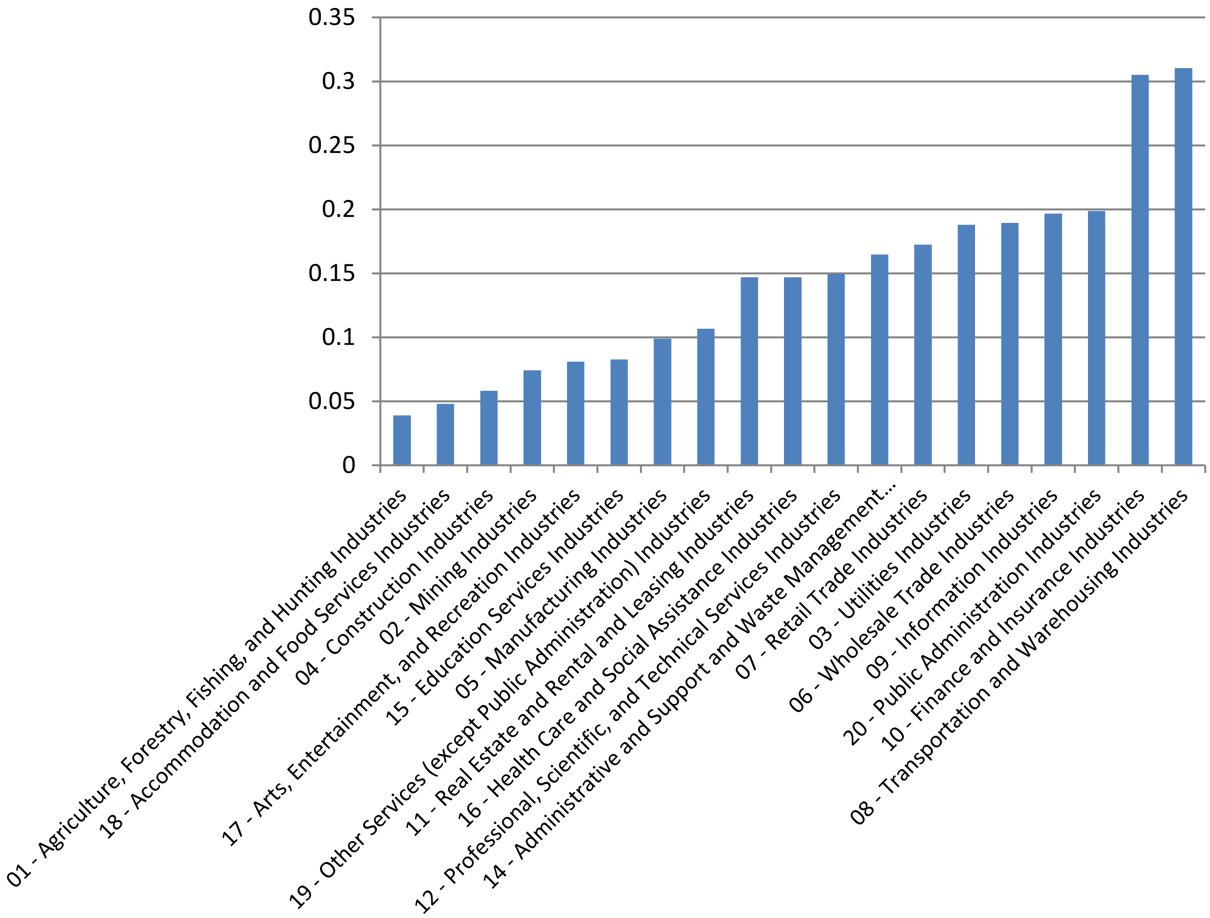


Figure 2b: Distribution of Paid Sick Leave across Industry (Administrative Workers, N=9,632)



Notes: Data Source is 2005-2013 National Health Interview Survey.

Figure 3
Distribution of Administrative Workers across Industry (N=66,535)



Notes: Data Source is 2005-2013 National Health Interview Survey.

