The Demand for Insurance and Rationale for a Mandate: Evidence from Workers' Compensation Insurance

ONLINE APPENDIX

Marika Cabral, Can Cui, and Michael Dworsky

A Description of Data Sources

Below, we provide more detail on the data sources used in this paper and variable construction based on these underlying data sources. There are three main sources of administrative data we obtained from the Texas Department of Insurance (TDI).

1. **Covered Employers.** Data on covered employers is obtained from the "Proof-of-Coverage Data," released by TDI in response to an open records request (TDI (2014)). This database includes information on covered employers including: employer identifiers (e.g., FEIN), policy effective dates, and employer governing classification. To construct our analytical dataset, we use employer identifiers (based on employer FEIN) to define unique employers.[1] We aggregate data across employers using information on employer governing classifications as described in the text. We also obtain supplemental data from regulatory documents on certified self-insured firms, which were also released through an open records request (TDI (2016*a*)). These data are reported monthly, and we use information about the certified self-insured status and number of covered employees reported in January of each calendar year.

2. **Covered Payroll and Claims.** Data on covered payroll and claims is obtained from the "Unit Statistical Data" (TDI (2019)). These data were released by TDI in response to an open records request. These actuarial data cover every workers' compensation insurance policy sold in Texas, including information on: industry-occupation classification code, coverage dates, covered payroll, premiums, experience rating modifiers, and data on each associated claim. For each workers' compensation claim, the data include information on: the date of claim, type of claim (e.g., major indemnity, minor indemnity, medical only), classification of injured employee, incurred medical benefits, and incurred indemnity benefits. Losses are valued at pre-specified intervals since the policy effective date, and our baseline cost measure draws on costs valued at 42 months after the policy effective date. The mean claim cost measure used in the selection analysis draws on underlying claim cost data that are winsorized at the 99$^{\text{th}}$ percentile and that exclude losses flagged as due to aggregate catastrophic events (e.g., natural disasters). See Section 1 for more description on the construction of our cost measure. To construct our analytical dataset, we aggregate payroll and claims data to the classification-time level as described in the text.

3. **Base Rates and Supplemental Data.** Data on classification base rates—commonly known as relativities—was obtained from the "Workers' Compensation Relativities Studies" files posted on the TDI website (TDI (2016*b*)). The data we extract from these files include: classification code, classification base rates (the final adopted rates and intermediate rates used in earlier steps within the update algorithm), and precise effective dates. We also obtain loss development factor data from the Workers' Compensation Relativities Studies files (TDI (2011)). Finally, we obtain data on insurer combined loss ratios from TDI publications (TDI WCREG (2016), TDI WCREG (2010)).

[1] In instances with multiple observations for the same employer (as defined by employer FEIN), we assign the employer the governing classification and NAICS industry code of the observation representing the largest share of premiums.

B Description of Base Rate Update Algorithm

Below, we describe the algorithm used by the Texas Department of Insurance (TDI) to update base rates. The data associated with the base rate update algorithm (e.g., inputs, outputs, intermediate outputs) come from the Workers' Compensation Relativities Studies files. We are thankful to employees of the TDI Actuarial Office for several helpful discussions as we worked to understand the details of the rate update process. We first outline the steps for updating base rates in a typical year with a revenue neutral update, and we then explain how this update algorithm is adjusted in years in which the overall level of base rates is adjusted (i.e., "re-basing years").

Step 1: The initial inputs into the algorithm are: (i) the raw loss experience for relevant policy years, which is a five-year window lagged by four years and (ii) the current base rates. For example, for base rates in 2007, the raw loss experience considered is the loss experience from policy years 1999 to 2003. Below, we will represent the year the update will take effect as t, and consider the window used as input as $[t-8, t-4]$. Indemnity losses were grouped into categories depending on the injury type. These categories are serious (i.e., death, permanent total, and major permanent partial) and non-serious (i.e., minor permanent partial and temporary total). Medical losses are similarly grouped into serious, non-serious, and medical only categories.

Step 2: Raw losses were adjusted to exclude all amounts in excess of per-claim or per-accident caps (e.g., \$350,000 per claim, \$700,000 per accident). These adjusted amounts are referred to as limited losses. The purpose of limiting the losses is to reduce the possibility of large random fluctuations that might otherwise occur from the occurrence or non-occurrence of a single large accident.

Step 3: The limited losses are adjusted for loss development and scaled so that the mean equals the mean of the current base rates.

Step 4: The adjusted limited losses summed across all the input policy years for each classification-category ($\text{AggLimitedLoss}_{jc}$) are used to determine a set of experience relative base rates. These experience relative base rates are then credibility weighted against the current relative base rates. The experience relative base rate, $expRel_{jc}$, for classification j and category c is defined as follows,

$$expRel_{jc} = \frac{(\text{AggLimitedLoss})_{jc} \times 100}{\text{AggPayroll}_{j}}. \tag{1}$$

These experience relative base rates are then weighted depending on whether a specified number of claims threshold is met using the following weights:

$$Cred_{jc} = \begin{cases} 1, & \text{if full credibility number of claims threshold met} \\ \left(\frac{(\text{AggPayroll}_{j} \times crtRel_{jc})/100}{\text{full credibility losses}_{c}}\right)^{0.4}, & \text{otherwise} \end{cases}$$

where the current relative base rate for classification j category c ($crtRel_{jc}$), the full credibility number of claims threshold, and full credibility losses are in TDI Documentation (Workers' Compensation Relativities Studies, Exhibits 21 and 22). Lastly, the *weighted relative base rate*, $wgtRel$, is defined as follows:

$$wgtRel_{jc} = Cred_{jc} expRel_{jc} + (1 - Cred_{jc}) crtRel_{jc}. \tag{2}$$

The final step works with the overall base rates, which is simply the sum across categories c. We denote overall base rates by dropping the c subscript.

Step 5: Next, the *balanced indicated relative base rate*, $balRel$, is calculated as follows:

$$balRel_{j} = \left(\frac{\Sigma_j crtRel_j \times \text{payroll in t-4}_j}{\Sigma_j wgtRel_j \times \text{payroll in t-4}_j}\right) wgtRel_j. \tag{3}$$

Lastly, the relative rates are capped so that the change is at most a 25% change in either direction to create the *limited relative base rate*, $limRel$:

$$limRel_j = \begin{cases} 1.25 \times crtRel_j, & \text{if } balRel_j > 1.25 \times crtRel_j \\ 0.75 \times crtRel_j, & \text{if } balRel_j < 0.75 \times crtRel_j \\ balRel_j, & \text{otherwise.} \end{cases}$$

In these terms, the *proposed relative base rate*, $proRel_j$, is:

$$proRel_j = \begin{cases} limRel_j, & \text{if } balRel_j > 1.25 \times crtRel_j \text{ or } balRel_j < 0.75 \times crtRel_j \\ (\frac{\Sigma_j crtRel_j \times \text{payroll in t-4}_j}{\Sigma_j limRel_j \times \text{payroll in t-4}_j}) limRel_j, & \text{otherwise.} \end{cases}$$

Note the above calculation yields a new set of relative base rates that are approximately revenue neutral.[2]

Step 6: Three updates during our analysis period (2008, 2009, and 2011) included across-the-board decreases in the level of base rates. These level decreases are made *after* all of the other steps described above. An X% drop in base rates is achieved by an adjustment of the following form:

$$\text{Final Base Rate}_j = (1 - X) proRel_j. \tag{4}$$

In a year with no across-the-board reduction, the *final base rate* is simply the proposed base rate ($X = 0$).

C Welfare Analysis: More Details on Empirical Implementation

C.1 Approach

The approach to empirically implementing the welfare analysis follows Einav and Finkelstein (2011), adapting the framework to accommodate the risk-adjusted premiums observed in this setting. Throughout the discussion below, the risk adjustment we refer to is employer-level experience rating. To ease notation, let us represent risk-adjusted payroll units as: $\mathbb{Q}$. Specifically, we use the variation in classification base rates to estimate reduced form elasticities in terms of risk-adjusted payroll for demand ($\epsilon_{\mathbb{Q},b} \equiv \frac{\partial \mathbb{Q}}{\partial b} \cdot \frac{b}{\mathbb{Q}}$) and average cost ($\epsilon_{AC,b} \equiv \frac{\partial AC(\mathbb{Q})}{\partial b} \cdot \frac{b}{AC(\mathbb{Q})}$). We can combine these elasticities to get the elasticity of the average cost curve with respect to risk-adjusted payroll:

$$\frac{\partial AC(\mathbb{Q})}{\partial \mathbb{Q}} \cdot \frac{\mathbb{Q}}{AC(\mathbb{Q})} = \frac{\frac{\partial AC(\mathbb{Q})}{\partial b} \cdot \frac{b}{AC(\mathbb{Q})}}{\frac{\partial \mathbb{Q}}{\partial b} \cdot \frac{b}{\mathbb{Q}}}. \tag{5}$$

Suppose that marginal costs are monotonic in $\mathbb{Q}$. Then, the sign of the above elasticity in equation 5 offers a test for selection: $\frac{\partial AC(\mathbb{Q})}{\partial \mathbb{Q}} > 0$ indicates advantageous selection, and $\frac{\partial AC(\mathbb{Q})}{\partial \mathbb{Q}} < 0$ indicates adverse selection.

To go beyond a test for selection in the quantitative welfare analysis, we need to make parametric assumptions on the form of the demand and cost curves. We proceed by making such assumptions and combining the reduced form elasticities with market-level data reported by the Texas Department of Insurance (TDI) on mean premiums, mean quantities, and mean insurer combined loss ratios to trace out the empirically relevant curves in this setting (analogous to those presented in the graphical illustration in Figure 5). Consider two different parametric forms for the demand and cost curves as a function of $\mathbb{Q}$: linear and constant elasticity.

We take as inputs our two elasticity estimates ($\epsilon_{\mathbb{Q},b} \equiv \frac{\partial \mathbb{Q}}{\partial b} \cdot \frac{b}{\mathbb{Q}}$; $\epsilon_{AC,b} \equiv \frac{\partial AC(\mathbb{Q})}{\partial b} \cdot \frac{b}{AC(\mathbb{Q})}$) and market-level aggregates from TDI on mean premium per risk-adjusted unit (p^*), mean cost per risk-adjusted unit (c^*),[3] and mean risk-adjusted quantity ($\mathbb{Q}^*$).

[2]In practice, there are two reasons why these rates may depart from revenue neutral updates slightly. First, in some years there seem to be some slight deviations from the above Step 5 description due to a rounding error. Second, Step 5 described above produces relative base rates that are close to (but not perfectly) revenue neutral. This is because the "capped" classifications are not re-normalized in the final stage. In practice, this does not make a difference because it is so close to revenue neutral.

[3]The mean costs are inferred from the reported mean insurer combined loss ratios and mean premiums.

