

Ciccarelli, Nicola

## Article

# The effects of employer-sponsored health insurance premiums on employment and wages: Evidence from US longitudinal data

Review of Economic Analysis (REA)

## Provided in Cooperation with:

International Centre for Economic Analysis (ICEA), Waterloo, Ontario

*Suggested Citation:* Ciccarelli, Nicola (2020) : The effects of employer-sponsored health insurance premiums on employment and wages: Evidence from US longitudinal data, Review of Economic Analysis (REA), ISSN 1973-3909, International Centre for Economic Analysis (ICEA), Waterloo (Ontario), Vol. 12, Iss. 1, pp. 89-127,  
<https://doi.org/10.15353/rea.v12i1.1697>

This Version is available at:

<https://hdl.handle.net/10419/328084>

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc/3.0/>

# The Effects of Employer-Sponsored Health Insurance Premiums on Employment and Wages: Evidence from US Longitudinal Data

NICOLA CICCARELLI\*<sup>†</sup>

*Independent Researcher*

We analyze the effect of employer-sponsored health insurance premiums on employment and annual wages in the US using a county-level panel dataset for the period 2005-2010. Using variation in medical malpractice payments and variation in medical malpractice legislation over time and within states as the source of identifying variation in the health insurance premiums, we estimate the causal effects of rising health insurance premiums on employment and annual wages. We find that a 10% increase in premiums reduces employment by 1.1 percentage points, and leads to a statistically insignificant reduction of annual wages. Since US counties are characterized by a varying degree of private health insurance coverage, we also test whether the private health insurance coverage is a *moderating variable* for the relationship between the health insurance premiums and the labor market outcomes analyzed in this study. We find that rising premiums negatively affect the labor market conditions faced by US workers, especially in areas that are characterized by high private health insurance coverage.

**Keywords:** Employer-sponsored health insurance premiums, Wages, Employment, Panel data, Instrumental variables

**JEL Classifications:** I1, I13, J0, J3

## 1 Introduction

The rate of growth in health care spending and the rate of growth in employer-sponsored health insurance premiums in the United States have outpaced the inflation rate for many years. For ex-

---

\* I especially thank Eric Bonsang, Bertrand Melenberg, Jan C. van Ours, Arthur van Soest, Ben A. Volvaard for extensive feedback on this paper. I thank Tobias Klein, Martin Salm, conference participants at AMES (Asia Meeting of the Econometric Society), and seminar participants at Tilburg University (GSS) for their helpful comments and suggestions. The author is indebted to Jerzy (Jurek) Konieczny, the Editor of Review of Economic Analysis, and to an anonymous referee for very helpful comments and suggestions that improved this paper. The dataset used in this paper is available on request from the author at [ciccarelli.nicola@gmail.com](mailto:ciccarelli.nicola@gmail.com). Additional results and copies of the computer programs used to generate the results presented in the paper are available from the author at [ciccarelli.nicola@gmail.com](mailto:ciccarelli.nicola@gmail.com). The usual disclaimers apply.

<sup>†</sup> Corresponding author: [ciccarelli.nicola@gmail.com](mailto:ciccarelli.nicola@gmail.com)

© 2020 Nicola Ciccarelli. Licenced under the Creative Commons Attribution-Noncommercial 3.0 Licence (<http://creativecommons.org/licenses/by-nc/3.0/>). Available at <http://rofea.org>.

ample, Catlin and Cowan (2015) report that national health care spending increased at an annual (nominal) rate of 5.4 per cent in the US during 2003–2013, and we find that employer-sponsored health insurance premiums for single health insurance coverage increased at an annual (nominal) rate of 5.7 per cent during 2003–2013 (see Kaiser/HRET (2013), Exhibit 1.11). Rapid growth in employer-sponsored health insurance premiums may harm US workers in the form of decreased employment and decreased wages. Since employer-sponsored health insurance coverage constitutes a source of costs for employers, increases in the cost of employer-sponsored health insurance may lead to lower employment and lower wages.

The impact of rising health insurance premiums for employer-sponsored health insurance on employment and wages tends to depend on how much workers value the increase in the health insurance costs paid by the employer. If workers fully value the health insurance benefit provided by the employers, workers will accept wage reductions that are equivalent to the increased employer spending on the health insurance benefit. That is, increases in the health insurance premiums will be “fully shifted” onto wages,<sup>1</sup> and employment levels will remain unvaried. There are, however, three cases under which we may expect that increases in the health insurance premiums will negatively affect employment. *First*, if rising premiums do not reflect increases in the quantity or quality of the health care provided by the employer but reflect health care inflation, employees may not fully value increases in the cost of the employer-provided health insurance benefit. In this case, workers will not accept wage reductions that are equivalent to the increased cost of the health insurance benefit, and the final labor market equilibrium might display lower employment. *Second*, the presence of labor unions is likely to restrict the employer’s ability to reduce wages as a response to rising health insurance premiums. Thus, as described in more detail below, the presence of labor unions may imply a detrimental effect of rising premiums on employment. *Third*, employers may prefer to avoid wage reductions when the health insurance premiums are rising, since wage reductions are likely to reduce working morale as well as productivity and output (see, e.g., Kube et al. (2013)). Consequently, some employers may prefer to reduce employment rather than cut wages when the health insurance premiums are rising. Finally, it is worth noting that, in addition to the potential effects on employment and wages, rising premiums may also imply a decrease of employer-sponsored health insurance coverage. Indeed, since employers provide health insurance coverage voluntarily in the period under study (i.e., 2005–2010), employers may react to increases in the health insurance premiums by dropping the health insurance coverage for their employees.

Reliable evidence on the economic impact of rapid growth in health insurance premiums for employer-sponsored health insurance is crucial to inform policy debates. To date, however,

---

<sup>1</sup> For example, Sheiner (1999) finds that increases in the cost of health insurance are fully shifted onto wages for older workers. Since older workers are likely to become sick and are characterized by high health care spending, older workers may fully value increases in the cost of the employer-sponsored health insurance benefit.

the available evidence is limited to studies that rely on repeated cross-sectional (and individual-level) datasets, and that cannot account for all sources of time-invariant unobserved heterogeneity which are correlated with the health insurance premiums and the labor market outcomes used in analysis (see, e.g., Cutler and Madrian (1998); Baicker and Chandra (2006)). In contrast to previous studies, this paper uses a newly constructed panel dataset with data on employer-sponsored health insurance premiums, employment, annual wages and private (nonelderly) health insurance coverage, which is used as a proxy for employer-sponsored health insurance coverage,<sup>2</sup> to estimate the economic impact of rapid growth in employer-sponsored health insurance premiums. All variables used in this study vary within counties and over time, except the health insurance premiums which vary within states and over time. In addition to controlling for all sources of time-invariant unobserved heterogeneity, we identify the economic impact of rising health insurance premiums by exploiting two plausibly exogenous sources of (state-level) variation in the cost of providing health insurance: We use variation in medical malpractice payments and variation in medical malpractice legislation to identify the effects of rising premiums on employment, wages and private health insurance coverage. As will be discussed below in more detail, increases in medical malpractice payments tend to increase the cost of health care and the health insurance premiums by increasing medical malpractice insurance premiums. Variation in medical malpractice legislation such as the implementation of medical malpractice tort reforms that limit doctors' malpractice liability is likely to decrease the employer-sponsored health insurance premiums by decreasing medical malpractice insurance premiums.

Using the two sources of identification outlined above, and controlling for time-invariant unobserved factors, we find that increases in the cost of health insurance imply a statistically significant decrease of the employment rate. We find that a 10% increase in the health insurance premiums implies a 1.1 percentage points decrease of the employment rate, *ceteris paribus*. On the other hand, increases in the cost of health insurance imply an insignificant decrease of annual wages. The previous results can be explained based on the following line of reasoning. Employees are not likely to fully value increases in the cost of health insurance that are driven by increases in medical malpractice payments (or by variation in medical malpractice legislation), since variation in medical malpractice payments (or variation in medical malpractice legislation) shift the price of the health care provided by employers without enhancing the quantity or quality of the same health care. Thus, increases in employer-sponsored health insurance premiums will not be fully shifted onto wages, and rising premiums will decrease employment. Since employer-sponsored health insurance premiums increased by 11.1% in real terms between 2005 and 2010, our results indicate that the increasing cost of health insurance had a strong negative effect on employment in the period 2005-2010.

---

<sup>2</sup> See the following discussion about the accuracy of private health insurance coverage as a proxy for employer-provided health insurance coverage.

Utilizing county-level estimates of private health insurance coverage as a proxy for employer-sponsored health insurance coverage, we also analyzed whether the private health insurance coverage acts as a *moderating variable* in the relationship between the health insurance premiums and the labor market outcomes used in analysis. Indeed, since increases in the cost of health insurance may negatively affect employment and wages via the increased cost of the health insurance benefit provided by the employer, rising premiums may imply stronger negative effects on employment and wages in counties that are characterized by high private coverage. We indeed find that rising premiums imply a (statistically significantly) stronger negative impact on employment and wages in counties with high private coverage. For example, we find that the effect of the health insurance premiums is approximately 30% more negative on the employment rate and annual wages for the county at the 75th percentile of private coverage than for the county at the 25th percentile of private coverage.

This paper is organized as follows. In Section 2 we outline a conceptual framework for our analysis. In Section 3, we provide a literature review on the effects of employer-sponsored health insurance premiums (or health care costs) on several labor market outcomes. Section 4 describes the data. The data sources for all variables used in this study are described in Appendix B. Section 5 describes the identification strategy used in this study. In Section 6 we discuss the empirical results for the effects of the health insurance premiums on employment and wages. Moreover, in the same section we analyze the effect of the health insurance premiums on private health insurance coverage. In Section 7 we analyze whether the private health insurance coverage is a moderating variable for the relationship between the health insurance premiums and the labor market outcomes used in analysis. Section 8 concludes.

## 2 Conceptual Framework

In this section, we first outline a microeconomic model that is the basis for the analysis of the effects of employer-sponsored health insurance premiums on employment and wages; the microeconomic model is reported in Section 2.1. Second, in Section 2.2 we outline and discuss some constraints on wage reductions that may have important implications on the effects of rising premiums on employment and wages.

### 2.1 The Microeconomic Model

Gruber and Krueger (1991) developed a microeconomic model to analyze the effect of the introduction of mandated employer-sponsored health insurance benefits on wages and employment. Subsequently, Baicker and Chandra (2006) adapted the microeconomic model developed by Gruber and Krueger (1991) to analyze the effect of increases in employer-sponsored health insurance premiums on wages and employment. In this section we briefly summarize the model of Baicker and Chandra (2006), and we refer the reader to Baicker and Chandra (2006) for further details.

The effect of increased (employer-sponsored) health insurance premiums on wages and employment depends on several factors. Let  $L_d = f_d(W + C)$  and  $L_s = f_s(W + \alpha C)$  be the labor demand and labor supply curves, respectively. The variable  $W$  represents wages,  $C$  represents the premium for the health insurance provided by the employer (which incorporates both the quantity of insurance and the price of health insurance),  $\alpha C$  represents the employees' monetary valuation of the health insurance benefit provided by the employer. Using this framework, the effect of an increase in the health insurance premium ( $C$ ) on wages is given by

$$\frac{dW}{dC} = -\frac{\eta^d - \alpha\eta^s}{\eta^d - \eta^s}, \quad (1)$$

where  $dC$  represents the change in the health insurance premium which is caused by an increase in either the price or the quantity of the health insurance provided by the employer, and where  $\eta^d$  and  $\eta^s$  are the price elasticities of labor demand and labor supply, respectively. Employment may also be affected by a change in the health insurance premium; the effect of an increase (or a decrease) in the health insurance premium on employment is given by eq. (2),

$$\frac{dL}{L} = \eta^d \frac{(W_0 - W_1 - dC)}{W_0}, \quad (2)$$

where  $W_0$  and  $W_1$  represent the initial and final levels of wages, respectively.

If there are no constraints on wages and compensation packages, and if employees fully value the health insurance benefit ( $\alpha = 1$ ), the microeconomic model outlined above predicts that wages will decrease by the full cost of the increased health insurance premium ("full shifting"), and employment levels will be unchanged. There are, however, other scenarios in which the effect of an increase in the health insurance premium will not be fully absorbed by a decrease in the wage. *First*, employees may not fully value the health insurance benefit ( $\alpha < 1$ ), or may simply assign no value to the health insurance benefit ( $\alpha = 0$ ). In these cases, the final labor market equilibrium will be characterized by both lower employment and lower wages.<sup>3</sup> *Second*, the employer's ability to reduce wages when the health insurance premiums rise may be constrained by some factors that are not taken into account in the above-mentioned model; see Section 2.2 for more details.

## 2.2 Constraints on Wage Reductions

In Section 2.1 we assumed that employers are not limited in their ability to reduce wages when the health insurance premiums rise. However, there are two cases under which the previous assumption might be violated and under which we may expect that increases in the health

---

<sup>3</sup> For simplicity, we exclude the scenario in which employees value the health insurance benefit more than they value wages ( $\alpha > 1$ ). Since US employees can use wages to buy the health insurance coverage on their own via nongroup health insurance, it is unlikely that employees will value the employer-sponsored health insurance benefit more than the wage –i.e., it is unlikely that  $\alpha$  is larger than one.

insurance premiums will not be fully absorbed by wage reductions. *First*, the presence of labor unions may restrict the employer's ability to reduce wages as a response to rising health insurance premiums. Indeed, labor market contracts with unionized workers may preclude wage cuts or even require wage increases, even if the cost of health insurance is rising (Sood et al. (2009)); moreover, competition for workers that are employed by unionized firms may further reduce the ability to absorb increases in the health insurance premiums via decreased wages (Sood et al. (2009)). Even though the US union membership rate is fairly low in the period 2005-2010 that is analyzed in this study (see, e.g., Western and Rosenfeld (2011)), the existence of labor unions is likely to represent a constraint on the ability of employers to reduce wages. Therefore, the existence of labor unions might imply that increases in the cost of the health insurance benefit will negatively affect employment levels. *Second*, some employers who face rising health insurance premiums may prefer not to reduce wages in order to avoid damaging working morale (Sood et al. (2009)). Indeed, wage reductions may reduce working morale as well as productivity and output (see, e.g., Kube et al. (2013)).<sup>4</sup> In light of the previous discussion, some employers may prefer to reduce employment (or to increase prices) rather than cut wages when the health insurance premiums are rising.