Table 3
Summary Statistics

	All Occupations		Admin Only	
	PSL=1	PSL=0	PSL=1	PSL=0
	(1)	(2)	(3)	(4)
Number of sick days from work	3.720	3.027	4.542	3.210
	(0.057)	(0.086)	(0.154)	(0.293)
Number of moderate days	2.278	1.792	2.712	1.516
	(0.045)	(0.065)	(0.119)	(0.155)
Number of severe days	1.752	1.726	2.216	2.241
	(0.032)	(0.060)	(0.096)	(0.239)
Hospitalizations? (0/1)	0.055	0.050	0.064	0.057
	(0.001)	(0.001)	(0.003)	(0.005)
5 Or More Sick Days? (0/1)	0.176	0.133	0.217	0.157
	(0.002)	(0.002)	(0.005)	(0.008)
10 Or More Sick Days? (0/1)	0.077	0.064	0.102	0.074
	(0.001)	(0.002)	(0.003)	(0.006)
Functional Limitations? (0/1)	0.223	0.219	0.271	0.271
	(0.002)	(0.003)	(0.005)	(0.010)
Never smoked? (0/1)	0.636	0.575	0.630	0.595
	(0.002)	(0.003)	(0.006)	(0.011)
Never Exercise? (0/1)	0.237	0.341	0.272	0.304
	(0.002)	(0.003)	(0.005)	(0.010)
BMI Overweight or Obese? (0/1)	0.654	0.637	0.653	0.599
	(0.002)	(0.003)	(0.006)	(0.011)
Abstain from alcohol? (0/1)	0.143	0.186	0.157	0.182
	(0.002)	(0.003)	(0.004)	(0.008)
Age in years	42.249	39.500	42.758	39.859
	(0.053)	(0.086)	(0.133)	(0.282)
Male? (0/1)	0.478	0.539	0.244	0.255
	(0.002)	(0.003)	(0.005)	(0.009)
Married? (0/1)	0.510	0.448	0.464	0.441
	(0.002)	(0.003)	(0.006)	(0.011)
Non-White? (0/1)	0.369	0.454	0.411	0.372
	(0.002)	(0.003)	(0.006)	(0.010)
High School Grad or Dropout? (0/1)	0.589	0.827	0.797	0.844
	(0.002)	(0.003)	(0.005)	(0.008)
Job Tenure (in years)	9.532	6.800	9.530	6.418
	(0.039)	(0.049)	(0.097)	(0.146)
Firm Size < 100? (0/1)	0.485	0.737	0.494	0.691
	(0.002)	(0.003)	(0.006)	(0.010)
Earnings Under $35000? (0/1)	0.344	0.713	0.511	0.815
	(0.002)	(0.003)	(0.006)	(0.008)
Hourly Worker? (0/1)	0.525	0.754	0.698	0.835
	(0.002)	(0.003)	(0.005)	(0.008)
Private Sector Worker? (0/1)	0.754	0.940	0.763	0.924
	(0.002)	(0.002)	(0.005)	(0.006)
Health insurance through employer plan? (0/1)	0.904	0.567	0.907	0.678
	(0.001)	(0.003)	(0.003)	(0.010)
Unemployment Rate (%)	7.150	7.234	7.093	7.266
	(0.009)	(0.013)	(0.023)	(0.043)
Low Google Flu Region? (0/1)	0.473	0.468	0.475	0.488
	(0.002)	(0.003)	(0.006)	(0.011)
Low CDC Flu Region? (0/1)	0.494	0.499	0.478	0.511
	(0.002)	(0.003)	(0.006)	(0.011)
Sample Size	45,250	21,285	7,467	2,165

Data Source: Variable means constructed from 2005-2013 National Health Interview Survey. Standard errors in parentheses.