- **Linear**: We use the reduced form estimates along with the aggregate TDI data and a linear parametric extrapolation to back out the parameters in the demand and average cost curves: $D(p) = A + Bp$; $AC(p) = C + Ep$. We can derive the MC curve from these curves using:

$$MC(p) = (\frac{\partial D}{\partial p})^{-1} \frac{\partial (AC(p) \times D(p))}{\partial p}. \tag{6}$$

  Using this relationship, we get that:

$$MC(p) = \frac{AE}{B} + C + 2Ep. \tag{7}$$

  We can re-write these in terms of $\mathbb{Q}$,

  - $P(\mathbb{Q}) = \frac{\mathbb{Q}}{B} - \frac{A}{B}$
  - $AC(\mathbb{Q}) = C - \frac{AE}{B} + \frac{\mathbb{Q}E}{B}$
  - $MC(\mathbb{Q}) = C - \frac{AE}{B} + \frac{2E\mathbb{Q}}{B}$.

  We can back out these parameters with our reduced form elasticity estimates and the available aggregates: $A \equiv \mathbb{Q}^*(1 - \epsilon_{\mathbb{Q},b})$; $B \equiv \epsilon_{\mathbb{Q},b}(\frac{\mathbb{Q}^*}{p^*})$; $C \equiv c^*(1 - \epsilon_{AC,b})$; $E \equiv \epsilon_{AC,b}(\frac{c^*}{p^*})$.

- **Constant Elasticity**: We use the reduced form estimates along with the aggregate TDI data and a constant elasticity parametric extrapolation to back out the parameters in the demand and average cost curves: (i) $AC(p) = Ap^{e_c}$ and (ii) $D(p) = Bp^{e_d}$. We can derive the MC curve from these curves using:

$$MC(p) = (\frac{\partial D}{\partial p})^{-1} \frac{\partial (AC(p) \times D(p))}{\partial p}. \tag{8}$$

  Using this relationship, we get that:

$$MC(p) = \frac{e_c + e_d}{e_d} AC(p). \tag{9}$$

  So, we can write $MC(p) = Cp^{e_c}$, where $C \equiv A\frac{e_c + e_d}{e_d}$. In terms of $\mathbb{Q}$ we can express the inverse demand and cost curves as:

  - $P(\mathbb{Q}) = (\frac{\mathbb{Q}}{B})^{\frac{1}{e_d}}$
  - $AC(\mathbb{Q}) = A(\frac{\mathbb{Q}}{B})^{\frac{e_c}{e_d}}$
  - $MC(\mathbb{Q}) = A\frac{e_c + e_d}{e_d}(\frac{\mathbb{Q}}{B})^{\frac{e_c}{e_d}}$.

  We can back out these parameters with our reduced form elasticity estimates and the available aggregates: $e_c \equiv \epsilon_{AC,b}$; $e_d \equiv \epsilon_{\mathbb{Q},b}$; $A \equiv \frac{c^*}{(p^*)^{\epsilon_{AC,b}}}$; $B \equiv \frac{\mathbb{Q}^*}{(p^*)^{\epsilon_{\mathbb{Q},b}}}$.

C.2 Definition of Data Elements

While Section 1.2 describes our data sources, this section elaborates on the available data and the definition of several variables of interest in our analysis. The administrative data focus on information about employers and payroll covered by workers' compensation insurance. To conduct the welfare analysis described in the text, we additionally need to measure the size of the market: the total eligible payroll that could be covered by the workers' compensation system. Following the methodology used by TDI for internal research on participation rates (Choi, 2011), we measure the size of the market through comparing the administrative covered payroll data to private sector covered payroll data from the Quarterly Census of Employment and Wages (QCEW) (U.S. Bureau of Labor Statistics (2014)). Because the administrative data on covered payroll exclude certified self-insured employers, we adjust the denominator of private sector payroll to exclude payroll represented by certified self-insured employers during our analysis period. Because there is no covered payroll information for the certified self-insured employers, we approximate covered payroll at these firms by combining administrative data on the number of covered employees at these firms with data on mean earnings in private sector employment from the QCEW.

Recall that premiums in this market are represented as in equation 1 described in Section 1.1 of the main text. We have data on several components of these premiums. We use data on regulatory base rates ($b_t(c_j)$) in our primary estimation. In addition, we use data on premiums before experience rating is applied and experience rating modifiers. The welfare analysis measures quantity in units of risk-adjusted (experience-rated) payroll. To measure the fraction insured, we need an estimate of the total eligible risk-adjusted payroll that could be insured in the market. In practice, we estimate the total eligible risk-adjusted payroll by calculating how the mean experience rating modifier varies with covered payroll, and we use this function—in combination with market-wide data on the total eligible payroll in Texas—to estimate the total eligible risk-adjusted payroll that could potentially be insured. To estimate how the experience rating modifier varies with the covered payroll, we estimate reduced form regressions relating: (i) the mean risk adjustment modifier to base rates and (ii) the mean covered payroll to base rates. We then use a linear extrapolation from these estimated elasticities to predict the average experience rating modifier if all eligible payroll were insured in the market, and scale the total eligible payroll in Texas by this prediction to obtain the total eligible risk-adjusted payroll.

D Additional Robustness Analysis

D.1 Workers' Compensation Classification Coding

The identification strategy outlined in the main text takes workers' compensation classification coding of employers as exogenous. In this appendix section, we investigate the possibility of problematic endogenous coding related to our identifying variation. Let j represent an employer and t represent year. Specifically, we estimate specifications such as the following:

$$I(c_{j,t} \neq c_{j,t-1}) = \beta \Delta ln(b)_{c_{j,t-1}} + \tau_t + \gamma_j + \alpha_{c_{j,t-1}} + \epsilon_{jt}, \tag{10}$$

where $c_{j,t}$ represents the classification of employer j in year t, and $\Delta ln(b)_k$ is defined as the difference in log base rate for classification k between year t and $t-1$ $\left(\Delta ln(b)_k \equiv ln(b_{k_t}) - ln(b_{k_{t-1}})\right)$. As noted above, additional controls include year fixed effects (τ_t), employer fixed effects (γ_j), and fixed effects for the classification in year $t-1$ ($\alpha_{c_{j,t-1}}$). Robust standard errors are clustered by classification in year $t-1$.

As noted in the text, in practice employers may have multiple classifications if they have a diverse workforce. In the employer-level data we use, we observe the employer's primary classification, often referred to as the *governing classification*, which covers most of the employer's payroll. Actual premiums paid are adjusted to account for the fraction of the employer's workforce dedicated to other categories (most commonly clerical and sales services), and the percent of payroll allocated to each classification is subject to verification with ex post payroll auditing. In the analysis here, we focus on whether there is endogenous coding of an employer's governing classification (i.e., an employer's primary classification). We note that any observed changes in the governing classification of an employer could represent true underlying changes in the workforce composition of an employer.

With the inclusion of employer fixed effects, the coefficient β in equation 10 measures the degree to which employer classification switching is correlated with regulatory base rate increases associated with an employer's classification. Specifically, a positive and significant coefficient estimate for β would indicate that employers are more likely to switch away from a particular classification when the relative price increases for this classification. Appendix Table A2 presents the results. There are a few important things to note. First, changes in employer governing classifications are uncommon. Among the classification-year observations in this data, 91% represent employers who have the same classification in this year as in the prior year. Second, there is no detectable association between the base rate variation and classification switching. Appendix Table A2 displays the estimates from equation 10 with the controls listed above (in column 1) and with additional controls (in column 2); both specifications yield estimates for β that are small and statistically indistinguishable from zero.

D.2 Impact of Governing Classification Base Rates on Overall Mean Base Rates

While our data does not allow us to investigate the prevalence of firms with multiple classifications, we conduct some conservative back-of-the-envelope calculations to assess the potential importance of this data limitation on the demand estimation. This analysis suggests that this data limitation has limited potential impact on the demand estimation.

According to the Texas Department of Insurance (TDI) actuarial office, it is common for large firms with multiple classifications to have 80-90% of payroll attributable to their governing classification, with adjustments for the remaining 10-20% of payroll attributable to other classifications, most commonly clerical and sales classifications. Because premium adjustments for secondary classifications are concentrated in clerical and sales classifications—classifications that are low risk with low base rates—these adjustments typically account for a small share of premiums for employers with multiple classifications. For instance, the most common clerical classification (classification 8810) has a classification base rate that is 0.17 times the mean classification base rate in the baseline year 2006.

We conduct conservative back-of-the-envelope analysis to understand how adjustments for secondary classifications may affect our estimates. Specifically, we analyze the impact of a firm's governing classification base rate on the total premiums paid by the firm for this coverage. Though most firms have a single classification, suppose we conservatively assume that all firms have 20% of payroll attributable to another classification—clerical services (classification 8810). We can estimate the impact of the employer's governing classification base rate on the associated employer's overall mean base rate (and premiums) by estimating the following specification:

$$ln(y_{j,t}) = \theta_0 + \theta_1 ln(b_{j,t}) + \rho_j + \gamma_t + u_{jt}, \tag{11}$$

where the overall mean base rate for firms with governing classification j in year t is represented by $y_{j,t} \equiv 0.8b_{j,t} + 0.2b_{clerical,t}$. We note that this regression accounts for any correlation in rate updates across the governing classifications and the secondary clerical classification and accounts for heterogeneity across classifications in the relative magnitude of these adjustments compared to the governing classification base rate. We also estimate a specification that replaces the overall mean base rate with the overall mean premiums per unit of risk-adjusted payroll: $y_{j,t} \equiv 0.8p_{j,t} + 0.2p_{clerical,t}$. The results are displayed in Appendix Table A3. Based on these estimates, we see that a 1% increase in the governing classification base rate leads to a 0.975% increase in the overall mean base rate and a 0.970% increase in premiums paid per unit of risk-adjusted payroll. This analysis illustrates that percent changes in governing classification base rates would translate nearly one-for-one in percent changes to employer overall mean base rates and premiums, even if adjustments for secondary classifications were more prevalent than indicated by TDI.

D.3 Exclusion of Certified Self-Insured Employers

Our baseline analysis excludes certified self-insured employers and associated employee payroll. We make this exclusion for two key reasons: (i) our identification strategy leverages variation in the premiums for coverage purchased from workers' compensation insurance providers, and (ii) the administrative data on covered payroll and claims are only available for the payroll covered through policies purchased from a workers' compensation insurance provider. As discussed in the text, there are strict requirements to become a certified self-insured firm. Perhaps because of these requirements, very few employers take up this option: only 95 firms are ever self-insured during our analysis period (2006-2011). Among these 95 firms that are ever self-insured from 2006-2011, 89 firms are continuously self-insured for the entire time period. In other words, there are only a handful of firms who ever switch between being self-insured and another status (purchased policy or no insurance). While the persistence in self-insurance implies it is unlikely that the exclusion of these firms affects our demand estimates, we directly analyze the robustness of the results with respect to our baseline sample definition, as described below.

We have administrative data on the identity of each certified self-insured firm in addition to each employer with a purchased policy. Thus, we can repeat the analysis analyzing the number of participating employers, either excluding or not excluding the certified self-insured firms. The baseline analysis reported in Table 3 columns 1 through 4 in the main text excludes certified self-insured firms, and Appendix Table A4 displays the analysis including all covered employers within the proof-of-coverage data (with no restriction to exclude the certified self-insured). Comparing these results, we see the results are very similar.