### 3 Literature Review

In this section we summarize previous studies which analyzed the effects of employer-sponsored health insurance premiums (or health care costs) on labor market outcomes that are related to the labor market outcomes used in our analysis.

Cutler and Madrian (1998) study the effect of rising health insurance costs on hours worked for the period 1979-1992. The authors use data on (per-worker) spending on health insurance for each US industry, and individual-level data on hours worked from the CPS (Current Population Survey). Using a pooled OLS regression model and assigning health insurance costs to employees based on their industry, the authors find that rising health insurance costs in the 1980s increased the hours worked of employees that are covered by employer-provided health insurance by 3 percent. Since the health insurance benefit is a fixed cost for employers at any fixed point in time, firms face an incentive to substitute hours per worker for employment when the health insurance costs rise.

Baicker and Chandra (2006) analyze the effect of rising health insurance premiums on wages, employment and other labor market outcomes for the period 1996-2002. The authors use state-level data on employer-sponsored health insurance premiums from the Medical Expenditure Panel Survey, and individual-level data on labor market outcomes from the CPS. Using a pooled OLS regression model with state fixed effects, and using instrumental variables

---

<sup>4</sup> Via a natural field experiment, Kube et al. (2013) show that wage reductions have a detrimental and persistent impact on productivity and firm output. Furthermore, the authors interpret the aforementioned effects as evidence that wage cuts damage working morale (see p. 11 of Kube et al. (2013)).

for the health insurance premiums, the authors find a substantial negative effect of rising health insurance premiums on employment and wages. Indeed, a 10% increase in the health insurance premiums reduces the probability of being employed by 1.2 percentage points, and implies a 2.3% decrease of wages for employees that are covered by employer-provided health insurance –i.e., rising premiums are fully shifted onto wages for these employees. On the other hand, rising premiums do not have a significant effect on wages for the full sample of respondents included in the CPS (see Table 2 of Baicker and Chandra (2006)). Finally, it is worth noting that increases in the cost of health insurance do not imply a decrease of employer-sponsored health insurance coverage for the full sample of individuals included in the CPS.

Sood et al. (2009) analyze the effect of rising health care costs on employment, gross output and value added of US industries for the period 1987-2005. The authors use data on health care costs for each US industry from the National Health Expenditures Accounts (NHEA) database, and use data on employment, gross output and value added for each US industry from the Bureau of Economic Analysis' Annual Industry Accounts database. The authors find substantial negative effects of rising health care costs on industry-level employment, gross output and value added, and these effects are stronger for industries in which large percentages of workers receive employer-provided health insurance.

#### **4 Data**

We use a panel of annual variables for counties in the 50 US states plus the District of Columbia. All variables used in this study vary at the county-year level, except the employer-sponsored health insurance premiums which vary at the state-year level. County-level estimates on employer-sponsored health insurance premiums are not available from governmental agencies, and we use state-level estimates for employer-sponsored health insurance premiums from the MEPS-IC database (see discussion below for more details). We focus our analysis on data for the period 2005-2010, since county-level estimates for private (nonelderly) health insurance coverage cannot be computed for the years before 2005 and the years after 2010.<sup>5</sup>

##### ***Employer-Sponsored Health Insurance Premiums***

We use annual state-level data on employer-sponsored health insurance premiums from the Medical Expenditure Panel Survey-Insurance Component for the period 2005-2010. The main explanatory variable used in this study is the state-level average of annual health insurance premiums for single (health insurance) coverage.<sup>6</sup> Data on health insurance premiums for the

---

<sup>5</sup> See the discussion below and Appendix B.3 for more details on data restrictions for private health insurance coverage. Moreover, data on control variables from the U.S. Census Bureau's "USA Counties" database are missing for the years after 2010, since the U.S. Census Bureau has terminated its support for the "USA Counties" database in 2010.

<sup>6</sup> Annual health insurance premiums for single coverage are the sum of the employee and employer contributions for single coverage. Employer contributions for health insurance amounted to approximately



year 2007 are missing for all US states. As in Burtless and Mikusheva (2012), data for the health insurance premiums in year 2007 are imputed as the average of the health insurance premiums in year 2006 and year 2008 (by state).

Since we use state-level estimates of health insurance premiums as a proxy for the (unknown) county-level premiums, our main explanatory variable is very likely to be measured with error; more details on the problem of measurement error in the health insurance premiums are provided in Section 5. Using the Bureau of Labor Statistics' CPI inflation calculator, we express the health insurance premiums in constant 2005 US dollars, and we additionally express the health insurance premiums in 1,000 US dollars. As shown in Panel A of Table 1, the average of the (real) health insurance premiums expressed in 1,000 US dollars is 4.040, that is, premiums for single coverage averaged 4,040 US dollars in the period 2005–2010. The health insurance premiums display high variability as they range between a minimum value of 3,384 US dollars and a maximum value of 5,505 US dollars in the period 2005–2010 (see Table 1). Moreover, the health insurance premiums are characterized by high variability over time. Indeed, as shown in the first row of Table 2, the yearly average of the health insurance premiums varies considerably over time and grows between 2006 and 2010. Lastly, in Appendix B.1 we report more details about the data sources and the data imputation technique that is used for the health insurance premiums.

### ***Labor Market Outcomes***

The labor market outcomes used in this study are the employment rate and annual wages. These two variables are used as they naturally arise from the microeconomic model reported in Section 2.1. Both the employment rate and annual wages vary within counties and over time.

The *employment rate* is constructed as the ratio between the number of employed persons and the number of 15–64 year old persons (by county). Data on the number of employed persons for each county is obtained from the Bureau of Labor Statistics' QCEW (Quarterly Census of Employment and Wages) dataset. Data on the number of 15–64 year old persons living in each county is obtained from the US Census Bureau's "USA Counties" dataset. The QCEW dataset does not include self-employed workers in the number of employed persons (see also Retzer et al. (2013), p. 169), thus the average of the employment rate used in this study is lower than the average of the employment rate found in official statistics for the United States. The average of the employment rate is 0.51 (see Panel A of Table 1). The yearly averages of the employment rate display considerable variability over time, since they range between a minimum value of 0.50 in the year 2010 to a maximum value of 0.53 in the year 2007 (see Table 2).

County-level estimates of *annual wages* are obtained from the QCEW dataset. In the QCEW dataset, annual wages are computed as the ratio between total annual wages and total annual

---

87% of the total premium for single coverage in the period 1999–2008 (see Claxton et al. (2007), p. 1410). The relative shares of the total premium paid by employers and employees remained approximately constant in the period 1999–2008 (Claxton et al. (2008), Exhibit 6.1, p. 70).

employment (by county). Using the Bureau of Labor Statistics' CPI inflation calculator, we express annual wages in constant 2005 US dollars. The average of annual wages is close to 29,800 US dollars in real terms in the period 2005-2010 (see Panel A of Table 1). The yearly averages of the annual wages exhibit limited variability over time, and they range between 29,229 US dollars and 29,975 US dollars. Finally, in Appendix B.2 we report more details about the data sources for the employment rate and annual wages.

### ***Private (Nonelderly) Health Insurance Coverage***

County-level estimates of employer-sponsored health insurance coverage are not available from US governmental agencies. For this reason we use county-level estimates of private nonelderly health insurance coverage as a proxy for (county-level) employer-sponsored health insurance coverage. Our estimates for private nonelderly health insurance coverage are obtained using data from the US Census Bureau's SAHIE (Small Area Health Insurance Estimates) dataset and the US Department of Health & Human Services' AHRF (Area Health Resource File) dataset. In Appendix A.2 we explain in detail how we constructed the county-level estimates for private nonelderly health insurance coverage.

Our county-level estimates for **private health insurance (PHI) coverage** represent the county-level share of nonelderly (0-64) individuals that are covered by either **employer-sponsored health insurance (ESHI)** or **nongroup health insurance (NHI)**. Even though our estimates for the PHI coverage are not identical to the ESHI coverage by definition, it is worth mentioning that the vast majority of privately insured individuals are covered by ESHI in the United States. Indeed, DeNavas-Walt et al. (2008) report that approximately 88% of US individuals covered by PHI are covered by ESHI while only 12% of US individuals covered by PHI are covered by NHI.

For two reasons which are discussed below in detail, it is highly likely that our county-level estimates of PHI coverage are a relatively accurate proxy of county-level ESHI coverage. *First*, Blavin et al. (2012) reports that NHI coverage remained relatively constant in the period 2000-2010.<sup>7</sup> Since the identification strategy used in this paper accounts for county fixed effects, and based on the above-mentioned finding reported in Blavin et al. (2012), it is highly likely that changes in PHI coverage over time are primarily driven by changes in ESHI coverage over time. *Second*, using data from March CPS datasets for the period 2006-2011, we computed estimates of *state-level ESHI coverage* for the years 2005 to 2010 (see Appendix B.4 for more details about these estimates), and we have found that the correlation between our county-level estimates of PHI coverage and the state-level estimates of ESHI coverage from the March CPS

---

<sup>7</sup> Blavin et al. (2012), p.4: "Figures 1 through 4 show trends between 2000 and 2010 in ESI [employer-sponsored insurance], Medicaid, and uninsurance among parents, adults without dependent children, and children. Rates of nongroup and other public coverage were relatively constant over the period of analysis and are therefore not included in any of the figures."

datasets ranges between 0.55 and 0.82.<sup>8</sup> The latter result seems to suggest that our county-level estimates of PHI coverage are a relatively accurate proxy for the (unknown) county-level ESHI coverage. As shown in Panel A of Table 1, the average of PHI coverage is 0.621, that is, approximately 62.1% of nonelderly US individuals are covered by PHI in the period 2005–2010.<sup>9</sup> Finally, the average of the PHI coverage decreases from a value of 0.63 in the year 2005 to a value of 0.60 in the year 2010 (see Table 2).

Table 1: Summary Statistics for the Health Insurance Premiums, Labor Market Outcomes and Private Health Insurance Coverage

|   | Obs.  | Mean     | Std. dev. | Min      | Max      |
|---|-------|----------|-----------|----------|----------|
| Health insurance premiums in 1,000 US \$ <sup>†</sup> | 15366 | 4.040    | 0.300     | 3.384    | 5.505    |
| Employment rate <sup>††</sup>                         | 15366 | 0.514    | 0.204     | 0.233    | 0.886    |
| Annual wages in US dollars <sup>§</sup>               | 15366 | 29781.62 | 6645.618  | 19844.16 | 53023.25 |
| Private health insurance coverage (share)             | 15366 | 0.621    | 0.117     | 0.391    | 0.811    |

<sup>†</sup>The health insurance premiums are expressed in real terms: We discount the nominal health insurance premiums back to the year 2005 using the Bureau of Labor Statistics' CPI inflation calculator, and we express the health insurance premiums in 1,000 US dollars.

<sup>††</sup>The employment rate is constructed as the ratio between the number of employed persons (QCEW dataset) and the number of 15–64 year old persons ("USA counties" dataset).

<sup>§</sup>Annual wages are expressed in real terms: we discount nominal wages back to the year 2005 using the Bureau of Labor Statistics' CPI inflation calculator.

<sup>8</sup> First, we computed the correlation between the county-level PHI coverage and the state-level ESHI coverage from the March CPS, and we have found that the correlation is 0.55. Second, since the ESHI coverage estimates from the March CPS vary at the state (and year) level while our estimates for the county-level PHI coverage vary at the county (and year) level, we averaged the PHI coverage at the state-year level, and we computed the correlation between the latter variable and the state-level ESHI coverage from the March CPS. In this case, we have found that the correlation is 0.82.

<sup>9</sup> Summary statistics for total and public (nonelderly) health insurance coverage are reported in Table C1. Around 82% of nonelderly individuals are covered by any form of health insurance in the period 2005–2010, and around 20% of nonelderly individuals are covered by public health insurance (i.e., Medicaid, Children's Health Insurance Program) in the period 2005–2010.