Office and Administrative Support Occupations (51 groups): First-line supervisors/managers of office and administrative support workers, Switchboard operators, including answering service, Telephone operators, Communications equipment operators, all other, Bill and account collectors, Billing and posting clerks and machine operators, Bookkeeping, accounting, and auditing clerks, Gaming cage workers, Payroll and timekeeping clerks, Procurement clerks, Tellers, Brokerage clerks, Correspondence clerks, Court, municipal, and license clerks, Credit authorizers, checkers, and clerks, Customer service representatives, Eligibility interviewers, government programs, File Clerks, Hotel, motel, and resort desk clerks, Interviewers, except eligibility and loan, Library assistants, clerical, Loan interviewers and clerks, New accounts clerks, Order clerks, Human resources assistants, except payroll and timekeeping, Receptionists and information clerks, Reservation and transportation ticket agents and travel clerks, Information and record clerks, all other, Cargo and freight agents, Couriers and messengers, Dispatchers, Meter readers, utilities, Postal service clerks, Postal service mail carriers, Postal service mail sorters, processors, and processing machine operators, Production, planning, and expediting clerks, Shipping, receiving, and traffic clerks, Stock clerks and order fillers, Weighers, measurers, checkers, and samplers, recordkeeping, Secretaries and administrative assistants, Computer operators, Data entry keyers, Word processors and typists, Desktop publishers, Insurance claims and policy processing clerks, Mail clerks and mail machine operators, except postal service, Office clerks, general, Office machine operators, except computer, Proofreaders and copy markers, Statistical assistants, Office and administrative support workers, all other.

See http://www.bls.gov/tus/census02iocodes.pdf , p. 12-13.

Table 4.A
Balanced Sample of Administrative Workers - Does PSL Affect Absences?

	Number of Sick Days	Moderate Sick Days	Severe Sick Days	Hospitalized During Year? (0/1)	5 or more Sick Days? (0/1)	10 or more Sick Days? (0/1)
	Administrative Workers (N=9,632)					
PSL?	1.230***	0.984***	0.257	0.011*	0.033***	0.014
(0/1)	(0.357)	(0.337)	(0.219)	(0.006)	(0.009)	(0.011)
	Administrative Workers With Kids (N=3,937)					
PSL?	1.196*	1.129***	-0.102	0.002	0.023	0.013
(0/1)	(0.699)	(0.381)	(0.311)	(0.010)	(0.023)	(0.020)
	Administrative Workers Without Kids (N=5,695)					
PSL?	1.272	0.880	0.565**	0.019**	0.039***	0.015
(0/1)	(0.874)	(0.686)	(0.263)	(0.010)	(0.010)	(0.015)
	Administrative Workers with Low Job Tenure (1-5 years, N=4,627)					
PSL?	0.899	0.909**	0.041	0.004	0.031*	0.015
(0/1)	(0.696)	(0.461)	(0.474)	(0.014)	(0.019)	(0.013)
	Administrative Workers with High Job Tenure (6+ years, N=5,005)					
PSL?	1.570***	1.033***	0.543*	0.023*	0.027*	0.012
(0/1)	(0.604)	(0.329)	(0.283)	(0.012)	(0.016)	(0.015)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011).

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 4.B

Balanced Sample of Administrative Workers - Is There Bunching at Sick Day Limits?

	0-4 days	5-7 days	8-9 days	10-14 days	15-21 days	>21 days
			Full Sample of Admin Workers (N=9,632)			
PSL?	-0.033***	0.020**	0.000	-0.001	0.001	0.013***
(0/1)	(0.009)	(0.009)	(0.004)	(0.009)	(0.008)	(0.004)
			Small Firms, 1-99 employees (N= 5,183)			
PSL?	-0.041**	0.031***	0.002	0.004	0.000	0.004
(0/1)	(0.016)	(0.010)	(0.004)	(0.012)	(0.006)	(0.007)
			Large Firms, 100 or more employees (N=4,449)			
PSL?	-0.017	-0.004	-0.005	-0.007	0.004	0.028***
(0/1)	(0.024)	(0.019)	(0.007)	(0.015)	(0.019)	(0.006)
			Exclude top 1% of users (N=9,513)			
PSL?	-0.025***	0.021**	0.000	0.000	0.002	0.002
(0/1)	(0.009)	(0.009)	(0.004)	(0.009)	(0.008)	(0.005)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011).