D.4 Eligible Population of Firms and Workers

Our baseline analysis uses dependent variables (the natural logarithm of covered employers, the natural logarithm of covered payroll) that are constructed solely from the administrative data. As discussed in Section

2, there is no administrative data on the universe of eligible firms and workers in each classification, so it is not possible to estimate demand in terms of the fraction of payroll insured (or the fraction of firms insured). A more detailed explanation is below. The ideal demand estimation would be in terms of the share of eligible firms or eligible payroll that is covered:

$$ln(\frac{\text{TotInsured}_{jt}}{\text{TotEligible}_{jt}}) = \gamma + \pi ln(b_{jt}) + \lambda_j + \tau_t + \mu_{jt}. \tag{12}$$

Rearranging terms we get:

$$ln(\text{TotInsured}_{jt}) = \gamma + \pi ln(b_{jt}) + \lambda_j + \tau_t - ln(\text{TotEligible}_{jt}) + \mu_{jt}, \tag{13}$$

where $ln(\text{TotEligible}_{jt})$ is unobserved. Suppose we can represent this term as:

$$ln(\text{TotEligible}_{jt}) = \phi + \rho ln(b_{jt}) + \eta_j + \sigma_t + e_{jt}. \tag{14}$$

Substituting this into the ideal demand specification we get:

$$ln(\text{TotInsured}_{jt}) = (\gamma + \phi) + (\pi + \rho) ln(b_{jt}) + (\eta_j + \lambda_j) + (\sigma_t + \tau_t) + (e_{jt} + \mu_{jt}). \tag{15}$$

Thus, the feasible regression will provide an estimate of $\pi + \rho$. This is a consistent estimate of the true demand elasticity π if and only if $\rho = 0$. Thus, to interpret the baseline estimates as reflecting the demand for insurance, a key assumption is that the eligible population of workers and firms in each classification is not changing in response to the identifying premium variation (i.e., $\rho = 0$). While the lack of classification-level data on the eligible population prevents us from testing this directly, we present some supporting evidence for this assumption by using North American Industry Classification System (NAICS) industry-year level data on the Texas workforce from the Quarterly Census of Employment and Wages (QCEW) and relating this to the classification-year-level variation in workers' compensation premiums using a crosswalk derived from the administrative data.

Specifically, we take aggregate data on the universe of firms and workers at the NAICS industry-year-level from the QCEW. We then match these to the classification-year-level workers' compensation premium variation using a crosswalk that is derived from the administrative data. We construct this crosswalk using the administrative proof-of-coverage data on employers participating in the Texas workers' compensation system. Importantly, these data include the workers' compensation governing classification code for each employer and these data also include information on the NAICS six-digit industry code.

In practice, there are a few challenges to creating a crosswalk from industry codes to classification codes. First, the NAICS industry code field is missing for approximately one-fifth of observations. Second, each NAICS code does not always map nicely to one workers' compensation classification code. In the face of these challenges, we proceed as follows. Starting with the pooled data across our analysis period, we use the observed NAICS industry-classification pairs to construct a frequency-weighted crosswalk under the assumption that the missing industry values are not selected. To remove outliers that may represent measurement error, we exclude industry-classification pairs that represent fewer than 10 observations or fewer than 5% of the observations associated with a particular NAICS industry code. In this analysis, we restrict attention to industries with mean annual employment exceeding 1,000 workers over the analysis period.

We examine whether the eligible population is related to the identifying variation by estimating variants of the following equation:

$$ln(y_{it}) = \alpha + \beta ln(b_{it}) + \delta_i + \theta_t + \lambda_i t + \epsilon_{it}, \tag{16}$$

where i is a NAICS industry, and t is a year. In this specification, $ln(b_{it})$ represents the natural logarithm of the mean base rate applicable in the industry based on the constructed NAICS-classification weighted crosswalk described above. All specifications include year and industry fixed effects, and we also estimate specifications with an additional control: a three-digit NAICS industry-specific time trend.

Appendix Table A6 presents the results. Overall, the results suggest that neither the aggregate number of firms nor the aggregate number of workers in an industry is responsive to the premium variation in classifications associated with the industry. This evidence builds confidence in our interpretation of the primary baseline regressions as reflecting the demand for insurance.

D.5 Demand Analysis: Additional Robustness

In addition to the alternative specifications discussed in the main text, we further probe the robustness of the demand estimates with respect to a few additional potential concerns.

Incidence of Premium Changes It is unclear how the burden of increased premiums (or the benefit from reduced premiums) is shared among employers and employees. To the extent that employers shift the cost of workers' compensation premiums onto workers, wages may be partially shifted upward or downward to reflect changes in workers' compensation premiums. Ideally, the demand estimation would use a pure quantity measure that is not sensitive to possibly endogenous wage adjustments. While we analyze covered employers which is a pure quantity measure, we also analyze covered payroll (wages multiplied by hours) which only represents a pure quantity measure if wages are not responsive to the identifying variation in workers' compensation premiums.[4] To evaluate the sensitivity of our estimates to potential endogenous wage adjustment, we repeat the covered payroll regression analysis under various assumptions on the fraction of premiums passed through to employees in the form of reduced wages. Specifically, these additional specifications repeat the baseline payroll regression replacing the dependent variable with the natural logarithm of normalized covered payroll: $ln(\frac{\text{payroll}_{jt}}{1-\theta\times\text{premium}_{jt}})$, where premium_{jt} represents the mean premium per dollar of payroll for classification j in year t, and θ represents the fraction of premiums shifted to employees in the form of reduced wages.

Appendix Table A8 Panel A displays the results of these additional specifications. These estimates illustrate that regardless of the division of premiums between employers and employees on the margin, increases in classification base rates lead to a decline in covered payroll. Across the range of possible assumptions on the division of premiums between employees and employers, a 1% increase in classification base rates leads to an estimated decline in normalized covered payroll of 0.22% to 0.29%.[5] For the purpose of our discussion of mandates in Section 4, we use demand estimates where quantity is measured using unadjusted covered payroll.

Alternative Samples We investigate the stability of our estimates when estimating alternative specifications in which we restrict attention to a subset of classifications. Appendix Table A8 Panel B displays estimates from a specification focusing on larger classifications (excluding classifications with annual insured payroll of less than $10 million) and estimates from a specification that excludes clerical and sales classifications (classifications that are the most common secondary classifications). The estimates in these alternative specifications are very similar to the baseline estimates.

Alternative Weighting While the descriptive statistics throughout represent market-level aggregates weighted by the payroll insured within each classification, the baseline regressions estimating the causal effect of rates on coverage are unweighted. There are two key reasons for this. First, it is not clear whether unweighted or weighted regressions are preferred when estimating causal effects.[6] Second, data is not available to construct

[4] Analyzing data from compulsory workers' compensation insurance systems, Gruber and Krueger (1991) find that workers' compensation premium changes in the 1980s in some high-risk industries were largely shifted into wages. As these authors discuss, their findings are consistent with multiple explanations, including that labor supply is more inelastic than labor demand (a typical finding in tax incidence analyses of labor markets) or that employees value workers' compensation coverage changes that were coincident with the premium changes they analyze. Because the present empirical setting is quite different from the setting these authors investigate (for example, in the present empirical setting coverage is optional, all occupational groups are included, etc.), it is not clear whether employers or employees bear the incidence of workers' compensation insurance premium updates. While our baseline approach is to analyze unadjusted covered payroll, the key results are not sensitive to which segment of consumers bears the incidence of workers' compensation insurance premiums.

[5] It is not surprising that the results are robust across the different possible divisions of premium updates across employers and employees. To see this, note that the average premium is $1.81 per $100 in payroll; thus, a 1% across-the-board increase in premiums would lead to approximately a 0.0181% decrease in covered payroll if coverage rates were held fixed and premium changes were fully shifted onto employees in the form of reduced wages. In other words, any mechanical effect of premiums on wages is expected to be an order of magnitude smaller than the estimated demand elasticity, regardless of the incidence of workers' compensation insurance premiums.

[6] As discussed in Solon, Haider and Wooldridge (2015), weighted regressions do not recover the average partial treatment effect in the presence of unmodeled treatment effect heterogeneity; if there is no heterogeneity in partial effects, both weighted and unweighted regressions provide consistent estimates of the homogeneous partial effect. Heteroskedasticity may be another motivation to weight regressions, though weighting may either ameliorate or exacerbate heteroskedasticity concerns depending on the degree to which outcomes are correlated within clusters. Solon, Haider and Wooldridge (2015) recommend comparing weighted and unweighted estimates to assess model mis-specification and recommend that researchers report heteroskedasticity robust standard errors. In line with these recommendations, we report heteroskedasticity robust standard errors throughout and assess robustness to weighting as described

the most natural weights for the demand estimation in this setting: total eligible payroll or total eligible employers within each classification. Though weighting by eligible payroll (or eligible employers) is infeasible, we present supplemental analysis which suggests that the results of this infeasible analysis would likely be similar to the estimates in the unweighted analysis. Appendix Table A8 Panel C illustrates that we obtain similar estimates in alternative specifications, where we weight the regressions using feasible proxies for eligible payroll within each classification: covered payroll, risk-adjusted covered payroll, and premiums paid in the first year of the analysis period.

D.6 Selection Analysis: Additional Robustness

In addition to the analysis in the main text, we further explore the robustness of the selection analysis. Appendix Table A9 presents estimates from additional specifications with alternative transformations of the cost measure and alternative cost measures. The results reported in Appendix Table A9 illustrate that these additional specifications yield similar findings. Appendix Figure A8 depicts a binned mean residual plot to graphically illustrate the baseline selection estimates. While the cost regression estimates are nosier than the demand estimates, these plots show no evidence of selection. Finally, Appendix Figure A9 displays estimates from the complementary event study approach outlined in equation 6, with mean claim costs (Panel A) and risk-adjusted payroll (Panel B) as outcomes. This figure shows no evidence of selection and no evidence of pre-existing trends.

D.7 Welfare Analysis and Empirical Cost Curves

Based on the empirical analysis which finds no evidence of selection in this market, the primary welfare calculations in the text are conducted under the assumption of no selection, meaning that there is a flat market-level average/marginal (risk-adjusted) cost curve. Appendix Table A11 presents alternative welfare calculations employing the small (and statistically indistinguishable from zero) risk-adjusted cost elasticity estimates reported in Table 4. The key patterns in these welfare estimates are similar to those in the baseline welfare analysis in Table 5.

E Interpretation of Demand

While the decision to purchase workers' compensation insurance is made by employers, the welfare analysis relies on the assumption that employer decisions reflect both employer and employee values for workers' compensation insurance. In Section 4 of the text, we describe one simple model that provides sufficient conditions for demand to reflect both employer and worker values of workers' compensation insurance. Below, we present a more detailed description of this model, and we discuss a simple alternative model that yields the same result.

E.1 Detailed Description of Model

Below, we apply the model of equalizing differentials outlined in Rosen (1986) to the setting of a labor market where firms choose wages and whether to purchase workers' compensation insurance over the outside option of settling workplace injuries through the tort system. The intuition behind this model is simple. Labor markets tie together two transactions: workers sell their labor services to firms and buy a set of job attributes from firms, while firms buy labor services from workers and sell a set of job attributes to workers. In this way, the labor market induces sorting of workers across firms, and job attributes reflect worker preferences and firm costs.