Table 2: Yearly Averages and Yearly Standard Deviations of the Dependent and Explanatory Variables

|  | Year 2005 |          | Year 2006 |          | Year 2007 |          | Year 2008 |          | Year 2009 |          | Year 2010 |          |
|--|-----------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|----------|
|  | Mean      | SD       | Mean      | SD       | Mean      | SD       | Mean      | SD       | Mean      | SD       | Mean      | SD       |
| Premium in 1,000 US dollars <sup>†</sup> | 3.916     | 0.247    | 3.896     | 0.224    | 3.900     | 0.211    | 3.904     | 0.231    | 4.156     | 0.276    | 4.351     | 0.242    |
| Employment rate <sup>††</sup>            | 0.519     | 0.195    | 0.523     | 0.199    | 0.526     | 0.212    | 0.522     | 0.209    | 0.499     | 0.198    | 0.496     | 0.197    |
| Annual wages in US dollars <sup>§</sup>  | 29228.500 | 6470.887 | 29318.600 | 6586.976 | 29927.950 | 6786.698 | 29736.440 | 6659.372 | 29956.890 | 6561.290 | 29974.720 | 6616.342 |
| Private h.i. coverage (share)            | 0.630     | 0.109    | 0.627     | 0.110    | 0.633     | 0.107    | 0.624     | 0.118    | 0.614     | 0.120    | 0.604     | 0.126    |
| Male                                     | 0.498     | 0.021    | 0.498     | 0.021    | 0.498     | 0.021    | 0.499     | 0.021    | 0.499     | 0.022    | 0.499     | 0.022    |
| Age (15-19)                              | 0.073     | 0.012    | 0.073     | 0.012    | 0.072     | 0.011    | 0.071     | 0.0116   | 0.070     | 0.011    | 0.069     | 0.011    |
| Age (20-24)                              | 0.064     | 0.026    | 0.063     | 0.026    | 0.062     | 0.026    | 0.061     | 0.026    | 0.060     | 0.026    | 0.060     | 0.026    |
| Age (65+)                                | 0.150     | 0.041    | 0.151     | 0.041    | 0.153     | 0.041    | 0.155     | 0.041    | 0.157     | 0.041    | 0.158     | 0.041    |
| Population density                       | 231.720   | 1688.366 | 232.584   | 1689.067 | 234.226   | 1693.960 | 236.179   | 1703.819 | 238.060   | 1710.640 | 239.820   | 1720.343 |
| White non-Hispanic (share)               | 0.798     | 0.194    | 0.795     | 0.195    | 0.792     | 0.196    | 0.788     | 0.197    | 0.786     | 0.198    | 0.782     | 0.199    |
| White Hispanic (share)                   | 0.067     | 0.122    | 0.069     | 0.123    | 0.071     | 0.124    | 0.072     | 0.124    | 0.074     | 0.125    | 0.075     | 0.126    |
| Black non-Hispanic (share)               | 0.087     | 0.144    | 0.088     | 0.144    | 0.088     | 0.144    | 0.088     | 0.145    | 0.088     | 0.144    | 0.088     | 0.144    |
| Black Hispanic (share)                   | 0.002     | 0.004    | 0.002     | 0.004    | 0.002     | 0.004    | 0.002     | 0.004    | 0.002     | 0.004    | 0.003     | 0.004    |
| Asian non-Hispanic (share)               | 0.010     | 0.023    | 0.011     | 0.024    | 0.011     | 0.024    | 0.011     | 0.025    | 0.011     | 0.025    | 0.012     | 0.025    |
| Asian Hispanic (share)                   | 0.0004    | 0.001    | 0.0004    | 0.001    | 0.0004    | 0.001    | 0.0005    | 0.001    | 0.0005    | 0.001    | 0.0005    | 0.001    |
| Other non-Hispanic (share)               | 0.031     | 0.081    | 0.031     | 0.081    | 0.032     | 0.081    | 0.033     | 0.081    | 0.033     | 0.082    | 0.035     | 0.081    |
| Other Hispanic (share)                   | 0.004     | 0.005    | 0.004     | 0.005    | 0.004     | 0.006    | 0.005     | 0.006    | 0.005     | 0.006    | 0.005     | 0.009    |
| Share of small firms                     | 0.920     | 0.035    | 0.918     | 0.035    | 0.916     | 0.036    | 0.913     | 0.037    | 0.912     | 0.038    | 0.909     | 0.037    |
| Number of observations per year          | 3079      |          | 3079      |          | 3080      |          | 3080      |          | 3080      |          | 3048      |          |

Notes: this table displays the yearly averages and yearly standard deviations for the dependent variables and the explanatory variables that are used in our analysis. See Section 4 for more detailed information about the dependent and explanatory variables that are used in this study.

<sup>†</sup>The health insurance premiums are expressed in real terms: We discount the nominal health insurance premiums back to the year 2005 using the Bureau of Labor Statistics' CPI inflation calculator, and we express the health insurance premiums in 1,000 US dollars.

<sup>††</sup>The employment rate is constructed as the ratio between the number of employed persons (QCEW dataset) and the number of 15-64 year old persons ("USA counties" dataset).

<sup>§</sup> Annual wages are expressed in real terms: we discount nominal wages back to the year 2005 using the Bureau of Labor Statistics' CPI inflation calculator.

### Control Variables

We control for an extensive set of variables that may be correlated to the labor market outcomes and the health insurance premiums. All control variables used in this study vary at the county-year level. We report panel-level summary statistics for the control variables in Table 3, and we report the yearly averages and yearly standard deviations for the control variables in Table 2.

We control for several *socio-economic variables*. Data on *socio-economic variables* are obtained from the US Census Bureau’s “USA counties” data set (see Appendix B.5 for more details). Socio-economic control variables include: (1) share of males; (2) three variables for the age distribution (i.e., share of 15–19 year old persons, share of 20–24 year old persons, share of 65+ year old persons); (3) population density (i.e., population per square mile); (4) eight variables for the racial composition of the county population. Individuals can be White, Black, Asian, of another race; moreover, individuals of any race can be either (I) Hispanic, or (II) non-Hispanic. We control for eight variables indicating the county-level share of individuals belonging to a specific combination of race and Hispanic origin; the baseline variable is the share of non-Hispanic Whites.

We control for an additional variable that in this study is defined as *share of small firms*. This variable is the ratio between the number of small firms and total number of firms (by county). In line with the classification of firm size used in the US, firms that employ less than 500 employees are defined as *small firms* for the construction of the variable *share of small firms*. Since small (big) firms have low (high) bargaining power in extracting price concessions from health insurers (Stevens (2003), p. 74), it is likely that the *share of small firms* is correlated with the health insurance premiums (see also Lee (2002)). Lastly, the data sources that are used to construct *share of small firms* are reported in Appendix B.1.

Table 3: Summary Statistics for the Control Variables

|   | Obs.  | Mean    | Std. dev. | Min   | Max   |
|---|-------|---------|-----------|-------|-------|
| Male (share)                              | 15366 | 0.499   | 0.022     | 0.426 | 0.721 |
| Age distribution:                         |       |         |           |       |       |
| Age (15–19) (share)                       | 15366 | 0.072   | 0.011     | 0.022 | 0.157 |
| Age (20–24) (share)                       | 15366 | 0.062   | 0.025     | 0.013 | 0.283 |
| Age (65+) (share)                         | 15366 | 0.154   | 0.041     | 0.029 | 0.434 |
| Population density (pop. per square mile) | 15366 | 235.939 | 1702.654  | 0.038 | 69559 |
| Race/Hispanic origin variables:           |       |         |           |       |       |
| White non-Hispanic (share)                | 15366 | 0.791   | 0.196     | 0.025 | 0.994 |
| White Hispanic (share)                    | 15366 | 0.072   | 0.125     | 0     | 0.959 |
| Black non-Hispanic (share)                | 15366 | 0.087   | 0.143     | 0     | 0.859 |
| Black Hispanic (share)                    | 15366 | 0.002   | 0.004     | 0     | 0.133 |
| Asian non-Hispanic (share)                | 15366 | 0.011   | 0.023     | 0     | 0.440 |
| Asian Hispanic (share)                    | 15366 | 0.0005  | 0.001     | 0     | 0.015 |
| Other non-Hispanic (share)                | 15366 | 0.032   | 0.081     | 0     | 0.970 |
| Other Hispanic (share)                    | 15366 | 0.005   | 0.006     | 0     | 0.070 |
| Share of small firms                      | 15366 | 0.915   | 0.036     | 0.531 | 1     |

**Panel Structure**

The longitudinal dataset used in this study is almost balanced, and we use 15,366 observations in analysis. For each year, our longitudinal dataset contains data on slightly more than 3,040 counties (see Table A1). This represents almost 97% of the total number of US counties, since the total number of US counties is 3,142 in the period under study.

**5 Empirical Strategy**

We start this section by providing details on the estimation methodology used in analysis; see Section 5.1. In Section 5.2 we provide details on the main identification strategy used in analysis.

**5.1 Methodology**

As reported in Section 4, all dependent and explanatory variables used in this study vary at the county-year level, except the health insurance premiums which vary at the state-year level. Consider the following regression model for county  $c$  (and state  $s$ ) in year  $t$ :

$$y_{ct} = \beta_1 \text{premiums}_{st} + \text{controls}_{ct}^\top \beta_2 + T_t + v_{ct}, \quad \text{where } v_{ct} = \alpha_c + \epsilon_{ct}. \quad (3)$$

Here,  $y_{ct}$  is the labor market outcome of interest (employment rate or annual wages) or private health insurance coverage. The expression  $\text{premiums}_{st}$  represents the health insurance premiums in state  $s$  and year  $t$ . The expression  $\text{controls}_{ct}$  measures county-level covariates, including socio-economic controls and share of small firms. The expression  $T_t$  represents year fixed effects, and  $v_{ct}$  is a composite error term which is composed by time-invariant county-level unobserved heterogeneity ( $\alpha_c$ ) and a county-specific idiosyncratic error ( $\epsilon_{ct}$ ).

As a first step, we use pooled ordinary least squares (OLS) to study the association between the health insurance premiums and the dependent variables. In the case of pooled OLS estimation, time-invariant county-level unobserved heterogeneity ( $\alpha_c$ ) is treated as part of the error term, and is not eliminated from the regression model – i.e., the error term is the composite error term ( $v_{ct}$ ). Since the error term in eq. (3) may be heteroskedastic and clustered, we use cluster-robust standard errors; the *standard errors are clustered at the state level* since our main explanatory variable – i.e., the health insurance premiums – varies at the state level. As time-invariant and unobserved county characteristics might be correlated with the dependent variable and the health insurance premiums, the OLS estimates cannot be interpreted as causal.

Time-invariant unobserved heterogeneity can be eliminated from the regression model (3) using within-group transformation or first-difference transformation of the longitudinal dataset. We use the latter estimation methodology, since in this case we can use both “external instrumental variables” and “internal instrumental variables” for the endogenous variables: External instrumental variables are represented by random variables that do not arise from the regression model (3) while internal instrumental variables arise from the regression model itself in the

form of lagged explanatory variables (see, e.g., p. 372 of Wooldridge (2010)).<sup>10</sup> The panel data model in first differences for county  $c$  (and state  $s$ ) in year  $t$  takes the following form:

$$\Delta y_{ct} = \beta_1 \Delta \text{premium}_{s_{st}} + \Delta \text{controls}_{s_{ct}}^\top \beta_2 + \Delta T_t + \Delta \epsilon_{ct}, \quad (4)$$

and the FD (first difference) estimator is the OLS estimator applied to equation (4). Since the error term in eq. (4) may be heteroskedastic and clustered, we use cluster-robust standard errors, and the *standard errors are clustered at the state level*.

Since the estimate of the main parameter of interest ( $\beta_1$ ) may be biased due to endogeneity problems (e.g., measurement errors in the health insurance premiums), we will also analyze the effect of the health insurance premiums on the dependent variables using instrumental variables for the health insurance premiums; see Section 5.2 for more details.

## 5.2 Identification Strategy

Measurement errors in the health insurance premiums may introduce a bias toward zero in the estimated coefficient on the premiums. The true (and unknown) county-level premiums are the sum of the state-level premiums and an error term which varies at the county level:

$$\text{premium}_{s_{ct}} = \text{premium}_{s_{st}} + m_{ct}, \quad (5)$$

where  $s$  indicates the state,  $c$  indicates the county, and  $t$  indicates the year. Here,  $m_{ct}$  represents an error term that is idiosyncratic to county  $c$  in year  $t$ . Since our preferred identification strategy relies on first difference estimation (see above), eq. (5) can be replaced by eq. (6),

$$\Delta \text{premium}_{s_{ct}} = \Delta \text{premium}_{s_{st}} + \Delta m_{ct}, \quad (6)$$

where  $\Delta m_{ct}$  denotes the measurement error in first differences. If measurement error in levels ( $m_{ct}$ ) is a unit root process – i.e.,  $m_{ct} = m_{c,t-1} + \xi_{ct}$  where  $\xi_{ct}$  is a white noise process – then measurement errors in the health insurance premiums would not lead to a bias toward zero in the estimated coefficient on the premiums. Excluding the above case, there are, however, several cases in which measurement error bias may arise, even in the case that we estimate the parameters of interest using the regression model in first differences (4). Thus, for simplicity we assume that measurement errors are likely to affect our empirical results – i.e., we do not assume that  $\text{Cov}(\Delta \epsilon_{ct}, \Delta m_{ct}) = 0$ .

In the case that the dependent variable is either annual wages or the employment rate, an additional source of endogeneity may be reverse causality bias. In the case of annual wages,

<sup>10</sup> Internal instrumental variables require an additional validity assumption compared to external instrumental variables: The error term in levels ( $\epsilon_{ct}$ ) must be serially uncorrelated. For this reason, we use internal instruments only in the case that external instruments are jointly weak in the first-stage regression. Internal instruments are used exclusively in Section 7. See Section 7 for more details about internal instruments.

reverse causality bias may be present for the following reason. Since increases in income or wages imply rising health care spending (see, e.g., Acemoglu et al. (2013)), and since premiums depend positively on health care spending,<sup>11</sup> we believe that increases in annual wages are likely to increase the health insurance premiums. Second, following the same line of reasoning reported above, increases in (local) employment rates are likely to increase the health insurance premiums by increasing (local) average income.