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 5
Does Living in a High Flu Region Increase Sick Days?
(Balanced Sample of Administrative Workers)

High Google Flu Region (0/1)	0.816***	
	(0.265)	
High CDC Flu Region (0/1)		-0.004
		(0.524)
Regional Unemployment Rate	0.185	0.236
	(0.323)	(0.310)
Region 2 (Midwest Region)	-0.314	-0.346
	(0.392)	(0.453)
Region 3 (South Region)	-0.816***	-0.535
	(0.243)	(0.589)
Region 4 (West Region)	-0.901*	-0.962*
	(0.492)	(0.519)
Year 2006	-0.259	-0.554
	(0.472)	(0.509)
Year 2007	0.442	0.155
	(0.362)	(0.286)
Year 2008	-0.638	-0.673
	(0.510)	(0.456)
Year 2009	-2.130	-1.850
	(1.475)	(1.349)
Year 2010	-1.500	-2.041
	(1.609)	(1.425)
Year 2011	-1.639	-1.863
	(1.479)	(1.372)
Year 2012	-1.783*	-1.427
	(0.980)	(0.908)
Year 2013	-1.785**	-1.400*
	(0.771)	(0.728)
Constant term	4.188***	4.154***
	(1.371)	(1.289)

Notes: Sample adults in 2005-2013 NHIS. Sample size in all specifications is 9,632. Standard errors in parentheses. High flu region defined at the REGION*YEAR level from Google Flu Trends (See https://www.google.org/flutrends/us/#US) or CDC. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011).

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 6
Balanced Sample of Administrative Workers - Using Regional Flu Shocks to Create Exogenous Variation in Replacement Rate

	Number of Sick Days	Moderate Sick Days	Severe Sick Days	Hospitalized During Year? (0/1)	5 or more Sick Days? (0/1)	10 or more Sick Days? (0/1)
	Administrative Workers in High Google Flu Regions (N=5,027)					
PSL?	0.728*	0.727	-0.062	0.011*	0.021	0.006
(0/1)	(0.434)	(0.479)	(0.322)	(0.006)	(0.013)	(0.013)
	Administrative Workers in Low Google Flu Regions (N=4,605)					
PSL?	1.758***	1.221***	0.627***	0.012	0.047***	0.022
(0/1)	(0.473)	(0.353)	(0.186)	(0.011)	(0.018)	(0.019)
	Administrative Workers in All Google Flu Regions (N=9,632)					
PSL*Low Google Region	0.940* (0.562)	0.475 (0.489)	0.696 (0.585)	0.002 (0.009)	0.021 (0.025)	0.014 (0.021)
PSL?	0.773*	0.738	-0.058	0.010*	0.023**	0.007
(0/1)	(0.465)	(0.473)	(0.409)	(0.006)	(0.012)	(0.013)
	Administrative Workers in High Google Flu Regions, Upper tercile (N=3,875)					
PSL?	0.609	0.694	-0.137	0.013	0.012	0.006
(0/1)	(0.530)	(0.677)	(0.354)	(0.008)	(0.016)	(0.013)
	Administrative Workers in Low Google Flu Regions, Lower tercile (N=2,321)					
PSL?	2.097***	1.298***	0.622***	0.028**	0.062***	0.044*
(0/1)	(0.495)	(0.408)	(0.239)	(0.011)	(0.018)	(0.024)
	Administrative Workers in All Google Flu Regions, exclude middle tercile (N=6,196)					
PSL*Low Google Region	1.157* (0.631)	0.545 (0.587)	0.581 (0.577)	0.011 (0.010)	0.042 (0.027)	0.032 (0.034)
PSL?	0.730	0.713	-0.058	0.014**	0.015	0.009
(0/1)	(0.501)	(0.503)	(0.390)	(0.007)	(0.016)	(0.016)
	Administrative Workers in High CDC Flu Regions (N=4,956)					
PSL?	1.006**	0.983***	0.010	0.008	0.022**	0.007
(0/1)	(0.436)	(0.156)	(0.216)	(0.007)	(0.009)	(0.010)
	Administrative Workers in Low CDC Flu Regions (N=4,676)					
PSL?	1.478**	0.947	0.576***	0.016	0.044**	0.020**
(0/1)	(0.697)	(0.596)	(0.189)	(0.010)	(0.018)	(0.010)
	Administrative Workers in CDC Flu Regions (N= 9,632)					

PSL*Low CDC Region	0.470 (0.924)	0.177 (0.815)	0.369 (0.717)	-0.002 (0.012)	0.017 (0.026)	0.008 (0.014)
PSL? (0/1)	0.994* (0.570)	0.878*** (0.312)	0.099 (0.428)	0.013 (0.009)	0.025 (0.016)	0.009 (0.013)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011). Models that interact PSL with flu region also include interactions of flu region with individual characteristics.