Consider a competitive labor market where workers have homogeneous productivity, and there are two types of jobs. Let I index the job type, where $I = 1$ in jobs with workers' compensation insurance and $I = 0$ in jobs without workers' compensation insurance. Let w_1 and w_0 represent the wages earned in the associated job type, and let the wage differential be represented by $\Delta w \equiv w_0 - w_1$. Both workers and firms make rational, privately optimal decisions.

Worker Preferences Worker i's preferences are represented by utility function U^i, which is a function of consumption (C), and workers' compensation insurance at his/her job (I), where $I = 1$ if insured and 0 otherwise. Worker utility is increasing in consumption ($U^i_C > 0$) and workers may place positive or negative value on workers' compensation insurance relative to the outside option of legal recourse through the tort system ($U^i_I \lessgtr 0$).

above.

Let C_i^1 denote market consumption for individual i when $I = 1$. Let C_i^* denote the consumption level that would give the same utility to the worker in a job without workers' compensation insurance as would have been attained in a job with workers' compensation at consumption level C_i^1: $U^i(C_i^*, 0) = U^i(C_i^1, 1)$. If the worker places a positive (negative) value on workers' compensation insurance then $C_i^* \geq C_i^1$ ($C_i^* \leq C_i^1$).

Define $\beta_i = C_i^* - C_i^1$, which represents worker i's value of a job with workers' compensation insurance relative to a job with the outside option of tort liability (i.e., the compensating differential for $I = 0$ compared to $I = 1$). Worker i chooses to apply to a job with workers' compensation insurance if and only if $\beta_i \geq \Delta w$.

Market Supply Holding total employment fixed, the labor supply for each type of job is simply the fraction of workers applying to jobs in each market segment: jobs with and without workers' compensation insurance.[7] Let G represent the distribution of β in the worker population. Let L_I^s be the fraction of workers applying to jobs of type I. In this notation, we can represent labor supply in each market segment as:

$$\begin{aligned} L_1^s(\Delta w) &= P(\beta \geq \Delta w) = 1 - G(\Delta w) \\ L_0^s(\Delta w) &= P(\beta < \Delta w) = G(\Delta w). \end{aligned} \tag{17}$$

Firm Production and Costs Firms choose which job type ($I = 0$ or $I = 1$) offered to the market. Suppose firms' production scales linearly with the number of workers L. Further, suppose the following describes profits for firm j:

$$\pi = \underbrace{\delta L}_{= \text{Total Production}} - \underbrace{\big((p + w_1)I + (\alpha_j + w_0)(1 - I)\big)L}_{= \text{Total Costs}}, \tag{18}$$

where δ represents per-worker productivity of labor, α_j represents the per-worker expected costs associated with tort liability, and p represents the per-worker price of workers' compensation insurance. The expected costs of tort liability are weakly positive ($\alpha_j \geq 0$), and this represents the per-worker cost savings (or value) to the firm from purchasing workers' compensation insurance. While firms face the same price for workers' compensation insurance, firms may differ in the cost savings they get from avoiding tort liability (α_j). For instance, there may be variation across firms in the transaction costs and legal fees associated with injury settlements in the outside option. Firm j will purchase workers' compensation insurance if and only if $p \leq \Delta w + \alpha_j$, or equivalently when $\alpha_j \geq p - \Delta w$.

Market Demand Suppose the number of firms and firm size are fixed. Let F represent the distribution of α across jobs offered by firms. This distribution incorporates the size of each firm as well as production technology, meaning $F(\alpha)$ indicates the fraction of potential jobs in the economy for which the expected per-worker costs from tort liability are less than or equal to α. Let L_I^d represent labor demand for job type I. This can be represented as:

$$\begin{aligned} L_1^d(p - \Delta w) &= P(\alpha \geq p - \Delta w) = 1 - F(p - \Delta w) \\ L_0^d(p - \Delta w) &= P(\alpha < p - \Delta w) = F(p - \Delta w). \end{aligned} \tag{19}$$

Market Equilibrium The market clears when labor supply equals labor demand in each segment of the market: $L_1^d(p - \Delta w) = L_1^s(\Delta w)$ (or equivalently, $L_0^d(p - \Delta w) = L_0^s(\Delta w)$). In the notation above, this equilibrium condition can be represented as:

$$1 - F(p - \Delta w) = 1 - G(\Delta w). \tag{20}$$

A direct consequence of this model is that there will be positive assortative matching of firms and workers, meaning workers with higher values for workers' compensation insurance sort toward firms with higher per-worker values from purchasing workers' compensation insurance (i.e., greater cost savings from avoiding tort liability). In equilibrium, the sum of the per-worker value of this coverage to the marginal firm and the value of this coverage to the marginal worker equals the price of workers' compensation insurance. Thus, the demand for workers' compensation insurance reflects the value of this coverage to marginal employers and employees. Given the equilibrium wage differential Δw^*, the marginal worker i' is indifferent between working at a job with and without workers' compensation insurance, $\beta_{i'} = \Delta w^*$. The marginal firm j' is

[7] Because this exercise holds fixed total employment, the labor supply in each market segment only depends on the wage differential Δw rather than the wage level in each segment (w_0 and w_1). A general equilibrium model would be required to determine the wage level in each market segment.

indifferent between purchasing workers' compensation insurance or not, $p = \alpha_{j'} + \Delta w^*$. Combining these expressions, we see that in equilibrium, the sum of the per-worker value of this coverage to the marginal firm and the value of this coverage to the marginal worker equals the price of workers' compensation insurance, $p = \alpha_{j'} + \beta_{i'}$. Hence, at any given price, the share of the workforce covered by workers' compensation insurance is the share of the workforce for which the sum of the per-worker employer value and the employee value exceeds the price.

E.2 Simple Alternative Model

We present one simple alternative model employing different assumptions that produces the same basic result: the demand for workers' compensation insurance reflects the value of this coverage to marginal employers and employees. Consider an employer's decision to allocate employee compensation across wage and non-wage job attributes, where we focus on one non-wage attribute: the provision of workers' compensation insurance. Suppose employer-employee matches and total employee compensation, c_i (the aggregate value of wage and non-wage compensation for employee i), are determined within the broader labor market and are taken as given by an employer. Let α_j represent employer j's per-worker expected costs of tort liability. The expected costs of tort liability are weakly positive ($\alpha_j \geq 0$), may vary across employers, and represent the cost savings (or value) to the firm from purchasing workers' compensation insurance. Let β_i represent employee i's value of workers' compensation insurance relative to the outside option of tort liability. There may be heterogeneity in employee values, and employee values may be either positive or negative, as workers' compensation insurance and recourse through the tort system are horizontally differentiated from a worker's perspective.

Let p represent the per-employee premium for workers' compensation insurance. Suppose employers know employees' values for workers' compensation insurance, and employers can flexibly adjust employee wages. If an employer elects to purchase workers' compensation insurance, this insurance must be provided to all employees. Let N_j represent the number of workers employed by employer j. Employer j will choose to purchase workers' compensation coverage if and only if it minimizes the total compensation costs to do so:

$$\sum_{i=1}^{N_j} (w_i^{WC} + p - \alpha_j) \leq \sum_{i=1}^{N_j} w_i^0, \tag{21}$$

where w_i^{WC} is employee i's wage compensation if offered workers' compensation insurance and $w_i^0 \equiv c_i$ is employee i's wage compensation if not offered workers' compensation insurance. Total compensation from employee i's perspective is held constant by setting $w_i^{WC} = w_i^0 - \beta_i$. Thus, employer j will offer workers' compensation insurance if and only if the per-capita benefits accruing to the employer and associated employees exceed the per-capita premiums paid, $p \leq \frac{1}{N_j} \sum_{i=1}^{N_j} (\alpha_j + \beta_i)$. At a given price, the share of the workforce covered by workers' compensation insurance reflects the share of the workforce for which the sum of the per-worker value of this coverage to firms and the mean value of this coverage to employees exceeds the price of coverage.

References

Choi, Soon-Yong. 2011. "Costs to Employers and Efficiencies In the Texas Workers' Compensation System." Texas Department of Insurance Workers' Compensation Research and Evaluation Group. `https://www.tdi.texas.gov/reports/wcreg/documents/Employers_cost.pdf`. (Date accessed: 12/2021).

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives*, 25(1): 115.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. 2021. "Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]." Minneapolis, MN: IPUMS. `https://doi.org/10.18128/D030.V9.0`.

Gruber, Jonathan, and Alan B Krueger. 1991. "The Incidence of Mandated Employer-Provided Insurance: Lessons from Workers' Compensation Insurance." In *Tax Policy and the Economy, Volume 5*. 111–144. The MIT Press.

Rosen, Sherwin. 1986. "Chapter 12 The theory of equalizing differences." In . Vol. 1 of *Handbook of Labor Economics*, 641 – 692. Elsevier.

Solon, Gary, Steven J. Haider, and Jeffrey Wooldridge. 2015. "What are we weighting for?" *Journal of Human Resources*, 50(2): 301–316.

Texas Department of Insurance (TDI). 2011. "Workers' Compensation Relativities Studies 2011." Exhibit 12-1. Texas Department of Insurance. https://www.tdi.texas.gov/reports/report9.html. (Date accessed: 9/2014).

Texas Department of Insurance (TDI). 2014. "Texas Workers' Compensation Proof-of-Coverage Electronic Data Interchange Data (2005-2014Q1)." Texas Department of Insurance.

Texas Department of Insurance (TDI). 2016*a*. "Certified Self-Insured Extracts (2006-2011)." Texas Department of Insurance.

Texas Department of Insurance (TDI). 2016*b*. "Workers' Compensation Relativities Studies." Studies available within period 2003 to 2015. Texas Department of Insurance. https://www.tdi.texas.gov/reports/report9.html. (Date accessed: 9/2016).

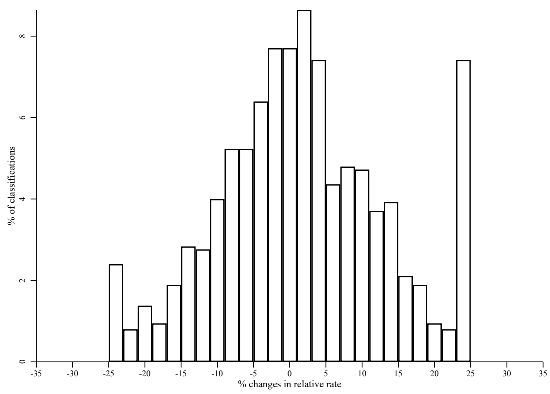
Texas Department of Insurance (TDI). 2019. "Unit Statistical Data (2004-2014)." Texas Department of Insurance.

Texas Department of Insurance Workers' Compensation Research and Evaluation Group (TDI WCREG). 2010. "Setting the Standard: An Analysis of the Impact of the 2005 Legislative Reforms on the Texas Workers' Compensation System, 2010 Results." Figure 2.1 pg. 8 (2006 estimates). Texas Department of Insurance. https://www.tdi.texas.gov/reports/wc/documents/wcsetstd2010.pdf. (Date accessed: 12/2021).