A solution to the problem of endogeneity bias is to instrument for the health insurance premiums using variables that are uncorrelated with  $\Delta\epsilon_{ct}$  and  $\Delta m_{ct}$  but are correlated with the health insurance premiums. In our analysis we use *per capita medical malpractice payments* and variation in *medical malpractice legislation* within states and over time to generate external instrumental variables for the health insurance premiums. In the following discussion, we describe the external instrumental variables and we explore the exogeneity of the external instrumental variables. Moreover, we report and discuss the first-stage estimates for the effects of the instrumental variables on the health insurance premiums.

### ***The External Instrumental Variables and the First-Stage Estimates***

#### *Variation in Medical Malpractice Legislation*

Medical malpractice tort reforms tend to decrease medical malpractice payments by limiting doctors' medical malpractice liability. Indeed, medical malpractice tort reforms usually set caps on the awards that the victim of malpractice can obtain from the negligent doctor (see Table A3 for an example of tort reform). Moreover, and more importantly, Kessler and McClellan (1996) show that medical malpractice tort reforms decrease the likelihood that doctors practice "defensive medicine" – i.e., to perform excessive tests and procedures because of concerns about malpractice liability. Since the demand for health care is relatively inelastic (Baicker and Chandra (2005)), the decreased costs of medical malpractice, which are caused by the implementation of malpractice tort reforms, will be passed along to consumers of health care in the form of lower health care prices and lower health insurance premiums (Baicker and Chandra (2006)). In line with this theoretical argument, Avraham et al. (2012) find that the implementation of medical malpractice tort reforms reduces employer-sponsored health insurance premiums.

Several states enacted a cap on non-economic damage awards since 2005 (see Table A4).<sup>12</sup> Non-economic damage awards are awards that the victim of malpractice can obtain from the negligent doctor for pain and suffering caused by the negligent doctor. Based on the previous discussion, states that enacted caps on non-economic damage awards may have experienced a decrease of employer-sponsored health insurance premiums. We construct an instrumental variable for the health insurance premiums using the introduction of caps on non-economic

<sup>11</sup> Indeed, health insurers set the health insurance premiums equal to health care spending plus a markup over costs.

<sup>12</sup> See Appendix B.6 for details on data sources for medical malpractice tort reforms.



damage awards. **First**, we generated a dummy variable (*caps on damages<sub>st</sub>*) which is equal to 1 for state *s* in the year of the reform and in successive years, 0 otherwise. For instance, as reported in Table A4, *caps on damages<sub>st</sub>* takes the value of 0 for the state of Alaska in 2005, and *caps on damages<sub>st</sub>* takes the value of 1 for the state of Alaska in 2006–2010 since a cap on non-economic damage awards was introduced in Alaska in the year 2006. **Second**, since health insurance contracts are typically finalized 3+ months prior to the calendar year in which they take effect (Avraham et al. (2012), p. 13), the final instrumental variable takes a value of 1 in the year following the implementation of the tort reform (and successive years). That is, we use *caps on damages<sub>s,t-1</sub>* as the instrumental variable for the health insurance premiums (see Avraham et al. (2012) for a similar specification).

Medical malpractice tort laws can also be repealed by state supreme courts if a medical malpractice tort law is incompatible with the state or federal constitution, or with state or federal laws. Following the same line of reasoning reported above, repeals of caps on non-economic damage awards may lead to higher health insurance premiums. We constructed an additional instrumental variable for the health insurance premiums using repeals of caps on non-economic damage awards; information about repeals of caps on non-economic damage awards is provided in the second row of Table A4. We use *repeals<sub>s,t-1</sub>* as the instrumental variable for the health insurance premiums, where *repeals<sub>s,t-1</sub>* takes a value of 1 in the year following the repeal of the tort reform (and successive years). In line with the theoretical arguments reported above, we find that enactments of caps on non-economic damages are negatively (but insignificantly) correlated to the health insurance premiums, and we find that repeals of caps on non-economic damage awards are positively and significantly correlated to the health insurance premiums at the 5% significance level; see the regression results in Table D1.<sup>13</sup> Finally, we find that repeals of caps on non-economic damage awards lead to a 2.7% statistically significant increase of the health insurance premiums.

The passage of tort reforms is a political process whose outcome depends on political and ideological factors that are unlikely to be related to the labor market outcomes used in analysis. The main opposition to the implementation of tort reforms, such as the implementation of caps on non-economic damage awards, is based on the argument that doctors will operate less carefully after the implementation of tort reforms. Moreover, states that enact tort reforms may attract the riskiest doctors from other US states (see, e.g., Seabury (2009)). Political and ideological factors that are unrelated to the labor market outcomes used in our analysis are the primary factors that determined the timing of the implementation of caps on non-economic damage awards. Moreover, supreme courts' decisions about the repeal of caps on non-economic damage awards are unlikely to be affected by present or past trends in employment or wages,

<sup>13</sup> The regression results reported in Table D1 are the estimates of equation (4), with the health insurance premiums on the left hand side of equation (4), and with *caps on damages<sub>s,t-1</sub>* and *repeals<sub>s,t-1</sub>* on the right hand side of equation (4).

and are likely to be related to innovations in the (state or federal) law environment. Consequently, we believe that the instrumental variables reported above are unlikely to be correlated with the error term on the labor market outcomes used in analysis.

#### *Per Capita Medical Malpractice Payments*

An external instrumental variable that is used in this study is per capita medical malpractice payments. Per capita medical malpractice payments are likely to significantly affect the health insurance premiums for the following reasons. Rising medical malpractice payments imply an increase of medical malpractice insurance premiums; since the demand for health services is relatively inelastic (Baicker and Chandra (2005)), rising medical malpractice insurance premiums will be passed along to consumers of health care in the form of higher health care prices and higher health insurance premiums (Baicker and Chandra (2006)).

The usage of (state-level) per capita medical malpractice payments as an instrument for the health insurance premiums is likely to remove the bias from any residual correlation between the health insurance premiums ( $\Delta premiums_{st}$ ) and the error term of the structural equation ( $\Delta \epsilon_{ct}$ ). Indeed, the instrument ( $\Delta payments_{st}$ ) only picks up that part of the within-state variation in premiums that is caused by changes in the state malpractice environment. Thus, it is likely that variation in medical malpractice payments over time ( $\Delta payments_{st}$ ) are uncorrelated to the error term of the structural equation.

The data for per capita medical malpractice payments are obtained from the National Practitioner Data Bank (NPDB) for the period 2005-2010 (see Appendix B.6 for more details on data sources for malpractice payments). Per capita medical malpractice payments are the ratio between total malpractice payments in each US state and the state population. Moreover, we express per capita medical malpractice payments in real terms (constant 2005 US dollars). Summary statistics for medical malpractice payments are reported in Table A2. The average of per capita medical malpractice payments is close to 10 US dollars in the period 2005-2010.

#### *First-stage Estimates*

Our preferred instrumental variable for the health insurance premiums is per capita medical malpractice payments, since per capita medical malpractice payments vary over time and within states for all US states, while the remaining instrumental variables affect only few US states in the period 2005-2010. The instrumental variables which are based on variation in medical malpractice legislation are used in order to improve the first-stage model fit and test the over-identifying restrictions. Thus, in the following discussion we will report first-stage estimates for the effects of the instruments on the health insurance premiums using two sets of instrumental variables: **(a)** per capita medical malpractice payments (alone); **(b)** all three instrumental variables described above. First stage estimates using per capita medical malpractice payments as the exclusive instrumental variable are shown in column 1 of Table 4 while first stage estimates using all three instrumental variables are shown in column 2 of Table 4.

Table 4: First-stage FDIV Estimates

|  | (1)<br>FDIV (first-stage) | (2)<br>FDIV (first-stage) |
|--|---------------------------|---------------------------|
|  | Dependent variable        |                           |
|  | HI Premium (in 1,000 \$)  | HI Premium (in 1,000 \$)  |
| Medical malpractice payments                           | 0.011***<br>(0.0038)      | 0.011***<br>(0.0038)      |
| Caps on damages (enactments)                           | -                         | 0.015<br>(0.048)          |
| Caps on damages (repeals)                              | -                         | 0.125***<br>(0.016)       |
| F-statistic on the excluded instrument(s) <sup>‡</sup> | 7.966                     | 21.485                    |
| P-value  | 0.0069                    | 0.0000                    |
| Observations   | 15,366                    | 15,366                    |
| Adjusted R-squared                                     | 0.399                     | 0.401                     |
| Year fixed effects?                                    | YES                       | YES                       |
| County fixed effects? <sup>†</sup>                     | YES                       | YES                       |
| Controls? <sup>††</sup>                                | ALL                       | ALL                       |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. The outcome variable is health insurance premiums. Year FE (fixed effects) are included in the regression model. The first-stage (FDIV) estimates are obtained using the Stata<sup>TM</sup> command “ivreg2, cluster(state identifier)” using the model in first differences.

<sup>‡</sup>The F-statistic on the excluded instrument(s) is the Kleibergen-Paap rank Wald F statistic. The Kleibergen-Paap rank Wald F statistic takes into account the usage of cluster-robust standard errors.

<sup>†</sup>County FE (fixed effects) are eliminated from the regression model using first-difference transformation of the regression model.

<sup>††</sup>Control variables include eight variables for race and Hispanic origin, share of males, three variables for the population distribution by age, population density, share of small firms.

In the case that we exclusively use per capita medical malpractice payments to instrument for the health insurance premiums (see column 1 of Table 4), we find that per capita medical malpractice payments are significantly correlated to the health insurance premiums at the 5% significance level. The first-stage estimate reported in column 1 of Table 4 suggests that when medical malpractice payments double, health insurance premiums increase by 2.7%. It is worthwhile noting that the latter result is quite similar to the effect of per capita medical malpractice payments on health insurance premiums found by Baicker and Chandra (2006), as the latter authors have found that a 100% increase of medical malpractice payments leads to a 1%-2% increase of the health insurance premiums (see pp. 622-624 of Baicker and Chandra (2006)).

As shown in column 1 of Table 4, we find that the first-stage  $F$ -statistic is 7.966 in the case that we exclusively use per capita medical malpractice payments to instrument for the health insurance premiums. On the contrary, the first-stage  $F$ -statistic is larger than 10 ( $F$ -

statistic=21.485) in the case that we use all available instruments (see column 2 of Table 4), indicating that the instrumental variables are jointly strong to predict the health insurance premiums in the first-stage regression. Since the first-stage  $F$ -statistic is lower than 10 when we exclusively use per capita medical malpractice payments to instrument for the health insurance premiums, we also estimated LIML (limited information maximum likelihood) models, using per capita medical malpractice payments as the exclusive instrument for the health insurance premiums. The second-stage results from LIML models were approximately identical to the second-stage results from standard TSLS (two-stage least squares) models. Since LIML estimation is more reliable than TSLS estimation when the instrumental variable is weak (see Stock et al. (2002)), it is very likely that the second-stage estimates (based on medical malpractice payments as the exclusive instrumental variable) are not attributable to the weak instrumentation problem.

## 6 Empirical Results

In this section we analyze the effects of the health insurance premiums on the labor market outcomes and private health insurance coverage. We start this section by discussing the effects on labor market outcomes in Section 6.1. Since increases in the health insurance premiums may lead to a decline of private health insurance coverage, we report and discuss estimates for the effect of the health insurance premiums on private health insurance coverage in Section 6.2. Lastly, in Section 6.3 we explore the robustness of our identification strategy.

### 6.1 The Causal Effects of the Health Insurance Premiums on Employment and Wages

We begin with an examination of the effect of increases in the health insurance premiums on the **employment rate**. Column 1 of Table 5 shows the results of the OLS estimation of equation (3). This regression includes year fixed effects but does not account for county fixed effects. We find that premiums are not significantly correlated with the employment rate, and the coefficient on premiums is positive. The OLS estimate cannot be interpreted as causal, since the health insurance premiums and the employment rate may be correlated with unobserved county characteristics that do not vary over time.<sup>14</sup>

Once we additionally control for time-invariant unobserved county characteristics using the FD estimator (see eq. (4)), we find that increases in the health insurance premiums lead to a significant decrease of the employment rate at the 5% significance level (see column 2 of

---

<sup>14</sup> Since a direct comparison of pooled OLS and FD (first differences) estimates cannot be implemented, we compared RE (random effects) and FE (fixed effects) estimates via the Hausman test, and we have found that RE and FE estimates are systematically different. Thus, time-invariant county characteristics seem to be correlated with the explanatory variables and the dependent variable. In turn, this implies that the results from the pooled OLS estimator are unreliable. RE and FE estimates are not reported and are available from the author.

Table 5). This estimate suggests that a 10% increase in premiums implies a 0.5% decrease (0.2-percentage-points decrease) of the employment rate. However, this estimate may be biased due to endogeneity problems. First, the FD estimate may be biased toward zero due to measurement errors in the health insurance premiums. Second, reverse causality may lead to an upward bias in the coefficient on the health insurance premiums. As a matter of fact, increases in local employment rates are very likely to imply increases in local incomes, which in turn are likely to increase health care spending<sup>15</sup> and the health insurance premiums (see also Section 5.2).

Since the health insurance premiums may be endogenous with respect to the error term on the employment rate, we also used instrumental variables to account for the potential endogeneity of premiums. We start by using per capita medical malpractice payments as the exclusive instrument for the health insurance premiums; see column 1 of Table 4 for the first-stage regression estimate. Second-stage estimates from the first difference instrumental variables (FDIV) estimator are shown in column 3 of Table 5. Even though the null hypothesis of exogeneity of the health insurance premiums is not rejected at the 5% level (Hausman test's p-value=0.11), the second-stage FDIV estimate is relevantly more negative than the corresponding FD estimate. The latter result may be explained by either measurement errors in premiums or reverse causation or both endogeneity problems. The second-stage coefficient for the health insurance premiums is statistically significant at the 5% significance level, and indicates that a 10% increase in premiums leads to a 2.2% decrease (1.1-percentage-points decrease) of the employment rate. We can use our estimates to study the economy-wide impact of the growth of health insurance premiums in the period 2005–2010. Since premiums increased by 11.1% in real terms between 2005 and 2010 (see Table 2), rising premiums decreased the employment rate by 2.4% (1.2-percentage-points) in the period 2005–2010, *ceteris paribus*.