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 7
Full Population Results

	Number of Sick Days	Moderate Sick Days	Severe Sick Days	Hospitalized During Year? (0/1)	5 or more Sick Days? (0/1)	10 or more Sick Days? (0/1)
			All Workers (N=66,535)			
PSL?	0.663***	0.532***	0.067	0.004*	0.026***	0.009***
(0/1)	(0.167)	(0.110)	(0.141)	(0.002)	(0.006)	(0.002)
			All Workers in High Google Flu Regions (N=35,148)			
PSL?	0.482	0.457*	0.042	0.005	0.024***	0.008
(0/1)	(0.299)	(0.246)	(0.097)	(0.004)	(0.005)	(0.005)
			All Workers in Low Google Flu Regions (N=31,387)			
PSL?	0.864***	0.620***	0.091	0.002	0.028***	0.011*
(0/1)	(0.258)	(0.111)	(0.215)	(0.006)	(0.006)	(0.006)
			All Workers in All Google Flu Regions (N=66,535)			
PSL*Low Google Region	0.395	0.160	0.077	-0.002	0.005	0.003
	(0.395)	(0.354)	(0.190)	(0.010)	(0.007)	(0.011)
PSL?	0.477	0.458*	0.030	0.004	0.023***	0.007
(0/1)	(0.301)	(0.273)	(0.148)	(0.004)	(0.007)	(0.006)
			All Workers in High Google Flu Regions, Upper tercile (N=27,305)			
PSL?	0.244	0.309	-0.012	0.005	0.018***	0.005
(0/1)	(0.250)	(0.195)	(0.112)	(0.005)	(0.006)	(0.006)
			All Workers in Low Google Flu Regions, Lower tercile (N=15,985)			
PSL?	0.864***	0.591***	0.052	0.006	0.026***	0.015**
(0/1)	(0.295)	(0.173)	(0.235)	(0.010)	(0.008)	(0.006)
			All Workers in All Google Flu Regions, exclude middle tercile (N=43,290)			
PSL*Low Google Region	0.672**	0.298	0.137	0.002	0.008	0.010
	(0.311)	(0.347)	(0.196)	(0.015)	(0.013)	(0.009)
PSL?	0.222	0.299	-0.037	0.005	0.018**	0.005
(0/1)	(0.262)	(0.226)	(0.128)	(0.005)	(0.008)	(0.006)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011). Models that interact PSL with flu region also include interactions of flu region with individual characteristics.

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 8
Vulnerable Workers Results

	Number of Sick Days	Moderate Sick Days	Severe Sick Days	Hospitalized During Year? (0/1)	5 or more Sick Days? (0/1)	10 or more Sick Days? (0/1)
	Vulnerable Workers (N=13,106)					
PSL?	0.489	0.449	-0.044	-0.003	0.020	0.006
(0/1)	(0.384)	(0.278)	(0.282)	(0.008)	(0.013)	(0.008)
	Vulnerable Workers in High Google Flu Regions (N=6,937)					
PSL?	0.406	0.435	0.005	-0.002	0.007	-0.002
(0/1)	(0.258)	(0.290)	(0.134)	(0.011)	(0.005)	(0.006)
	Vulnerable Workers in Low Google Flu Regions (N=6,169)					
PSL?	0.542	0.436	-0.130	-0.006	0.033	0.014
(0/1)	(0.575)	(0.375)	(0.486)	(0.011)	(0.021)	(0.017)
	Vulnerable Workers in All Google Flu Regions (N=13,106)					
PSL*Low	0.370	0.034	0.092	0.000	0.036*	0.021
Google Region	(0.599)	(0.462)	(0.551)	(0.010)	(0.021)	(0.017)
PSL?	0.285	0.408	-0.098	-0.004	0.002	-0.004
(0/1)	(0.287)	(0.314)	(0.289)	(0.011)	(0.008)	(0.007)
	Vulnerable Workers in High Google Flu Regions, Upper tercile (N=5,389)					
PSL?	0.135	0.094	0.170*	-0.003	0.007	-0.006
(0/1)	(0.197)	(0.232)	(0.096)	(0.014)	(0.006)	(0.006)
	Vulnerable Workers in Low Google Flu Regions, Lower tercile (N=3,141)					
PSL?	0.291	0.242	-0.430	-0.005	0.032	-0.003
(0/1)	(0.684)	(0.401)	(0.736)	(0.007)	(0.043)	(0.017)
	Vulnerable Workers in All Google Flu Regions, exclude middle tercile (N=8,530)					
PSL*Low	0.233	0.099	-0.438	0.001	0.035	0.007
Google Region	(0.794)	(0.631)	(0.698)	(0.015)	(0.039)	(0.018)
PSL?	0.081	0.094	0.106	-0.004	0.003	-0.008
(0/1)	(0.190)	(0.233)	(0.148)	(0.014)	(0.007)	(0.007)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011). Models that interact PSL with flu region also include interactions of flu region with individual characteristics.