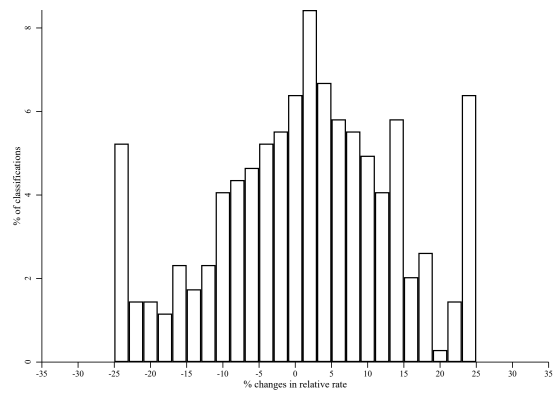
Texas Department of Insurance Workers' Compensation Research and Evaluation Group (TDI WCREG). 2016. "Setting the Standard: An Analysis of the Impact of the 2005 Legislative Reforms on the Texas Workers' Compensation System, 2016 Results." Figure 2.1 pg. 7 (2007-2011 estimates). Texas Department of Insurance. https://www.tdi.texas.gov/reports/wcreg/documents/biennial2016.pdf. (Date accessed: 12/2021).

U.S. Bureau of Labor Statistics. 2014. Quarterly Census of Employment and Wages (QCEW), Retrieved from https://www.bls.gov/cew/downloadable-data-files.htm, June 15, 2014.

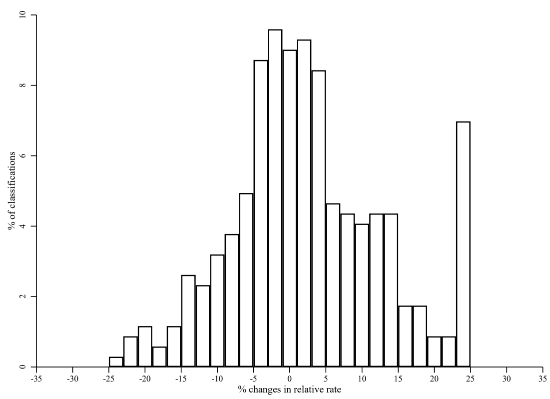
Figure A1: Histogram of Proposed Base Rate Updates



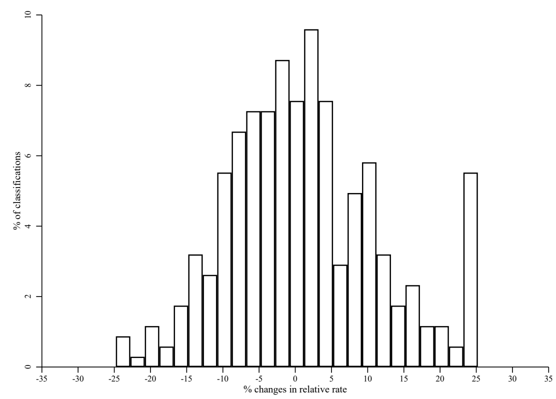
(a) Pooled



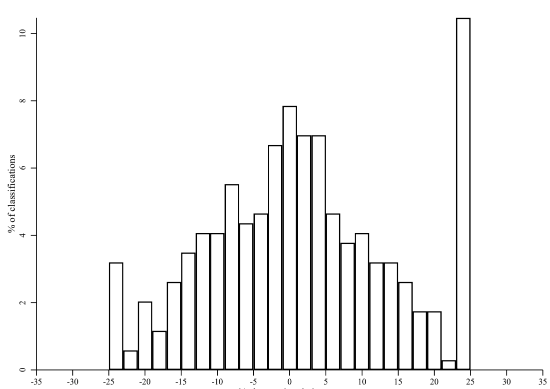
(b) Update 2007



(c) Update 2008



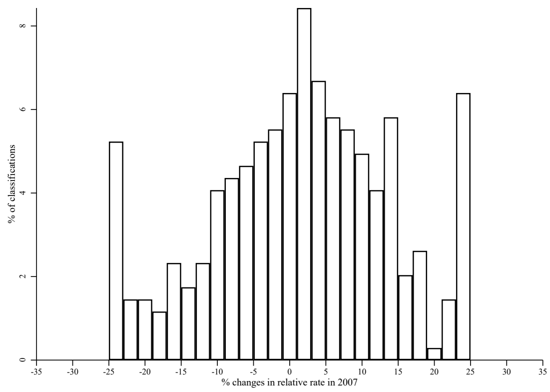
(d) Update 2009



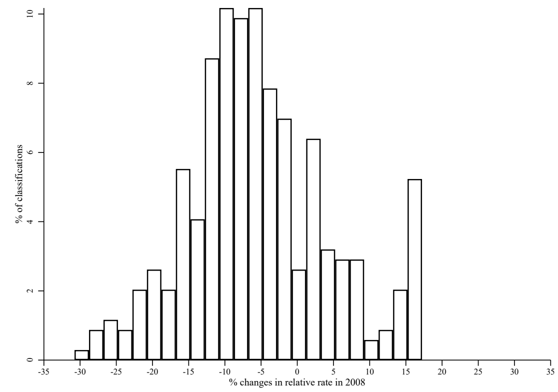
(e) Update 2011

Notes: The above histograms describe the proposed updates to the base rates (before any across-the-board adjustments). Following the definitions in Appendix Section B, the percent change here is defined as: $\frac{proRel_j - crtRel_j}{crtRel_j}$ for classification j. The updates in the final implemented base rates (after across-the-board adjustments) are depicted in Appendix Figure A2.

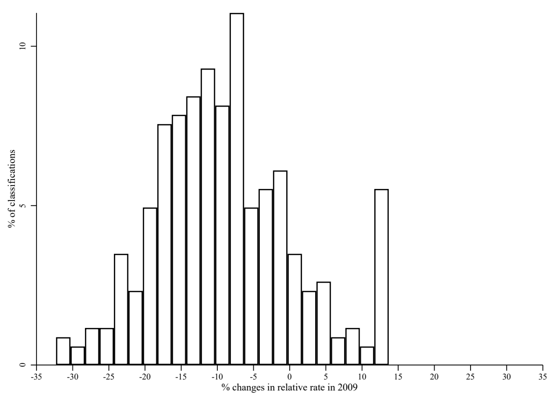
Figure A2: Histogram of Percent Change in Final Base Rates



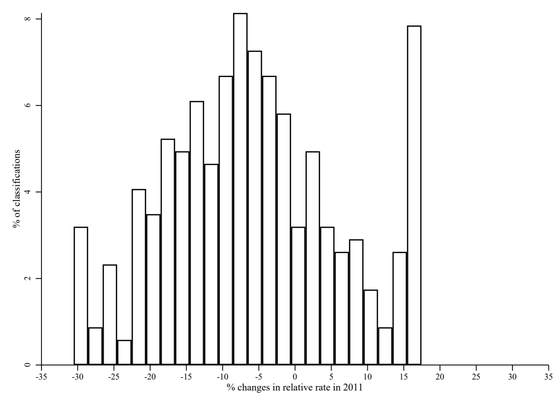
(a) Update 2007



(b) Update 2008



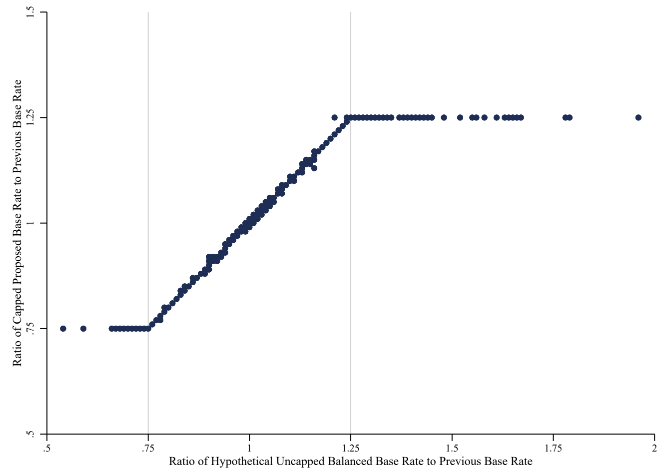
(c) Update 2009



(d) Update 2011

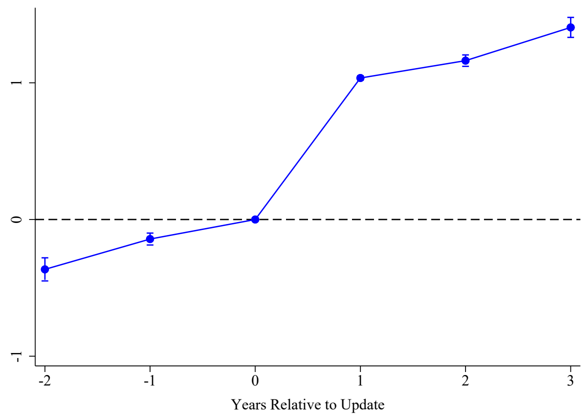
Notes: The above histograms describe the change in the final relative base rates. These histograms focus on the change in the final implemented base rates (after any across-the-board adjustments). Following the definitions in Appendix B, the percent change here is defined as: $\frac{\text{Final Base Rate}_{j-crtRel_j}}{crtRel_j}$ for classification j.

Figure A3: Base Rate Updates: Proposed Capped Rates and Hypothetical Uncapped Rates

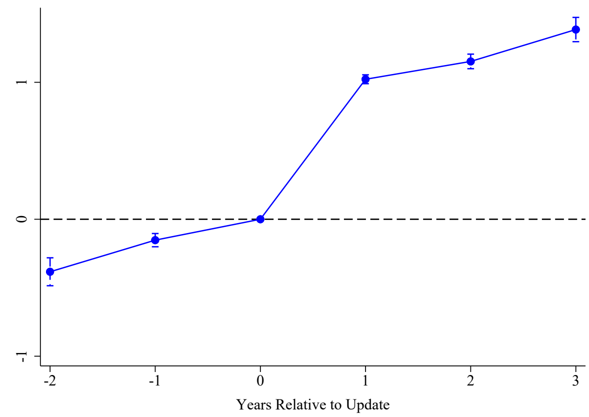


Notes: Each dot in the figure represents a classification update, where classification observations are pooled across updates in the analysis period (2006-2011). The figure displays a scatter plot of the following two ratios: the ratio of capped proposed relative base rate to previous base rate ($\frac{proRel_j}{crtRel_j}$ for classification j) and the ratio of hypothetical uncapped balanced base rate to previous base rate ($\frac{balRel_j}{crtRel_j}$ for classification j). See Appendix Section B for more details on these inputs into the base rate update algorithm.