We also report and discuss the second-stage FDIV estimate in the case that we use the full set of instruments for the health insurance premiums (see Section 5.2 for the list of instrumental variables for the health insurance premiums). The first-stage FDIV estimates are reported in column 2 of Table 4 while the second-stage FDIV estimate is reported in column 4 of Table 5. The second-stage coefficient on the health insurance premiums using the full set of instruments is approximately identical to the coefficient reported in column 3 of Table 5, in which case we used per capita medical malpractice payments as the only excluded instrument. For this reason we do not describe the empirical results from FDIV estimation which uses all available instruments for the health insurance premiums, and we refer the reader to the discussion reported above in which case we exclusively used per capita medical malpractice payments as the unique excluded instrumental variable. Lastly, it is worth mentioning that the over-identifying restrictions are not rejected at the 5% level (Hansen test's p-value=0.58), indicating that the instruments are very likely to be exogenous.

---

<sup>15</sup> E.g., see Acemoglu et al. (2013).

Table 5: The Effects of the Health Insurance Premiums on Employment – OLS, FD and Second-stage FDIV Estimates Using Two Sets of Instrumental Variables

|  | (1)<br>OLS       | (2)<br>FD            | (3)<br>FDIV<br>(second-stage) | (4)<br>FDIV<br>(second-stage) |
|--|------------------|----------------------|-------------------------------|-------------------------------|
| Dependent variable: employment rate                |                  |                      |                               |                               |
| HI Premium (in 1000 \$)                            | 0.049<br>(0.042) | -0.006***<br>(0.002) | -0.028**<br>(0.012)           | -0.029**<br>(0.012)           |
| Observations                                       | 15,366           | 15,366               | 15,366                        | 15,366                        |
| Adjusted R-squared                                 | 0.170            | 0.113                | -                             | -                             |
| Year fixed effects?                                | YES              | YES                  | YES                           | YES                           |
| County fixed effects? <sup>†</sup>                 | NO               | YES                  | YES                           | YES                           |
| Controls? <sup>††</sup>                            | ALL              | ALL                  | ALL                           | ALL                           |
| Additional tests for FDIV estimations              |                  |                      |                               |                               |
| F-statistic on excluded instrument(s) <sup>‡</sup> | -                | -                    | 7.966                         | 21.485                        |
| Hausman test's p-value <sup>§</sup>                | -                | -                    | 0.113                         | 0.063                         |
| Hansen <i>J</i> test's p-value <sup>§</sup>        | -                | -                    | -                             | 0.580                         |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. The outcome variable is the employment rate in Columns 1-4. Year FE (fixed effects) are included in the regression model. The second-stage (FDIV) estimates are computed using the Stata<sup>TM</sup> command “ivreg2, cluster(state identifier)” using the model in first differences.

<sup>‡</sup>The F-statistic on the excluded instrument(s) is the Kleibergen-Paap rank Wald F statistic. The Kleibergen-Paap rank Wald F statistic takes into account the usage of cluster-robust standard errors.

<sup>†</sup>County FE (fixed effects) are eliminated from the regression models using first-difference transformation of the regression model, except in the case of Column 1 (pooled OLS estimation).

<sup>††</sup>Control variables include eight variables for race and Hispanic origin, share of males, three variables for the population distribution by age, population density, share of small firms.

<sup>§</sup>The Hausman test and the Hansen *J* test are used to test the exogeneity of the health insurance premiums and the validity of the over-identifying restrictions, respectively. Under the null hypothesis of the Hausman test, the premiums are exogenous. Under the null hypothesis of the Hansen *J* test, the over-identifying restrictions are valid.

We also investigated the effect of the health insurance premiums on **annual wages**. We report the OLS estimate for the effect of the health insurance premiums on annual wages in column 1 of Table 6. Based on this estimate, we note that premiums are positively correlated with wages at the 5% significance level: Counties with highly productive workers earning high wages are likely to be characterized by high health care spending and high health insurance premiums (Baicker and Chandra (2006), p. 616). Controlling for time-invariant unobserved heterogeneity via the FD estimator reverses the sign of the coefficient on the health insurance

premiums, and makes the coefficient statistically insignificant (see column 2 of Table 6). In unreported regressions, we have found evidence that time-invariant county characteristics are correlated with annual wages and the explanatory variables,<sup>14</sup> making the OLS estimate unreliable.

Table 6: The Effects of the Health Insurance Premiums on Annual Wages – OLS, FD and Second-stage FDIV Estimates Using Two Sets of Instrumental Variables

|  | (1)<br>OLS                | (2)<br>FD             | (3)<br>FDIV<br>(second-stage) | (4)<br>FDIV<br>(second-stage) |
|--|---------------------------|-----------------------|-------------------------------|-------------------------------|
| Dependent variable: annual wages                   |                           |                       |                               |                               |
| HI Premium (in 1000 \$)                            | 4,424.718***<br>(919.221) | -165.911<br>(111.927) | -250.269<br>(727.216)         | -353.025<br>(711.050)         |
| Observations                                       | 15,366                    | 15,366                | 15,366                        | 15,366                        |
| Adjusted R-squared                                 | 0.429                     | 0.068                 | -                             | -                             |
| Year fixed effects?                                | YES                       | YES                   | YES                           | YES                           |
| County fixed effects? <sup>†</sup>                 | NO                        | YES                   | YES                           | YES                           |
| Controls? <sup>††</sup>                            | ALL                       | ALL                   | ALL                           | ALL                           |
| Additional tests for FDIV estimations              |                           |                       |                               |                               |
| F-statistic on excluded instrument(s) <sup>‡</sup> | -                         | -                     | 7.966                         | 21.485                        |
| Hausman test's p-value <sup>§</sup>                | -                         | -                     | 0.906                         | 0.760                         |
| Hansen <i>J</i> test's p-value <sup>§</sup>        | -                         | -                     | -                             | 0.489                         |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. The outcome variable is annual wages in Columns 1–4. Year FE (fixed effects) are included in the regression model. The second-stage (FDIV) estimates are computed using the Stata<sup>TM</sup> command “ivreg2, cluster(state identifier)” using the model in first differences.

<sup>‡</sup>The F-statistic on the excluded instrument(s) is the Kleibergen-Paap rank Wald F statistic. The Kleibergen-Paap rank Wald F statistic takes into account the usage of cluster-robust standard errors.

<sup>†</sup>County FE (fixed effects) are eliminated from the regression models using first-difference transformation of the regression model, except in the case of Column 1 (pooled OLS estimation).

<sup>††</sup>Control variables include eight variables for race and Hispanic origin, share of males, three variables for the population distribution by age, population density, share of small firms.

<sup>§</sup>The Hausman test and the Hansen *J* test are used to test the exogeneity of the health insurance premiums and the validity of the over-identifying restrictions, respectively. Under the null hypothesis of the Hausman test, the premiums are exogenous. Under the null hypothesis of the Hansen *J* test, the over-identifying restrictions are valid.

Second-stage FDIV estimates for the effect of the health insurance premiums on wages are reported in column 3 and column 4 of Table 6 while the corresponding first-stage FDIV estimates are reported in column 1 and column 2 of Table 4, respectively. In the case of column 3 of Table 6, we exclusively use per capita medical malpractice payments to instrument for

the health insurance premiums. In the case of column 4 of Table 6, we use the full set of instrumental variables for the health insurance premiums. The second-stage coefficients on the health insurance premiums using instrumental variables methods are more negative than the corresponding FD coefficient reported in column 2 of Table 6. The difference between the FD and FDIV estimates may be explained by measurement errors in the health insurance premiums or by reverse causality bias. The latter endogeneity problem may arise for the following reason: Increases in (local) wages are likely to increase health care spending (see, e.g., Acemoglu et al. (2013)) and the health insurance premiums. As shown in columns 3-4 of Table 6, we find that the second-stage coefficient on the health insurance premiums is statistically insignificant at the 5% significance level. This result is at least partially caused by the *large standard error associated to the coefficient on the health insurance premiums*. Finally, in the case that we use the full set of instruments (see column 4), we find that the over-identifying restrictions are not rejected at the 5% significance level (Hansen test's p-value=0.489). This indicates that the instrumental variables are likely to be exogenous.

The empirical results for the effects of the health insurance premiums on the employment rate and wages are consistent with the predictions of a model in which workers do not fully value the health insurance benefit provided by the employer (see Section 2.1). Variation in medical malpractice payments and variation in medical malpractice legislation within states and over time do not increase the quantity or quality of the health insurance benefit provided by the employer, but changes the price of the health care provided by the employer (see Section 5.2). Consequently, changes in the health insurance premiums, which are driven by variation in medical malpractice payments or variation in medical malpractice legislation, will not be fully valued by workers.

## **6.2 The Causal Effect of the Health Insurance Premiums on Private Health Insurance Coverage**

Since employers are not obliged to offer health insurance coverage to their employees in the period 2005-2010, we also investigate whether increases in the health insurance premiums imply a decrease of private health insurance coverage. As shown in column 1 of Table 7, the OLS estimate suggests that premiums are uncorrelated to the private health insurance coverage. However, the OLS estimate reported in column 1 of Table 7 is unreliable, since we have found evidence that time-invariant county characteristics are correlated with the explanatory variables and private health insurance coverage.<sup>14</sup>

In the case that we additionally control for time-invariant county characteristics (see column 2 of Table 7), we find that increases in the cost of health insurance imply a statistically insignificant decrease of private health insurance coverage. The previous result may be explained based on the following line of reasoning. Employer contributions and employee contributions for employer-sponsored health insurance are not subjected to taxation in the US (see Gruber



(2011)). On the other hand, costs for nongroup health insurance are subjected to taxation in the US. Consequently, employees and employers may agree on a compensation package which includes employer-sponsored health insurance coverage, even when premiums are rising.

Table 7: The Effects of the Health Insurance Premiums on Private Health Insurance Coverage – OLS, FD and Second-stage FDIV Estimates Using Two Sets of Instrumental Variables

|  | (1)<br>OLS  | (2)<br>FD           | (3)<br>FDIV (second-<br>stage)<br>(second-stage) | (4)<br>FDIV (second-<br>stage)<br>(second-stage) |
|--|---|---------------------|--|--|
|  | Dependent variable: private health insurance coverage |                     |  |  |
| HI Premium (in 1000 \$)                            | 0.0202<br>(0.0364)                                    | -0.0026<br>(0.0052) | -0.0030<br>(0.0217)                              | -0.0048<br>(0.0201)                              |
| Observations                                       | 15,366  | 15,366              | 15,366   | 15,366   |
| Adjusted R-squared                                 | 0.437   | 0.049               | -  | -  |
| Year fixed effects?                                | YES   | YES                 | YES  | YES  |
| County fixed effects? <sup>†</sup>                 | NO  | YES                 | YES  | YES  |
| Controls? <sup>††</sup>                            | ALL   | ALL                 | ALL  | ALL  |
|  | Additional tests for FDIV estimations                 |                     |  |  |
| F-statistic on excluded instrument(s) <sup>‡</sup> | -   | -                   | 7.966  | 21.485   |
| Hausman test's p-value <sup>§</sup>                | -   | -                   | 0.987  | 0.686  |
| Hansen <i>J</i> test's p-value <sup>§</sup>        | -   | -                   | -  | 0.660  |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. The outcome variable is private (nonelderly) health insurance coverage. Year FE (fixed effects) are included in the regression model. The second-stage (FDIV) estimates are computed using the Stata<sup>TM</sup> command “ivreg2, cluster(state identifier)” using the model in first differences.

<sup>‡</sup>The F-statistic on the excluded instrument(s) is the Kleibergen-Paap rank Wald F statistic. The Kleibergen-Paap rank Wald F statistic takes into account the usage of cluster-robust standard errors.

<sup>†</sup>County FE (fixed effects) are eliminated from the regression models using first-difference transformation of the regression model, except in the case of Column 1 (pooled OLS estimation).

<sup>††</sup>Control variables include eight variables for race and Hispanic origin, share of males, three variables for the population distribution by age, population density, share of small firms.

<sup>§</sup>The Hausman test and the Hansen *J* test are used to test the exogeneity of the health insurance premiums and the validity of the over-identifying restrictions, respectively. Under the null hypothesis of the Hausman test, the premiums are exogenous. Under the null hypothesis of the Hansen *J* test, the over-identifying restrictions are valid.

We also analyzed the effect of the health insurance premiums on private health insurance coverage using the first-difference instrumental variables (FDIV) estimator. The first-stage estimate for the effect of “per capita medical malpractice payments” on the health insurance premiums is reported in column 1 of Table 4, and the first-stage estimates using the full set of IVs are

reported in column 2 of Table 4. Second-stage estimates are reported in columns 3-4 of Table 7. Using either per capita medical malpractice payments as the only instrumental variable for the health insurance premiums (see column 3 of Table 7), or using the full set of instrumental variables for the health insurance premiums (see column 4 of Table 7), we find that the second-stage coefficient on the health insurance premiums is negative and statistically insignificant. We also find that the second-stage coefficient on premiums is more negative than the corresponding FD coefficient reported in column 2 of Table 7. The latter finding might be explained by measurement errors in the health insurance premiums, which lead to a bias towards zero in the estimated coefficient on premiums. Finally, it is likely that the instrumental variables are exogenous, since the over-identifying restrictions are not rejected at the 5% significance level (Hansen test's p-value=0.660).