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 9
Output-based (Sales) Workers Results

	Number of Sick Days	Moderate Sick Days	Severe Sick Days	Hospitalized During Year? (0/1)	5 or more Sick Days? (0/1)	10 or more Sick Days? (0/1)
	Output-based (Sales) Workers (N=6,122)					
PSL?	0.469	0.286	0.133	0.006	0.014	0.004
(0/1)	(0.584)	(0.464)	(0.295)	(0.007)	(0.017)	(0.016)
	Output-based (Sales) Workers in High Google Flu Regions (N=3,267)					
PSL?	0.020	0.007	0.151	0.002	0.013	-0.006
(0/1)	(0.288)	(0.291)	(0.393)	(0.004)	(0.020)	(0.011)
	Output-based (Sales) Workers in Low Google Flu Regions (N=2,855)					
PSL?	1.088	0.686	0.111	0.007	0.017	0.015
(0/1)	(0.965)	(0.426)	(0.592)	(0.010)	(0.018)	(0.018)
	Output-based (Sales) Workers in All Google Flu Regions (N=6,122)					
PSL*Low	0.957	0.597	-0.140	0.002	0.000	0.017
Google	(0.898)	(0.411)	(0.716)	(0.009)	(0.025)	(0.017)
Region						
PSL?	0.069	0.042	0.211	0.004	0.015	-0.004
(0/1)	(0.337)	(0.345)	(0.477)	(0.005)	(0.022)	(0.013)
	Output-based (Sales) Workers in High Google Flu Regions, Upper tercile (N=2,497)					
PSL?	-0.004	0.004	0.271	0.006**	0.013	-0.009
(0/1)	(0.344)	(0.418)	(0.349)	(0.003)	(0.028)	(0.016)
	Output-based (Sales) Workers in Low Google Flu Regions, Lower tercile (N=1,445)					
PSL?	1.749	1.080*	0.000	-0.001	0.018	0.034**
(0/1)	(1.076)	(0.622)	(0.755)	(0.008)	(0.028)	(0.016)
	Output-based (Sales) Workers in All Google Flu Regions, exclude middle tercile (N=3,942)					
PSL*Low	1.630	0.980	-0.386	-0.009	0.004	0.041**
Google	(1.087)	(0.607)	(0.829)	(0.011)	(0.024)	(0.019)
Region						
PSL?	0.081	0.069	0.348	0.007**	0.015	-0.008
(0/1)	(0.329)	(0.400)	(0.379)	(0.003)	(0.025)	(0.016)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011). Models that interact PSL with flu region also include interactions of flu region with individual characteristics.

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

Table 10
Cost Factors: Does the PSL effect vary by Firm Size?

	Number of Sick Days	Moderate Sick Days	Severe Sick Days	Hospitalized During Year? (0/1)	5 or more Sick Days? (0/1)	10 or more Sick Days? (0/1)
	Firm with Under 50 Workers (N=29,528)					
PSL?	0.243*	0.208**	-0.003	-0.001	0.018*	0.003
(0/1)	(0.132)	(0.099)	(0.102)	(0.003)	(0.011)	(0.003)
	Firm with 50 or more Workers (N=37,007)					
PSL?	1.058***	0.837***	0.129	0.008**	0.032***	0.015***
(0/1)	(0.275)	(0.272)	(0.205)	(0.004)	(0.007)	(0.004)

Notes: Sample adults in 2005-2013 NHIS. Standard errors in parentheses. Covariates include: quartic in age, region effects, year effects, job tenure, functional limitations, current/former smoker, never exercise, obesity, current alcohol consumption, gender, marital status, number of children of each age, education, race, earnings bins, health insurance status, hourly worker, firm characteristics (firm size, public/private sector, and industry), and the regional unemployment rate. Standard errors corrected for non-nested two-way clustering at the REGION and YEAR levels, using methods described in Cameron, Gelbach and Miller (2011).