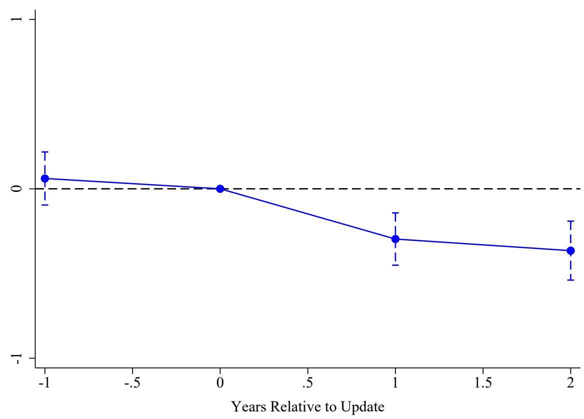
Figure A4: Event Study: Excluding Controls for Prior and Subsequent Rate Updates



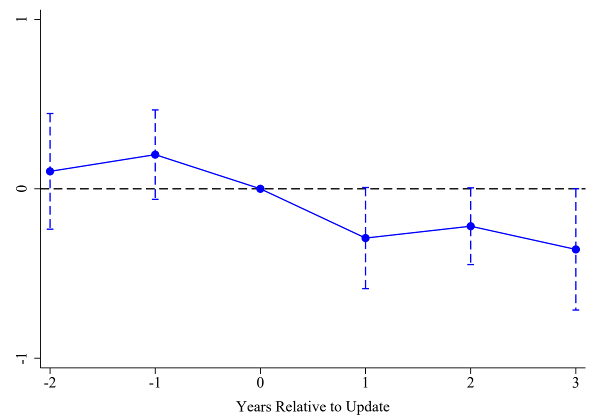
(a) ln(Base Rate)



(b) ln(Mean Premiums)



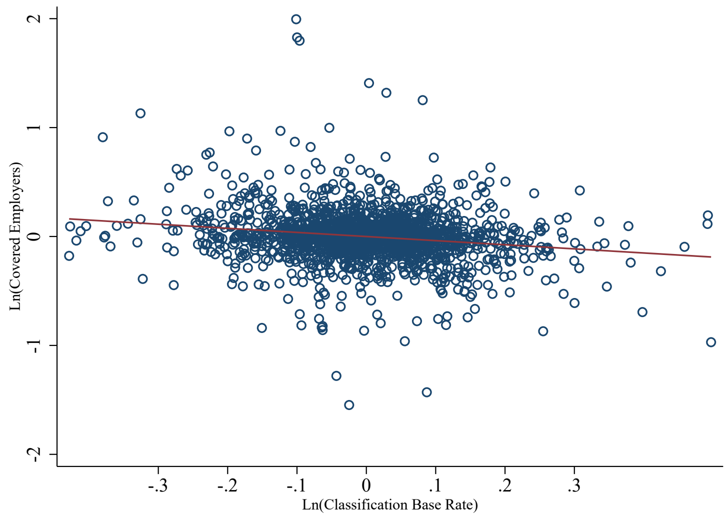
(c) ln(Covered Employers)



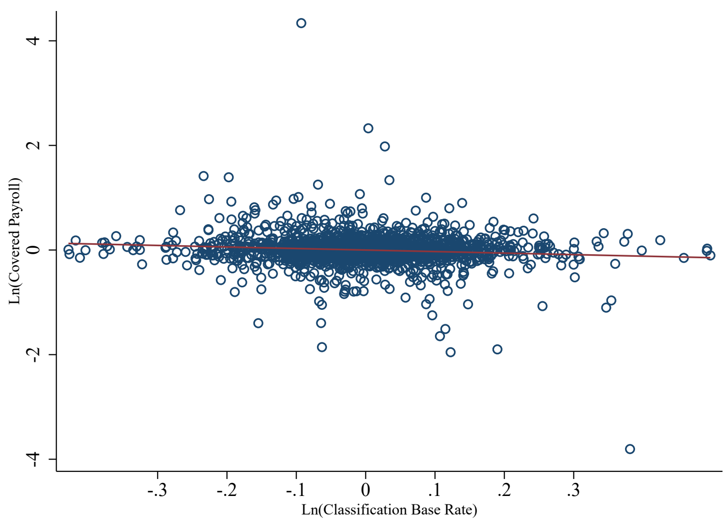
(d) ln(Covered Payroll)

Notes: The above figure displays estimates from the event study estimation described in equation 6 excluding controls for prior and subsequent rate updates. The dependent variables are as indicated in each panel. The horizontal axis displays time since the reference base rate update, where each point on the horizontal axis represents policies initiated in the indicated 12 month increment of event time. The event study representation focuses on the rate updates occurring between 2006 and 2011. The data used for this estimation is a series of balanced panels, where each panel includes data from three years (or two years in Panel C) pre- and post-update. Capped vertical bars indicate 95% confidence intervals, and robust standard errors are clustered at the classification level.

Figure A5: Graphical Depiction of Difference-in-Differences Demand Estimates



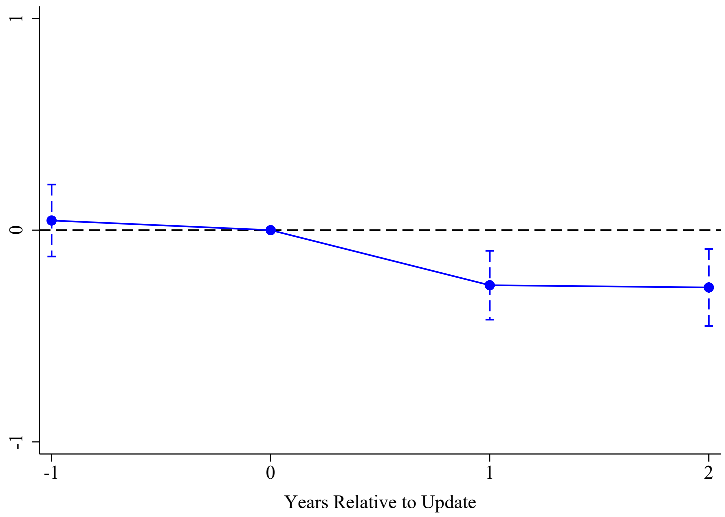
(a) Covered Employers



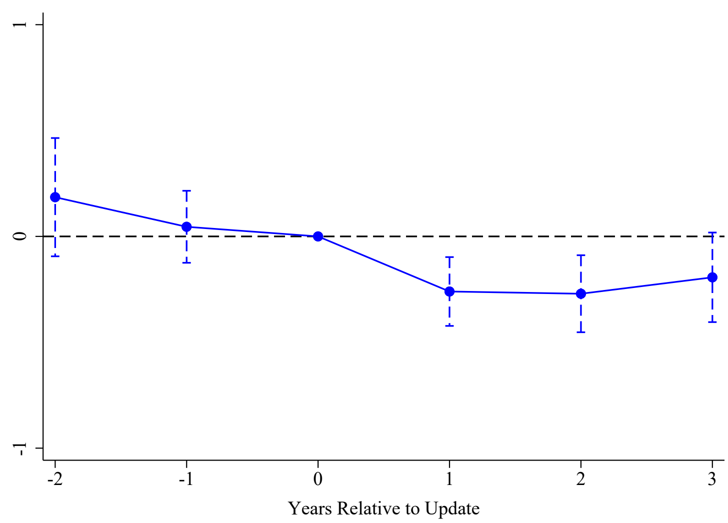
(b) Covered Payroll

Notes: This figure displays residual scatter plots for the baseline demand specifications. Each dot represents a classification-year observation in the baseline analysis data. Panel A displays the results for covered employers (analogous to the estimates in Table 3 column 1), and Panel B displays the results for covered payroll (analogous to the estimates in Table 3 column 5).

Figure A6: Event Study for Covered Employers: Alternative Specification



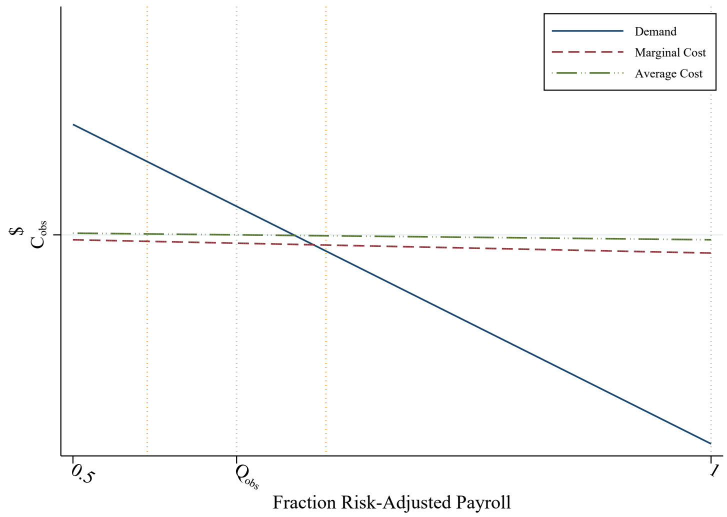
(a) Baseline: Balanced Panel



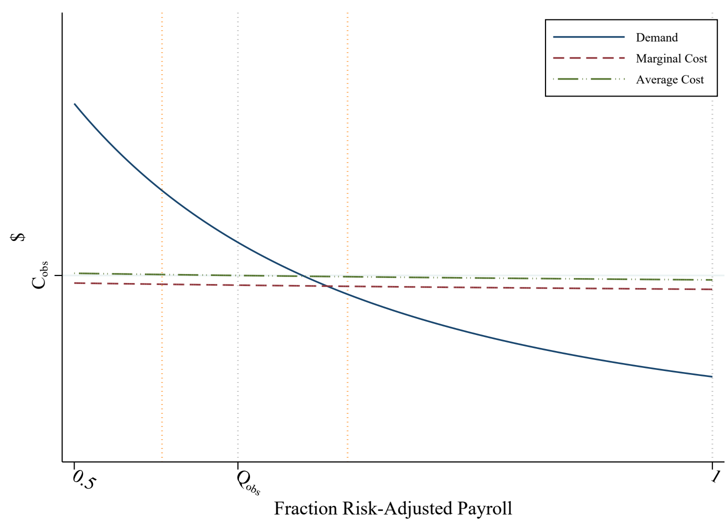
(b) Unbalanced Panel

Notes: The above figure displays estimates from the event study demand estimation described in equation 6. The dependent variable is the natural logarithm of covered employers. The horizontal axis displays time since the reference base rate update, where each point on the horizontal axis represents policies initiated in the indicated 12 month increment of event time. The event study representation focuses on the rate updates occurring between 2006 and 2011. The data used for this estimation is a series of panels, where each panel includes data from years pre- and post-update. Panel A focuses on a balanced panel (two years pre- and post- each update), while Panel B focuses on an unbalanced panel (three years pre- and post- update) excluding years for which the data is incomplete. Capped vertical bars indicate 95% confidence intervals, and robust standard errors are clustered at the classification level.