### 6.3 Robustness Checks

In this section we explore the robustness of our identification strategy. We analyze the relationship between the predicted health insurance premiums and variables, which should not be affected by our instruments. More in detail, we estimate eq. (4) using instrumental variables for the health insurance premiums, and using variables that should not be affected by our instruments (e.g., population density) as the dependent variable in eq. (4). The predicted premiums on the right-hand side of eq. (4) capture the variation in the instrumental variables that is used in the IV estimation; lastly, the predicted premiums are obtained by regressing the premiums on either per capita medical malpractice payments alone (see column 1 of Table 4) or by regressing the premiums on the full set of instrumental variables (see column 2 of Table 4).

Columns 1-8 of Table 8 demonstrate that the instrument “per capita medical malpractice payments” is unable to predict variation in the share of males, variables for age distribution (i.e., share of 15-19 year old persons, share of 20-24 year old persons, share of 65+ year old persons), population density, share of non-Hispanic Whites, share of non-Hispanic Blacks, share of small firms.<sup>16</sup> Since the above-mentioned variables are not statistically significantly affected by per capita medical malpractice payments, and since these variables could potentially affect the dependent variables used in this study (e.g., employment rate), it is likely that per capita medical malpractice payments affects the dependent variables only through the health insurance premiums.

Using the full set of instrumental variables (rather than using exclusively medical malpractice payments) to instrument for the health insurance premiums, we repeated the same tests

---

<sup>16</sup> In unreported regressions, the left-hand side variable for eq. (4) was represented by one of the following variables: share of Hispanic Whites, share of Hispanic Blacks, share of Hispanic Asians, share of non-Hispanic Asians, share of Hispanic persons belonging to another race, share of non-Hispanic persons belonging to another race. The instrument “per capita medical malpractice payments” was unable to predict variation in any of these dependent variables. These regression results are available from the author.

described above. The second-stage estimates are reported in Table 9. Columns 1–8 of Table 9 show that the instrumental variables are unable to predict variation in variables that could potentially affect the dependent variables used in analysis. Therefore, we think that the instruments reported in Section 5.2 affect the dependent variables only through the health insurance premiums.

Table 8: Specification Checks – Second-stage FDIV estimates (Medical Malpractice Payments as the Unique Instrumental Variable)

|                                    | (1)<br>FDIV         | (2)<br>FDIV         | (3)<br>FDIV       | (4)<br>FDIV      | (5)<br>FDIV         | (6)<br>FDIV          | (7)<br>FDIV          | (8)<br>FDIV        |
|------------------------------------|---------------------|---------------------|-------------------|------------------|---------------------|----------------------|----------------------|--------------------|
|                                    | Dependent variables |                     |                   |                  |                     |                      |                      |                    |
|                                    | Male                | Age (15–19)         | Age (20–24)       | Age (65+)        | Population density  | White (Non-Hispanic) | Black (Non-Hispanic) | Small firms        |
| HI Premium (in 1,000 \$)           | -0.002<br>(0.002)   | -0.00007<br>(0.001) | -0.002<br>(0.002) | 0.002<br>(0.002) | -14.664<br>(17.615) | -0.0003<br>(0.001)   | 0.004<br>(0.003)     | -0.0002<br>(0.003) |
| Observations                       | 15,366              | 15,366              | 15,366            | 15,366           | 15,366              | 15,366               | 15,366               | 15,366             |
| Year fixed effects?                | YES                 | YES                 | YES               | YES              | YES                 | YES                  | YES                  | YES                |
| County fixed effects? <sup>†</sup> | YES                 | YES                 | YES               | YES              | YES                 | YES                  | YES                  | YES                |
| Controls? <sup>††</sup>            | ALL                 | ALL                 | ALL               | ALL              | ALL                 | ALL                  | ALL                  | ALL                |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. Year FE (fixed effects) are included in the regression model. The second-stage (FDIV) estimates are computed using the Stata<sup>TM</sup> command “ivreg2, cluster(state identifier)” using the model in first differences.

<sup>†</sup>County effects are eliminated from the regression models using first-difference transformation of the regression model.

<sup>††</sup>Control variables include eight variables for race and Hispanic origin, share of males, three variables for the population distribution by age, population density, share of small firms, *excluding the dependent variable*.

Table 9: Specification Checks – Second-stage FDIV estimates (Full Set of Instrumental Variables)

|                                    | (1)<br>FDIV         | (2)<br>FDIV        | (3)<br>FDIV       | (4)<br>FDIV      | (5)<br>FDIV         | (6)<br>FDIV          | (7)<br>FDIV          | (8)<br>FDIV        |
|------------------------------------|---------------------|--------------------|-------------------|------------------|---------------------|----------------------|----------------------|--------------------|
|                                    | Dependent variables |                    |                   |                  |                     |                      |                      |                    |
|                                    | Male                | Age (15–19)        | Age (20–24)       | Age (65+)        | Population density  | White (Non-Hispanic) | Black (Non-Hispanic) | Small firms        |
| HI Premium (in 1,000 \$)           | -0.002<br>(0.001)   | -0.0003<br>(0.001) | -0.002<br>(0.002) | 0.002<br>(0.002) | -13.475<br>(16.154) | -0.0002<br>(0.001)   | 0.003<br>(0.003)     | -0.0003<br>(0.003) |
| Observations                       | 15,366              | 15,366             | 15,366            | 15,366           | 15,366              | 15,366               | 15,366               | 15,366             |
| Year fixed effects?                | YES                 | YES                | YES               | YES              | YES                 | YES                  | YES                  | YES                |
| County fixed effects? <sup>†</sup> | YES                 | YES                | YES               | YES              | YES                 | YES                  | YES                  | YES                |
| Controls? <sup>††</sup>            | ALL                 | ALL                | ALL               | ALL              | ALL                 | ALL                  | ALL                  | ALL                |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. Year FE (fixed effects) are included in the regression model. The second-stage (FDIV) estimates are computed using the Stata<sup>TM</sup> command “ivreg2, cluster(state identifier)” using the model in first differences.

<sup>†</sup>County effects are eliminated from the regression models using first-difference transformation of the regression model.

<sup>††</sup>Control variables include eight variables for race and Hispanic origin, share of males, three variables for the population distribution by age, population density, share of small firms, *excluding the dependent variable*.

## 7 The Heterogeneous Effects of the Health Insurance Premiums on Employment and Wages

In this section we analyze whether the relationship between the health insurance premiums and the labor market outcomes depends on the private health insurance coverage. That is, we ana-

lyze whether the private health insurance coverage is a *moderator* for the relationship between the health insurance premiums and the labor market outcomes used in analysis. We implement this analysis based on the following reasoning. Since we hypothesized that increases in the health insurance premiums may negatively affect the labor market outcomes via the increased employer's expenditure on the health insurance benefit (see Section 2.1), increases in the cost of health insurance may imply stronger negative effects on the labor market outcomes in counties with high private health insurance coverage (compared to counties with low private health insurance coverage).

In order to analyze whether the private health insurance coverage acts as a moderator for the relationship between the health insurance premiums and the labor market outcomes, we interact the health insurance premiums with (county-level) private health insurance coverage, and we use the following panel data regression model in first differences:

$$\Delta y_{ct} = \beta_1 \Delta \text{premiums}_{st} + \beta_2 \Delta \text{coverage}_{ct} + \beta_3 \Delta (\text{premiums}_{st} * \text{coverage}_{ct}) + \Delta \text{controls}_{ct}^T \beta_2 + \Delta T_t + \Delta \epsilon_{ct}, \quad (7)$$

where  $y_{ct}$  represents either the employment rate or annual wages; where  $\text{coverage}_{ct}$  represents the county-level private health insurance coverage; and where  $(\text{premiums}_{st} * \text{coverage}_{ct})$  represents the interaction between the health insurance premiums and private health insurance coverage. See Section 5.1 for the description of the remaining terms included in eq. (7). Below, we will report and discuss estimates from the FD (first difference) estimator for the regression model (7). As in previous sections, we use robust-cluster standard errors, and the standard errors are clustered at the state level.<sup>17</sup>

Since both the health insurance premiums and the private health insurance coverage as well as their interaction may be endogenous with respect to the error term on the labor market outcomes, we also use (static) GMM (generalized method of moments) estimation in order to estimate the parameters included in eq. (7). In the case of GMM estimation, we combine the external instruments reported in Section 5.2 and “internal instruments” in the same instrument set; the internal instruments are represented by the second, third, fourth and fifth lags of the three endogenous variables.<sup>18</sup> We use two-step GMM estimation and we use robust-cluster standard errors for GMM regressions, where the standard errors are clustered at the state level. The validity of the internal instruments relies on the assumption of serial uncorrelatedness of the error term in levels ( $\epsilon_{ct}$ ). In order to provide evidence that the aforementioned assumption

<sup>17</sup> For the sake of completeness, in the following regression tables we will also report estimates from the pooled OLS regression model (3) including premiums, private health insurance coverage and their interaction as the main explanatory variables. However, we will not discuss pooled OLS estimates, since we have found evidence that time-invariant county characteristics are correlated with the explanatory variables and the employment rate (or annual wages).

<sup>18</sup> We also estimated the regression model (7), using the three external instrumental variables reported in Section 5.2 and excluding the internal instruments from our instrument set -i.e., we used the FDIV estimator that was already presented in Section 6. The estimates from the (standard) FDIV estimator are not discussed below, since we have found that these estimates suffer from weak (joint) instrumentation.

is satisfied for our GMM regression models, in the following discussion we report and discuss results from Arellano-Bond tests for serial correlation of the error term (see Arellano and Bond (1991) for more details about these tests).<sup>19</sup>

### ***Empirical Estimates for the Employment Rate***

The results from FD and GMM estimations for the effect of premiums, private health insurance coverage and their interaction on the employment rate are reported in columns 2-3 of Table 10, respectively. It is worthwhile noting that the interaction term is statistically significant at the 5% significance level for the regressions reported in columns 2-3 of Table 10. This result implies that, regardless of the econometric specification, private health insurance coverage appears to be a moderator for the relationship between the health insurance premiums and the employment rate.

In the case that we control for county fixed effects via the first-difference transformation of the regression model (see column 2 of Table 10), we find that the coefficient on the health insurance premiums is statistically insignificant while the coefficient on the interaction term is significant at the 5% level. Since private health insurance coverage is always larger than zero for all counties, and since the coefficient on premiums is not statistically different from zero, the total effect of increases in the health insurance premiums on the employment rate is negative for all counties. Setting the coefficient on premiums equal to zero, we find that a 10% increase in premiums leads to a 0.7 percentage points and 0.9 percentage points decrease of the employment rate for the county at 25th percentile and 75th percentile of private health insurance coverage, respectively.<sup>20</sup>

Using GMM estimation, we find that the coefficient on the health insurance premiums is statistically insignificant and the coefficient on the interaction term is statistically significant at the 5% significance level (see column 3 of Table 10). Comparing the estimates from the FD and GMM estimations, we find evidence that the coefficients on premiums and the interaction term are downward biased for the FD estimator. This result is in line with the findings reported in Section 6.1. Setting the coefficient on premiums exactly equal to zero, we find that a 10% increase in premiums leads to a 1.8 percentage points and 2.4 percentage points decrease of the employment rate for the county at 25th percentile and 75th percentile of private health insurance coverage, respectively.<sup>21</sup> Since the health insurance premiums expressed in real terms

<sup>19</sup> Two-step GMM estimates are obtained using the “xtabond2, cluster(state identifier) nolevelq twostep” command in Stata 13™. The output of the “xtabond2” command also includes tests results for the Arellano-Bond tests for serial correlation of the error term.

<sup>20</sup> The 25th percentile of private health insurance coverage is equal to 0.545 while the 75th percentile of private health insurance coverage is equal to 0.706. Since the average of the health insurance premiums is 4.040, we obtain the following estimates:  $(0.404)*(-0.032)*(0.545)=-0.007$ ;  $(0.404)*(-0.032)*(0.706)=-0.009$ .

<sup>21</sup> The 25th percentile of private health insurance coverage is 0.545 while the 75th percentile of private health insurance coverage is 0.706. Since the average of the health insurance premiums is 4.040, we

increased by 11.1% in the period 2005-2010 (see Section 6.1), rising premiums decreased the employment rate by approximately 2 percentage points and 2.6 percentage points for the county at 25th percentile and 75th percentile of private health insurance coverage during 2005-2010, respectively.

The GMM estimation passes a series of specification tests. First, we find that the over-identifying restrictions are not rejected at the 5% level (Hansen test's p-value=0.375). Second, we find that the error term in levels ( $\epsilon_{ct}$ ) is uncorrelated with its first, second and third lags at the 5% significance level (see the results of the  $m_2$ ,  $m_3$  and  $m_4$  tests in Table 10, respectively),<sup>22</sup> indicating that internal instruments are very likely to be valid. To conclude, in line with our theoretical discussion, we find strong evidence that private health insurance coverage acts as a moderator for the relation between the health insurance premiums and the employment rate.