*** = Significant at 99% confidence level, ** = Significant at 95% confidence level, * = Significant at 90% confidence level.

APPENDIX – Construction of Google and CDC Flu Trends

Google Flu Trend Search Data: We collected data from Google Flu Trends for the years 2005-2013. For each state and year, Google collected weekly data on search interest for flu-related illnesses. Predictably, search interest peaks during flu season. Since the NHIS data is annual, we average flu searches for a state across all weeks during the year. A small number of state/year combinations
(such as Wyoming in 2005 and 2006, and South Dakota and North Dakota in 2005) did not have Google flu trend data.

The NHIS divides states into 4 regions – Northeast, Midwest, South and West. Based on 2005 populations by state, we create a regional index as weighted average of the state indexes for each of the 9 years from 2005-2013. Thus, there are 36 unique region-year values for the Google Flu data.

The median value of the index is 1607.5; the 18 region-year cells containing values below that are classified as “Low Google Flu” regions, while the rest are “High Google Flu” regions. In some regressions, the region-year cells are divided into terciles (excluding the middle tercile). The 12 region-year cells containing values at or below 1374 are then the “Low Google Flu” regions while the 12 region-year cells contain values at or above 1903 are the “High Google Flu” regions.

Appendix Table 1

Google Flu Trend Index Values, By Region and Year

	Region 1 Northeast	Region 2 Midwest	Region 3 South	Region 4 West
2005	1374	1232	1803	1723
2006	1225	1107	1642	1365
2007	1166	1079	1611	1329
2008	1574	1475	2070	1604
2009	2423	2124	2992	2748
2010	1161	1083	1554	1324
2011	1406	1339	1878	1564
2012	1886	1903	2856	2071
2013	2249	2078	3011	2726

Notes: Authors' tabulations from Google Flu Trends data. See https://www.google.org/flutrends/about/data/flu/us/data.txt . Region classifications derived from definitions in NHIS. Northeast contains: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. Midwest contains: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. South contains: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia. West contains: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

CDC Flu Hospitalization Data: The CDC provides – by nine Census divisions – weekly information on the count of flu-related hospitalizations and total hospitalizations. We aggregate these counts of flu hospitalizations and total hospitalizations to the annual level for 2005-2013, for each of the four NHIS-defined regions, and compute the flu-related hospitalization rate for each region and year. Again, there are 36 unique region-year values for the CDC Flu Hospitalization data.

The median value of the CDC Flu related hospitalization rate is 1.658%; the 18 region-year cells containing values below that are classified as "Low CDC Flu" regions, while the rest are "High CDC Flu" regions. In some regressions, the region-year cells are divided into terciles (excluding the middle tercile). The 12 region-year cells containing values at or below 1.553% are then the

“Low CDC Flu” regions while the 12 region-year cells contain values at or above 1.963% are the “High CDC Flu” regions.

Appendix Table 2
CDC Flu-Related Hospitalization Rates, By Region and Year

	Region 1 Northeast	Region 2 Midwest	Region 3 South	Region 4 West
2005	1.479%	2.009%	2.466%	1.554%
2006	1.342%	1.609%	2.233%	1.245%
2007	1.362%	1.608%	2.014%	1.187%
2008	1.606%	1.937%	2.273%	1.401%
2009	3.131%	2.803%	3.546%	2.752%
2010	1.300%	1.059%	1.733%	1.434%
2011	1.496%	1.553%	2.094%	1.698%
2012	1.465%	1.615%	1.963%	1.618%
2013	1.784%	1.697%	2.217%	1.786%

Notes: Authors’ tabulations from CDC Flu-Related Hospitalization data. Region classifications derived from definitions in NHIS. Northeast contains: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. Midwest contains: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. South contains: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia. West contains: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.