Figure A7: Selection: Graphical Illustration of Range of Magnitudes From Estimates



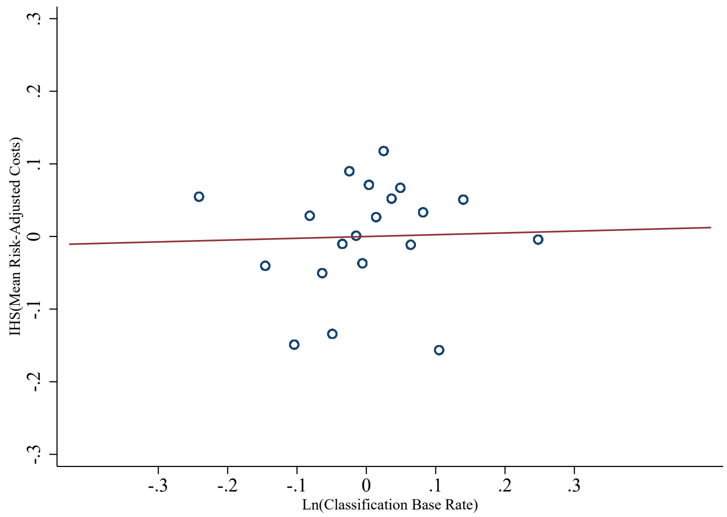
(a) Linear



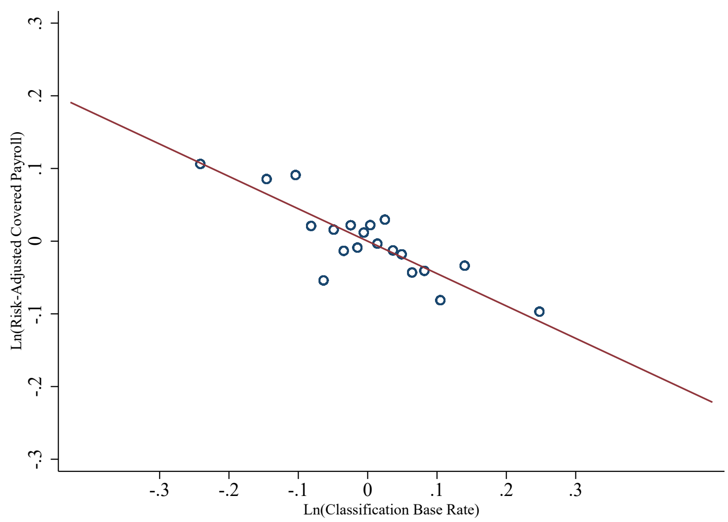
(b) Constant Elasticity

Notes: The above figure depicts a graphical representation of demand and costs based on the empirical estimates in Table 4. While the selection estimates presented in the text are not statistically distinct from zero, this figure plots the implied marginal and average cost curves based on the point estimates from Table 4 Panel A to give a sense of the magnitude of the point estimates. As discussed in the text, we obtain these curves by combining the estimated elasticities and aggregate summary statistics from the overall market on mean premiums, mean quantities, and mean insurer combined loss ratios. Panel A plots the estimates based on a linear extrapolation, while Panel B presents estimates based on a constant elasticity extrapolation. See Appendix Section C for further details on this calculation.

Figure A8: Graphical Depiction of Difference-in-Differences Selection Estimates



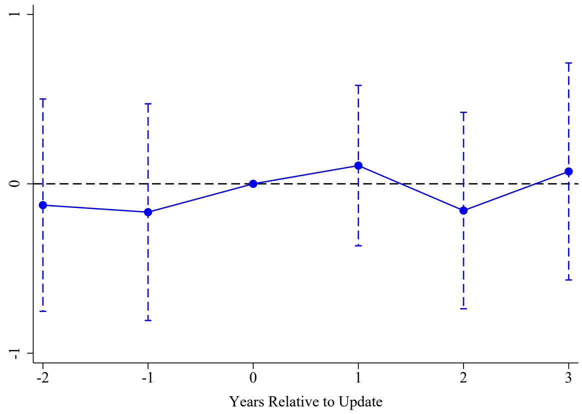
(a) Mean Risk-Adjusted Costs



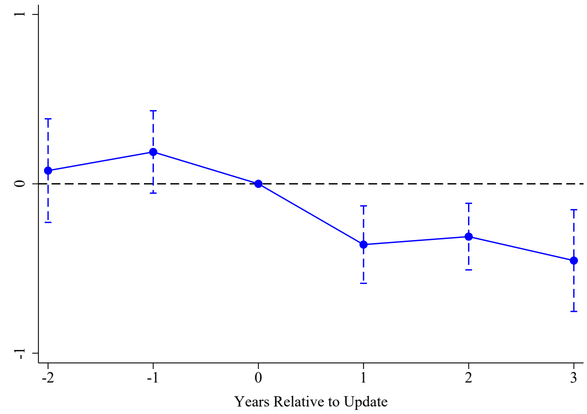
(b) Risk-Adjusted Covered Payroll

Notes: This figure displays binned mean residual scatter plots for the baseline selection analysis. Each dot represents 5% of the classification-year observations in the baseline analysis data, where bins are defined based on the values on the horizontal axis. Panel A displays the results for mean risk-adjusted costs (analogous to the estimates in Table 4 Panel A column 1), and Panel B displays the results for risk-adjusted covered payroll (analogous to the estimates in Table 4 Panel A column 2).

Figure A9: Event Study: Selection Estimates



(a) IHS(Mean Risk-Adjusted Costs)



(b) ln(Risk-Adjusted Covered Payroll)

Notes: The above figure displays estimates from the event study estimation described in equation 6. The dependent variables are: mean risk-adjusted costs (= costs per $10K of risk-adjusted payroll) (panel A) and risk-adjusted covered payroll (panel B). The horizontal axis displays time since the reference base rate update, where each point on the horizontal axis represents policies initiated in the indicated 12 month increment of event time. The event study representation focuses on the rate updates occurring between 2006 and 2011. The data used for this estimation is a series of balanced panels, where each panel includes data from three years pre- and post-update. Robust standard errors are clustered at the classification level.

Table A1: Comparison of Workers in Texas and Broader Populations

	Texas	States Recently Considering Mandate Repeal (Arkansas, Florida, Oklahoma, South Carolina, Tennessee)	All States
	(1)	(2)	(3)
Age	40.6	42.0	41.4
% Male	54.6%	52.3%	52.9%
% White	82.3%	79.9%	81.4%
% Married	57.1%	56.2%	55.8%
% Worked full time	82.7%	81.5%	79.3%
Family income	$72,743	$69,652	$78,668
Individual earnings	$35,757	$34,641	$37,293
Weekly earnings	$747	$724	$783
Industry (%)			
Agriculture/Forestry/Fishing/Hunting	2.5%	2.4%	2.6%
Arts/Entertainment/Accomodation/Food Services	11.6%	13.2%	12.3%
Finance/Real Estate/Professional Services	18.8%	18.7%	19.3%
Health Care/Educational Services	20.2%	20.9%	21.5%
Manufacturing	9.5%	9.2%	10.8%
Mining/Utilities/Construction	12.8%	10.3%	9.9%
Public Adminstration/Other Services	4.4%	5.1%	4.7%
Wholesale Trade/Retail Trade/Transportation	20.1%	20.1%	18.9%

Notes: This table compares the population of workers in Texas (column 1), in states recently considering mandate repeal (Arkansas, Florida, Oklahoma, South Carolina, Tennessee) (column 2) and in the entire United States (column 3) using data from the Current Population Survey Annual Social and Economic Supplement 2007-2012 (representing years 2006-2011) (Flood et al. (2021)). For this table, we define a worker as an individual with positive weeks of work reported in the prior year. All dollar values are CPI-U adjusted to 2006 dollars.

Table A2: Robustness: Workers' Compensation Classification Coding

	Dependent Variable: $I(c_{j,t} \neq c_{j,t-1})$	
	(1)	(2)
$\Delta ln(baseRate)_{c_{j,t-1}}$	-0.036	-0.019
	(0.039)	(0.046)
	[0.360]	[0.678]
Controls		
Employer Fixed Effects	x	x
Classification Fixed Effects	x	x
Year Fixed Effects	x	
Two-digit Classification X Year Fixed Effects		x
Mean Dep Var	0.087	0.087

Notes: The table above presents estimates from specifications as outlined in Appendix equation 10. These employer-year-level regressions include controls as listed above: employer fixed effects, classification fixed effects, year fixed effects (in column 1), and two-digit-classification-year fixed effects (in column 2). Each column represents a separate regression, where the estimated coefficients are displayed along with the associated standard errors in parentheses and p-values in brackets. The data used in these regressions cover the time period 2006-2011. This analysis focuses on employer-year observations where the employer is insured both in year t and year $t-1$ (N=789,223). Standard errors are clustered at the classification level. Both the classification-level clustering and the classification fixed effects described above are based on the classification in the prior year, $c_{j,t-1}$.

Table A3: Supplemental Evidence: Impact of Governing Base Rates on Overall Base Rates

	ln(Mean Overall Base Rate$_{jt}$)	ln(Mean Overall Premiums$_{jt}$)
	(1)	(2)
ln(baseRate$_{jt}$)	0.975	0.970
	(0.002)	(0.023)
	[<0.001]	[<0.001]

Notes: The table above presents estimates from specifications as outlined in Appendix equation 11. These classification-year-level regressions include year fixed effects and classification fixed effects. The dependent variables are as indicated in the table. Each column represents a separate regression, where the estimated coefficients are displayed along with the associated standard errors in parentheses and p-values in brackets. The data used in these regressions cover the time period 2006-2011. Standard errors are clustered at the classification level.

Table A4: Robustness: Demand Estimates Without Excluding Certified Self-Insured Employers

Dependent Variable: ln(covered employers$_{jt}$)				
	(1)	(2)	(3)	(4)
ln(baseRate$_{jt}$)	-0.380	-0.333	-0.389	-0.352
	(0.082)	(0.078)	(0.097)	(0.113)
	[<0.001]	[<0.001]	[<0.001]	[0.002]
ln(baseRate$_{j,t+2}$)			0.023	
			(0.113)	
			[0.841]	
ln(uncappedBaseRate$_{jt}$)*I(capBinding$_j$)				-0.049
				(0.130)
				[0.707]
Controls				
Classification Fixed Effects	x	x	x	x
Year Fixed Effects	x	x	x	x
Classification-specific Time Trend, 2-digit		x		

Notes: This table repeats the demand analysis in Table 3 columns 1 through 4 using all employers within the proof-of-coverage data, without excluding certified self-insured employers. The data used in these regressions cover the time period 2006-2011, where each observation represents a classification-year (N=2,058). The dependent variable is: ln(covered employers). Each column represents a separate regression, where the estimated coefficients are displayed along with the associated standard errors in parentheses and p-values in brackets. These classification-year-level regressions include year fixed effects and classification fixed effects. While column 1 reports the baseline specification, the remaining columns report alternative specifications with additional variables: 2-digit classification-specific time trends (column 2), leads of the legislated base rates (column 3) and uncapped base rates that were not ultimately adopted (column 4). These uncapped base rates correspond to the *balanced indicated relative base rates* discussed in Appendix Section B. Robust standard errors are clustered at the classification level.