Table 10: The Moderating Effect of the Private Health Insurance Coverage on the Relationship between the Health Insurance Premiums and the Labor Market Conditions

|   | (1)<br>OLS                             | (2)<br>FD           | (3)<br>GMM          | (4)<br>OLS                          | (5)<br>FD                | (6)<br>GMM                 |
|---|--|---------------------|---------------------|-------------------------------------|--------------------------|----------------------------|
|   | Dependent variable:<br>employment rate |                     |                     | Dependent variable:<br>annual wages |                          |                            |
| HI premium (in 1000 \$)                   | 0.175**<br>(0.068)                     | 0.012*<br>(0.007)   | 0.022<br>(0.021)    | 4,358.716*<br>(2,460.614)           | -55.780<br>(224.669)     | 258.589<br>(505.232)       |
| Private HI coverage                       | 1.624***<br>(0.400)                    | 0.175***<br>(0.051) | 0.264**<br>(0.132)  | 20,193.930<br>(18,461.358)          | 1,730.920<br>(1,670.301) | 8,573.110**<br>(4,264.930) |
| Interaction (premium*coverage)            | -0.241***<br>(0.088)                   | -0.031**<br>(0.012) | -0.084**<br>(0.034) | -517.736<br>(4,387.644)             | -179.106<br>(414.453)    | -1,733.887**<br>(877.124)  |
| Observations                              | 15,366                                 | 15,366              | 15,366              | 15,366                              | 15,366                   | 15,366                     |
| R-squared                                 | 0.249                                  | 0.115               | -                   | 0.486                               | 0.068                    | -                          |
| Year fixed effects?                       | YES                                    | YES                 | YES                 | YES                                 | YES                      | YES                        |
| County fixed effects?                     | NO                                     | YES                 | YES                 | NO                                  | YES                      | YES                        |
| Controls?                                 | ALL                                    | ALL                 | ALL                 | ALL                                 | ALL                      | ALL                        |
| Additional tests for GMM estimations      |  |                     |                     |                                     |                          |                            |
| Hansen $J$ test's p-value <sup>§</sup>    | -                                      | -                   | 0.375               | -                                   | -                        | 0.437                      |
| Arellano-Bond $m_2$ p-value <sup>§§</sup> | -                                      | -                   | 0.306               | -                                   | -                        | 0.131                      |
| Arellano-Bond $m_3$ p-value <sup>§§</sup> | -                                      | -                   | 0.592               | -                                   | -                        | 0.593                      |
| Arellano-Bond $m_4$ p-value <sup>§§</sup> | -                                      | -                   | 0.182               | -                                   | -                        | 0.144                      |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. The outcome variable is the employment rate in Columns 1-3, and wages in Columns 4-6. Year FE (fixed effects) are included in the regression model. The second-stage GMM estimates are computed using the Stata<sup>TM</sup> command "xtabond2, cluster(state identifier) twostep nolevelq".

<sup>§</sup>The Hansen  $J$  test is used to test the validity of the over-identifying restrictions. Under the null hypothesis of the Hansen  $J$  test, the over-identifying restrictions are valid.

<sup>§§</sup>Under the null hypothesis of the Arellano-Bond's  $m_2$  test, the error term in levels ( $\epsilon_{ct}$ ) is uncorrelated with its lag ( $\epsilon_{c,t-1}$ ). Under the null hypothesis of the Arellano-Bond's  $m_3$  test, the error term in levels ( $\epsilon_{ct}$ ) is uncorrelated with its second-order lag ( $\epsilon_{c,t-2}$ ). Under the null hypothesis of the Arellano-Bond's  $m_4$  test, the error term in levels ( $\epsilon_{ct}$ ) is uncorrelated with its third-order lag ( $\epsilon_{c,t-3}$ ). See Arellano and Bond (1991) for more details about the  $m_2$ ,  $m_3$  and  $m_4$  tests.

obtain the following estimates:  $(0.404)*(-0.084)*(0.545) = -0.018$ ;  $(0.404)*(-0.084)*(0.706) = -0.024$ .

<sup>22</sup> For more details about the null hypothesis of the  $m_2$ ,  $m_3$  and  $m_4$  tests, see the notes at the bottom of Table 10.

***Empirical Estimates for Annual Wages***

The results from FD and GMM estimations for the effects of premiums, private health insurance coverage and their interaction on annual wages are reported in columns 5-6 of Table 10, respectively. In line with the theoretical discussion reported above, we find that the coefficient on the interaction term is always negative for the regressions reported in columns 5-6; however, the coefficient on the interaction term is statistically significant at the 5% significance level only in the case of the GMM estimation. Since the health insurance premiums and the interaction term are statistically insignificant in the case of the FD estimation (see column 5), we focus our discussion on the results from GMM estimation.

The GMM estimation passes two specification tests. We find that the over-identifying restrictions are not rejected at the 5% significance level (Hansen test's p-value=0.437), suggesting that both the external and internal instrumental variables are exogenous. We also find that the error term in levels ( $\epsilon_{ct}$ ) is uncorrelated with its first, second and third lags at the 5% significance level (see the results of the  $m_2$ ,  $m_3$  and  $m_4$  tests in Table 10). The latter result suggests that the internal instrumental variables are very likely to be exogenous.

The second-stage GMM estimate for the coefficient on the interaction term is negative and significant at the 5% significance level, indicating that the effect of increases in the health insurance premiums on wages varies depending on private health insurance coverage. On the other hand, we find that the coefficient on the health insurance premiums is statistically insignificant. Based on the estimates reported in column 6 and setting the coefficient on the health insurance premiums exactly equal to zero, we find that a 10% increase in premiums –i.e., a 404\$ increase in premiums– decreases annual wages by 382\$ and 495\$ for the county at 25th percentile and 75th percentile of private health insurance coverage, respectively.<sup>23</sup> These estimates imply that increases in premiums are “fully shifted” onto wages only in the case of counties with high private health insurance coverage while increases in premiums are less than fully shifted onto wages in the case of counties with low private health insurance coverage. Since the health insurance premiums increased by 11.1% in real terms during 2005-2010 (see Section 6.1), rising premiums decreased annual wages by slightly more than 500 US dollars for the county at the 75th percentile of private health insurance coverage.

Finally, it is worth noting that if we repeat the above calculations including the direct effect of the health insurance premiums on annual wages,<sup>24</sup> we find evidence that increases in premiums are less than fully shifted onto wages for both the county at the 25th percentile and 75th percentile of private health insurance coverage. In conclusion, our empirical results con-

<sup>23</sup> I.e., these estimates are obtained using the following formulas:  $(0.404)*(-1733.887)*(0.545)=-381.767$ ;  $(0.404)*(-1733.887)*(0.706)=-494.546$ .

<sup>24</sup> I.e., we calculate the effect of the health insurance premiums on wages as  $[(\beta_{premium} \times 0.404) + (\beta_{interaction} \times private\ coverage \times 0.404)]$ , where  $\beta_{premium}$  is the estimated coefficient on premiums;  $\beta_{interaction}$  is the estimated coefficient on the interaction term; (*private coverage*) denotes the private health insurance coverage. Finally, a 10% increase in premiums is represented by the value 0.404.

firm the conjecture that the private health insurance coverage *moderates* the effect of the health insurance premiums on annual wages.

## 8 Conclusion

This study analyzes the effects of increases in employer-sponsored health insurance premiums on employment and annual wages in the US for the period 2005-2010. Previous studies analyzed the effects of rising health insurance premiums (or rising health care costs) on several labor market outcomes using repeated cross-sectional (and individual-level) datasets (see Cutler and Madrian (1998); Baicker and Chandra (2006)). In contrast to previous studies, data on the labor market outcomes used in this study –employment rate and annual wages– vary at the county-year level and enable to control for time-invariant unobserved heterogeneity which may be correlated with the labor market outcomes and the health insurance premiums.

In addition to controlling for time-invariant unobserved heterogeneity, we also use instrumental variables for the health insurance premiums in order to estimate the causal effects of the health insurance premiums on employment and wages. Using variation in per capita medical malpractice payments and variation in medical malpractice legislation as the sources of identifying variation in the health insurance premiums, we find that a 10% increase in the health insurance premiums implies a statistically significant 2.2% decrease (1.1 percentage points decrease) of the employment rate, *ceteris paribus*. On the other hand, increases in the cost of employer-sponsored health insurance imply an insignificant decrease of annual wages. Since the health insurance premiums increased by 11.1% in real terms between 2005 and 2010, our results indicate that the increasing cost of health insurance had a strong negative effect on employment during 2005-2010.

Utilizing estimates of private health insurance coverage as a proxy for employer-sponsored health insurance coverage, we also analyzed whether private health insurance coverage acts as a *moderating variable* in the relationship between the health insurance premiums and the labor market outcomes used in analysis. Indeed, since increases in the cost of health insurance may negatively affect employment and wages via the increased cost of the health insurance benefit provided by the employer, rising premiums may imply stronger negative effects on employment and wages in counties that are characterized by high private coverage. We find that rising premiums imply a (statistically significantly) stronger negative impact on employment and wages in counties with high private coverage: The negative effect of rising premiums on the employment rate and annual wages is approximately 30% larger for the county at the 75th percentile of private health insurance coverage than for the county at the 25th percentile of private health insurance coverage.

Our study reaches two main conclusions. First, rising health insurance premiums imply a negative effect on workers in the form of decreased employment, and this effect is stronger for counties that display high private health insurance coverage. Second, rising premiums exert



an additional negative effect on workers in the form of decreased wages in counties that are characterized by high private health insurance coverage. Since rising health insurance premiums worsen employment prospects and tend to reduce wages, the US federal and state governments may consider to use fiscal stimulus in order to counteract the negative effects of rising premiums on employment and wages.

## References

- Acemoglu, D., Finkelstein, A. and Notowidigdo, M.J. (2013). Income and Health Spending: Evidence from Oil Price Shocks, *The Review of Economics and Statistics* 95(4), 1079–1095.
- Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *The Review of Economic Studies* 58(2), 277–297.
- Avraham, R., Dafny, L.S. and Schanzenbach, M.M. (2012). The Impact of Tort Reform on Employer-Sponsored Health Insurance Premiums, *The Journal of Law, Economics and Organization* 28(4), 657–686.
- Avraham, R. (2014). Database of State Tort Law Reforms (5th), University of Texas Law, Law and Economics Research Paper No. e555.
- Baicker, K. and Chandra, A. (2005). The effect of malpractice liability on the delivery of health care, *Forum for Health Economics & Policy* 8(1), Article 4.
- Baicker, K. and Chandra, A. (2006). The Labor Market Effects of Rising Health Insurance Premiums, *Journal of Labor Economics* 24(3), 609–634.
- Blavin, F., Holahan, J., Kenney, G., Chen, V. (2012). *A decade of coverage losses: implications for the Affordable Care Act*, Urban Institute, Washington (DC), 2012, February.
- Burtless, G. and Mikusheva, A. (2012). Effects of Employer Health Costs on the Trend and Distribution of Social Security-Taxable Wages, Center for Retirement Research at Boston College, WP (working paper) 2012-11.
- Catlin, A.C., and Cowan., C.A. (2015). History of Health Spending in the United States, 1960–2013, Baltimore: Centers for Medicare & Medicaid Services.
- Claxton, G., Gabel, J., DiJulio, B., Pickreign, J., Whitmore, H., Finder, B., Jacobs, P. and Hawkin, S. (2007). Health Benefits In 2007: Premium Increases Fall To An Eight-Year Low, While Offer Rates And Enrollment Remain Stable, *Health Affairs* 26(5), 1407–1416.
- Claxton, G., Gabel, J., DiJulio, B., Pickreign, J., Whitmore, H., Finder, B., Jarlenski, M. and Hawkins, S. (2008). *Employer Health Benefits: 2008*, The Kaiser Family Foundation, Menlo Park, California, US.
- Cutler, D. and Madrian, B. (1998). Labor market responses to rising health insurance costs: Evidence on hours worked, *RAND Journal of Economics* 29(3), 509–530.
- DeNavas-Walt, C., Proctor, B.D. and Smith, J.C. (2008). *Income, Poverty, and Health Insurance*

- Coverage in the United States: 2007*, U.S. Government Printing Office, Washington, DC, 2008.
- Gruber, J. and Krueger, A.B. (1991). The Incidence of Mandated Employer-Provided Insurance: Lessons from Workers' Compensation Insurance, *Tax Policy and the Economy* 5, 111–143.
- Gruber, J. (2011). The tax exclusion for employer-sponsored health insurance, *National Tax Journal* 64(2), 511–530.
- The Kaiser Family Foundation and HRET (2013). *Employer Health Benefits 2013 Annual Survey – Survey Design and Methods*, Kaiser Family Foundation.
- Kessler, D. and McClellan, M. (1996). Do Physician's Practice Defensive Medicine?, *Quarterly Journal of Economics* 111(2), 353–390.
- Kube, S., Marechal, M.A. and Puppe, C. (2013). Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment, *Journal of the European Economic Association* 11(4), 853–870.
- Lee, J. (2002). Are Health Insurance Premiums Higher for Small Firms?, Synthesis Project, Report No. 2, The Robert Wood Johnson Foundation, New Jersey, US.
- Retzer, K.D., Hill, R.D. and Pratt, S.G. (2013). Motor vehicle fatalities among oil and gas extraction workers, *Accident Analysis & Prevention* 51, 168–174.
- Seabury, S.A. (2009). Does Malpractice Liability Reform Attract High Risk Doctors?, RAND, 2009.
- Sheiner, L. (1999). Health costs, aging, and wages, working paper, U.S. Federal Reserve, Finance and Economics Discussion Series (FEDS), Discussion Paper No. 99-19.
- Sood, N., Ghosh, A. and Escarce, J.J. (2009). Employer-Sponsored Insurance, Health Care Cost Growth, and the Economic Performance of U.S. Industries, *Health Services Research*, 44(5), Part 1, 1449–1464.
- Stevens, W.S. (2003). *Health Insurance: Current Issues and Background*, Nova Publishers, 2003.
- Stock, J.H., Wright, J. and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business and Economic Statistics*, 20(4), 518–529.
- Western, B. and Rosenfeld, J. (2011). Unions, norms, and the rise in US wage inequality, *American Sociological Review*, 76(4), 513–537.
- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data, Second Edition*, Cambridge, MA: MIT Press.

## A Appendix A

### A.1 Structure of the Panel Dataset

Table A1: Panel structure

| Year      | Observations | Percentage | Cumulative percentage |
|-----------|--------------|------------|-----------------------|
| 2006      | 3079         | 20.04      | 20.04                 |
| 2007      | 3079         | 20.04      | 40.08                 |
| 2008      | 3080         | 20.04      | 60.12                 |
| 2009      | 3080         | 20.04      | 80.16                 |
| 2010      | 3048         | 19.84      | 100.00                |
| All years | 15366        | 100        | 100                   |

Notes: Since our methodology relies on first-difference estimation, the number of observations in the year 2005 is zero. See Section 5.1 for more details about the methodology used in this study.

### A.2 The Construction of County-Level Estimates for Private Health Insurance Coverage

County-level estimates for private nonelderly health insurance coverage are not available from governmental agencies for the period 2005–2010. Using county-level estimates on total nonelderly health insurance coverage, and county-level estimates on public nonelderly health insurance coverage, we computed estimates of private nonelderly health insurance coverage by subtracting the latter variable from the first variable reported above.

#### A.2.1 Total (Nonelderly) Health Insurance Coverage

County-level estimates for total (nonelderly) health insurance coverage are obtained from the US Census Bureau’s SAHIE (Small Area Health Insurance Estimates) dataset. The SAHIE dataset contains estimates for the share of nonelderly (0–64) individuals that are covered by either private or public health insurance in each US county.<sup>25</sup> Data on total health insurance coverage are available from 2005 onward. For this reason, the initial year included in our longitudinal dataset is the year 2005. Finally, see Appendix B.3 for additional details on data sources for total health insurance coverage.

<sup>25</sup> Private health insurance coverage can be obtained via the employer or via the nongroup health insurance market. Public health insurance coverage can be obtained through Medicaid or CHIP (Children’s Health Insurance Program).

### A.2.2 Public (Nonelderly) Health Insurance Coverage

We use the US Department of Health & Human Services' AHRF (Area Health Resource File) dataset to construct county-level estimates of public (nonelderly) health insurance coverage. The AHRF (Area Health Resource File) dataset contains county-level estimates for the *number of nonelderly individuals that are eligible for public health insurance* –i.e., Medicaid or CHIP (Children's Health Insurance Program). We divide the latter variable, which varies at the county level, by county-level estimates of the *number of nonelderly (0-64) persons living in each US county*, thus obtaining county-level estimates of public health insurance coverage. Data on Medicaid/CHIP eligibility from the AHRF dataset are not available after the year 2010. For this reason, the final year included in our longitudinal dataset is the year 2010. Finally, see Appendix B.3 for additional details on data sources for public health insurance coverage.

### A.3 Summary Statistics for Per-Capita Medical Malpractice Payments

Table A2: Summary Statistics for Per Capita Medical Malpractice Payments

|  | Obs.  | Mean   | Std. Dev. | Min   | Max    |
|--|-------|--------|-----------|-------|--------|
| Per-capita medical malpractice payments in US dollars <sup>†</sup> | 15366 | 10.048 | 6.358     | 2.714 | 45.548 |

Notes: data on medical malpractice payments are obtained from the U.S. Department of Health and Human Services' NPDB (National Practitioner Data Bank) dataset.

<sup>†</sup>Medical malpractice payments are expressed in per capita terms, and are expressed in constant 2005 US dollars using the Bureau of Labor Statistics' CPI inflation calculator.

### A.4 Variation in Medical Malpractice Legislation

Table A3: Legal Reform Used in Analysis

| Reform                             | Description of the reform   | Example   |
|------------------------------------|---|---|
| Caps on non-economic damage awards | Non-economic (pain and suffering) damage awards payable to the victim of malpractice are capped at a statutorily specified dollar amount. | A cap on non-economic damage awards was introduced in Georgia in 2005, limiting non-economic damage awards to 350,000 \$ per healthcare provider. |

Table A4: List of States that Enacted (or Repealed) Caps on Non-economic Damage Awards

|                   | 2005 | 2006       | 2007 | 2008 | 2009 | #Changes |
|-------------------|------|------------|------|------|------|----------|
| <b>Enactments</b> | GA   | AK, IL, SC | -    | -    | -    | 4        |
| <b>Repeals</b>    | -    | WI         | -    | -    | -    | 1        |

Notes: since health insurance contracts are typically finalized 3+ months prior to the calendar year in which they take effect (see Avraham et al. (2012), p. 13), and since we use data for the health insurance premiums from 2005 to 2010, the relevant years for identification are changes from 2005 to 2009. Data source: DSTLR (Database of State Tort Law Reforms), 5th edition, 1980-2012; see Avraham (2014) for more details.

## B Appendix B. Data Sources

### B.1 Data Sources and Data Imputation for the Employer-Sponsored Health Insurance Premiums

- Data sources for the employer-sponsored health insurance premiums

County-level data on health insurance premiums are not available from governmental agencies. We use state-year-level data on health insurance premiums, and the data is obtained from the MEPS-IC database. More in detail, the data on health insurance premiums used in this study are obtained from Table II.C.1 of the MEPS-IC database.

The MEPS-IC data can be downloaded at the following link: [https://meps.ahrq.gov/mepsweb/data\\_stats/quick\\_tables\\_results.jsp?component=2&subcomponent=2&year=2005&tableSeries=-1&tableSubSeries=&searchText=&searchMethod=1&Action=Search](https://meps.ahrq.gov/mepsweb/data_stats/quick_tables_results.jsp?component=2&subcomponent=2&year=2005&tableSeries=-1&tableSubSeries=&searchText=&searchMethod=1&Action=Search)

The methodology used for data collection is available at the following link: [https://meps.ahrq.gov/mepsweb/data\\_stats/Pub\\_ProdLookup\\_Results.jsp?ProductType=Methodology\%20Report\&Comp=Insurance](https://meps.ahrq.gov/mepsweb/data_stats/Pub_ProdLookup_Results.jsp?ProductType=Methodology\%20Report\&Comp=Insurance)

- Data imputation for the health insurance premiums

The MEPS-IC dataset does not include data on health insurance premiums for the year 2007. As in Burtless and Mikusheva (2012), we reconstructed the health insurance premiums in year 2007 as the average of the health insurance premiums in year 2006 and year 2008 by state.

To check whether this imputation technique is valid to approximate the health insurance premiums in 2007, we constructed a synthetic measure for the health insurance premiums (henceforth “synthetic premiums”) in year 2009 as the average of the health insurance premiums in year 2008 and year 2010 (by state), and we have found that the “synthetic premiums” in year 2009 are highly correlated to the real health insurance premiums in 2009 (correlation=0.85). Thus, it is also likely that the imputed premiums for the year 2007 are highly correlated to the real (and unknown) premiums for the year 2007.

## **B.2   Data Sources for the Labor Market Outcomes**

- Data sources for the employment rate

The employment rate at the county level is constructed as the number of employed persons by county (QCEW dataset) divided by the number of 15-64 year old persons by county (“USA counties” dataset).

County-level (annual) data on the number of employed persons by county are obtained from the Bureau of Labor Statistics’ QCEW (Quarterly Census of Employment and Wages) dataset. QCEW datasets can be downloaded at the following link: <https://www.bls.gov/cew/datatoc.htm>

Data on the number of 15-64 year old persons by county are obtained from the US Census Bureau’s “USA counties” dataset. The “USA counties” dataset can be downloaded at this link: <http://www.census.gov/support/USACdataDownloads.html>.<sup>26</sup>

- Data sources for annual wages

County-level data on annual wages are obtained from the Bureau of Labor Statistics’ QCEW (Quarterly Census of Employment and Wages) dataset. The variable used in analysis is “Annual average pay”: “Annual average pay” is equal to total annual wages divided by annual average employment (by county). County-level data can be downloaded at the following link: <https://www.bls.gov/cew/datatoc.htm>

## **B.3   Data Sources for Private (Nonelderly) Health Insurance Coverage**

- Total (Nonelderly) Health Insurance Coverage

SAHIE estimates for the period 2005-2010 can be downloaded at the following link: <https://www.census.gov/data/datasets/time-series/demo/sahie/estimates-cps.html>

- Public (Nonelderly) Health Insurance Coverage

The AHRF data set can be downloaded at <http://ahrf.hrsa.gov/download.htm>. Data on the number of nonelderly individuals in each county-year cell are obtained from the SAHIE dataset; see <https://www.census.gov/data/datasets/time-series/demo/sahie/estimates-acs.html>

---

<sup>26</sup> See the Excel file “Mastdata.xls” for more information on the county level variables that are included in the “USA counties” dataset; the aforementioned Excel file can be found at <http://www.census.gov/support/USACdataDownloads.html>

#### **B.4 Data Sources for State-Level Estimates of Employer-Sponsored Health Insurance Coverage**

The March CPS (Current Population Survey) datasets contain information on health insurance coverage of US citizens. The March CPS asks several questions about the types of health insurance coverage that the respondent had during the calendar year prior to interview. Among all questions on health insurance coverage included in the March CPS, we exclusively selected the question on employer-sponsored health insurance coverage. Each respondent is asked whether (or not) the respondent is covered by employer-sponsored health insurance. The dummy variable for employer-sponsored health insurance coverage takes value 1 if the respondent receives health insurance from her employer, 0 otherwise. In order to compute estimates for employer-sponsored health insurance coverage at the state-year level, we computed the average of the above-mentioned dummy variable at the state level, and we repeated this procedure for all March CPS datasets ranging from March CPS 2006 to March CPS 2011.<sup>27</sup>

The March CPS datasets are obtained from the CEPR Uniform Extracts of the March CPS. The datasets and codebooks can be found at the following link: <http://ceprdata.org/cps-uniform-data-extracts/march-cps-supplement/march-cps-data/>

#### **B.5 Data Sources for Control Variables Used in Analysis**

- Data sources for socio-economic variables

County-level data on socio-economic variables are obtained from the US Census Bureau's "USA counties" dataset. The "USA counties" dataset can be downloaded at <http://www.census.gov/support/USACdataDownloads.html>

- Data sources for "share of small firms"

The share of small firms is defined as the ratio of number of small firms to total number of firms, where a firm is defined as a "small firm" if it employs less than 500 employees. Data on firm size can be downloaded at <https://www.sba.gov/advocacy/firm-size-data> (see the "counties excel file" therein).

#### **B.6 Data Sources for the Instrumental Variables**

- Data sources for "per capita medical malpractice payments"

---

<sup>27</sup> Due to the retrospective nature of the March CPS, we use the March CPS 2006 to generate estimates of employer-sponsored health insurance coverage for the year 2005; the same line of reasoning applies to the remaining March CPS datasets.

Data on “per capita medical malpractice payments” are obtained from the US Department of Health & Human Services’ NPDB (National Practitioner Data Bank) dataset. Data on medical malpractice payments can be downloaded at the following link: <http://www.npdb.hrsa.gov/resources/publicData.jsp>

- Data sources for medical malpractice tort reforms

Data on tort legislation at the state-year level are obtained from the DSTLR (Database of State Tort Law Reforms), fifth edition. See the excel file “Database of State Tort Law Reforms Clever (5.1)” at the following link: <http://www.utexas.edu/law/faculty/ravraham/dstlr.html>

## C Appendix C. Summary Statistics for Total and Public (Nonelderly) Health Insurance Coverage

Table C1: Summary Statistics for Total and Public Health Insurance Coverage

|   | Observations | Mean  | Std. Dev. | Min   | Max   |
|---|--------------|-------|-----------|-------|-------|
| Total nonelderly health insurance coverage (share) <sup>†</sup>   | 15366        | 0.819 | 0.059     | 0.692 | 0.906 |
| Public nonelderly health insurance coverage (share) <sup>††</sup> | 15366        | 0.197 | 0.094     | 0.053 | 0.388 |

<sup>†</sup>Total (nonelderly) health insurance coverage is the share of persons that are covered by health insurance in each county.

<sup>††</sup>Public (nonelderly) health insurance coverage is the share of persons that are covered by public health insurance in each county. Public health insurance coverage can be obtained through Medicaid or CHIP.

## D Appendix D. Malpractice Legislation and the Health Insurance Premiums

Table D1: Variation in Medical Malpractice Legislation and the Health Insurance Premiums

|  | FD estimator             |
|--|--------------------------|
|  | Dependent variable       |
|  | HI Premium (in 1,000 \$) |
| <i>caps on damages</i> <sub><i>s,t-1</i></sub> (enactments of caps on damage awards) | -0.003<br>(0.050)        |
| <i>repeals</i> <sub><i>s,t-1</i></sub> (repeals of caps on damage awards)            | 0.110***<br>(0.014)      |
| Observations   | 15,366                   |
| Adjusted R-squared   | 0.376                    |
| Year fixed effects?  | YES                      |
| County fixed effects?  | YES                      |
| Controls?  | ALL                      |

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Clustered (and robust) standard errors at the state level are reported in parenthesis. The outcome variable is health insurance premiums. Year FE (fixed effects) are included in the regression model. The regression model is estimated using the Stata<sup>TM</sup> command “reg, cluster(state identifier)” using the model in first differences.