Table A5: Descriptive Statistics

All Classification-Year Observations, 2006-2011		
	Mean	Std. Dev.
Classification Base Rate ($ per $100 in payroll)	2.41	3.32
Mean Premium ($ per $100 in risk-adjusted payroll)	2.19	3.07
Mean Premium ($ per $100 in payroll)	1.81	2.54
Mean Claim Cost ($ per $100 in payroll)		
All	0.69	1.22
Medical	0.43	0.83
Indemnity	0.26	0.48
Mean Claim Cost ($ per $100 in risk-adjusted payroll)		
All	0.84	1.53
Medical	0.52	1.03
Indemnity	0.32	0.61
Mean Claims (# per $50K in payroll)		
All	0.031	0.040
Serious Indemnity	4.03E-04	7.36E-04
Non-Serious Indemnity	0.008	0.011
Medical Only	0.022	0.030

Notes: This table describes the classification-year data from 2006 to 2011 used in the baseline demand analysis (N=2,064). The means in this table represent market-wide averages, weighting by the payroll covered within each classification. The mean claim cost measures described above capture mean claim costs (total unwinsorized incurred losses per $100 payroll or risk-adjusted payroll), where these claim costs are inclusive of both insurer costs and employer out-of-pocket costs. These cost measures reflect losses reported by the 42^{nd} month after the policy effective date, and we adjust these losses by TDI reported loss development factors to account for expected future reported costs related to these claims. In the above table, dollar quantities are adjusted using the CPI-U to be 2006 dollars.

Table A6: Eligible Population of Workers and Firms

Panel A: All Industries				
	ln(Total Number of Establishments)		ln(Total Number of Workers)	
	(1)	(2)	(3)	(4)
ln(baseRate$_{jt}$)	-0.107	-0.049	-0.106	-0.019
	(0.066)	(0.045)	(0.074)	(0.063)
	[0.109]	[0.273]	[0.149]	[0.764]
Controls				
Industry Fixed Effects	x	x	x	x
Year Fixed Effects	x	x	x	x
Industry-specific Time Trend, 3-digit		x		x

Panel B: Industries Mapping to Only One Classification				
	ln(Total Number of Establishments)		ln(Total Number of Workers)	
	(1)	(2)	(3)	(4)
ln(baseRate$_{jt}$)	0.003	0.062	0.110	0.113
	(0.080)	(0.064)	(0.144)	(0.100)
	[0.974]	[0.335]	[0.444]	[0.262]
Controls				
Industry Fixed Effects	x	x	x	x
Year Fixed Effects	x	x	x	x
Industry-specific Time Trend, 3-digit		x		x

Notes: The table above presents estimates from specifications as outlined in Appendix equation 16. In this table, i is a 6-digit NAICS industry, and t is a year. In this specification, $ln(b_{it})$ represents the natural logarithm of the mean base rate applicable in the industry based on a crosswalk between NAICS-classification codes. All specifications include year and industry fixed effects, and we estimate specifications with an additional control: a 3-digit NAICS industry-specific time trend. Each column represents a separate regression, where the estimated coefficients are displayed along with the associated standard errors in parentheses and p-values in brackets. Panel A focuses on a balanced sample of industry-year observations from industries with average annual employment exceeding 1,000 workers during 2006-2011 (N=3,582 industry-year observations), and Panel B further restricts attention to industries where there is a unique associated classification (N=540 industry-year observations). The dependent variables are as indicated in the table. Robust standard errors are clustered at the industry level. See Appendix Section D.4 for more details on this analysis.

Table A7: Event Study Regression Estimates

	ln(mean premiums $_{jt}$)	ln(covered employers$_{jt}$)	ln(covered payroll$_{jt}$)
	(1)	(2)	(3)
$I_{\{-2\}}$ X [ln(b_{1jp})-ln(b_{0jp})]	-0.017		0.028
	(0.021)		(0.159)
	[0.405]		[0.861]
$I_{\{-1\}}$ X [ln(b_{1jp})-ln(b_{0jp})]	-0.007	0.045	0.140
	(0.012)	(0.087)	(0.133)
	[0.578]	[0.601]	[0.294]
$I_{\{1\}}$ X [ln(b_{1jp})-ln(b_{0jp})]	0.999	-0.260	-0.268
	(0.016)	(0.083)	(0.124)
	[<0.001]	[0.002]	[0.031]
$I_{\{2\}}$ X [ln(b_{1jp})-ln(b_{0jp})]	0.977	-0.271	-0.196
	(0.022)	(0.093)	(0.111)
	[<0.001]	[0.004]	[0.078]
$I_{\{3\}}$ X [ln(b_{1jp})-ln(b_{0jp})]	0.989		-0.301
	(0.029)		(0.164)
	[<0.001]		[0.068]

Notes: The table above presents estimates from specifications as outlined in equation 6. The dependent variables are as indicated in each column: the natural logarithm of mean premiums per $100 of risk-adjusted payroll (column 1), the natural logarithm of covered employers (column 2), and the natural logarithm of covered payroll (column 3). Time in this specification indicates time since the reference base rate update, where observations represent policies initiated in the indicated 12 month increment of event time. The event study specification focuses on the rate updates occurring between 2006 and 2011. The data used for this estimation is a series of balanced panels, where each panel includes data from three years (or two years in column 2) pre- and post-update. Robust standard errors are clustered at the classification level.

Table A8: Demand Estimates: Additional Robustness Analysis

Panel A: Robustness, Alternative Assumption on Incidence of Workers' Compensation Premium Changes					
	Dependent Variable: ln(covered payroll, normalized$_{jt}$)				
	% of premiums borne by employees				
	0% (baseline)	10%	25%	50%	100%
	(1)	(2)	(3)	(4)	(5)
ln(baseRate$_{jt}$)	-0.293	-0.287	-0.277	-0.260	-0.222
	(0.122)	(0.122)	(0.122)	(0.122)	(0.122)
	[0.017]	[0.019]	[0.024]	[0.034]	[0.070]

Panel B: Robustness, alternative samples			
	Dependent Variable: ln (covered payroll$_{jt}$)		
	(1)	(2)	(3)
ln(baseRate$_{jt}$)	-0.293	-0.292	-0.371
	(0.122)	(0.123)	(0.154)
	[0.017]	[0.018]	[0.016]
Sample	baseline	drop common secondary classes	drop small classes

Panel C: Robustness, alternative weighting				
	Dependent Variable: ln(covered payroll$_{jt}$)			
	(1)	(2)	(3)	(4)
ln(baseRate$_{jt}$)	-0.293	-0.210	-0.241	-0.197
	(0.122)	(0.104)	(0.114)	(0.095)
	[0.017]	[0.044]	[0.036]	[0.038]
Weights	unweighted	baseline covered payroll	baseline risk-adjusted covered payroll	baseline premiums

Notes: The table above presents robustness analysis from the difference-in-differences demand estimation outlined in equation 2. The data used in these regressions cover the time period 2006-2011, and each regression includes year fixed effects and classification fixed effects. Each column represents a separate regression, where the estimated coefficients are displayed along with the associated standard errors in parentheses and p-values in brackets. Panel A displays robustness analysis under alternative assumptions on the incidence of changes in workers' compensation premiums. Specifically, these additional specifications repeat the baseline payroll regression replacing the dependent variable with the natural logarithm of normalized covered payroll: $ln(\frac{\text{payroll}_{jt}}{1-\theta\times\text{premium}_{jt}})$, where premium_{jt} represents the mean premium per dollar of payroll for classification j in year t and θ represents the fraction of premiums shifted to workers in the form of reduced wages. The corresponding assumption on the incidence of premium changes (the value of θ) is denoted in each column. Panel B displays robustness analysis using alternative samples: the baseline analysis data (column 1; N=2,064), dropping common secondary classifications (column 2; N=2,046), and dropping small classifications with less than $10 million in mean annual covered payroll (column 3; N=1,716). Panel C displays alternative specifications weighting regressions by the indicated weights. In all panels, robust standard errors are clustered at the classification level.

Table A9: Robustness: Selection Estimates

	ln(baseRate$_{jt}$)			Implied ΔQ : $Q^{optimal}$-Q^{CE}		Implied DWL from selection relative to optimal	
	Est	Std Err	p-value	Linear	Const Elasticity	Linear	Const Elasticity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) baseline	0.025	(0.265)	[0.926]	0.015	0.018	0.00080	0.00090
Alternative Specifications							
(2) shifted natural logarithm, ln(x+1)	0.030	(0.244)	[0.901]	0.018	0.023	0.00123	0.00139
(3) natural logarithm	0.086	(0.192)	[0.653]	0.057	0.076	0.01103	0.01250
Alternative Cost Measures							
(4) unwinsorized	0.009	(0.280)	[0.974]	0.005	0.006	0.00011	0.00012
(5) winsorize at 99.9 percentile	0.012	(0.275)	[0.965]	0.007	0.009	0.00019	0.00022
(6) winsorize at 98 percentile	0.028	(0.258)	[0.914]	0.016	0.020	0.00102	0.00115
(7) undeveloped losses	0.034	(0.259)	[0.897]	0.020	0.025	0.00151	0.00170

Notes: The table above presents alternative specifications for the cost regressions. The coefficients reported above are from a difference-in-differences specification as outlined in equation 2. These classification-year-level regressions include the following controls: year fixed effects and classification fixed effects. Each row represents a separate regression, where the table displays the estimated coefficient (column 1), standard error (column 2) and p-value (column 3). Robust standard errors are clustered at the classification level. The data used in these regressions cover the time period 2006-2011, where each observation represents a classification-year (N=2,030 in natural logarithm specification in row 3 and N=2,064 in all other specifications). Unless otherwise noted, the dependent variable is the inverse hyperbolic sine of costs per $10K of risk-adjusted payroll. See Appendix Section C for more details on risk-adjustment used in this analysis. Columns 4 and 5 report the implied difference in insured quantity, comparing the optimal quantity to the competitive equilibrium under a linear and constant elasticity extrapolation, respectively. Columns 6 and 7 report the implied deadweight loss of selection, comparing the competitive equilibrium to the optimal allocation, based on a linear and constant elasticity extrapolation, respectively.

Table A10: Selection Estimates: Empirical Cost Curves

	Linear		Constant Elasticity	
	Est	Std Err	Est	Std Err
	(1)	(2)	(3)	(4)
Demand Curve				
Constant	7.1229	(1.56)	0.7736	(0.21)
Slope	-7.8431	(2.49)	-2.2441	(0.71)
Average Cost Curve				
Constant	1.9478	(1.21)	1.7985	(0.53)
Slope	-0.1629	(1.93)	-0.0555	(0.66)
Marginal Cost Curve				
Constant	1.9478	(1.21)	1.6988	(1.73)
Slope	-0.3258	(3.85)	-0.0555	(0.66)

Notes: The table above reports the implied linear and constant elasticity parameters for demand as a function of the quantity insured based on the estimates in Table 4 Panel A. In this table, the "constant" and "slope" in the constant elasticity specification ($P = AQ^{\beta}$) refer to A and β, respectively; in the linear specification ($P = A + \beta Q$), the "constant" and "slope" refer to A and β, respectively. The table reports bootstrapped standard errors clustered at the classification level, where 1,000 randomly drawn bootstrap samples are used.

Table A11: Robustness: Welfare Calculations with Empirical Cost Curves

	Linear		Constant Elasticity	
	Baseline	Alternative	Baseline	Alternative
	(1)	(2)	(3)	(4)
Counterfactuals				
Quantity (fraction risk-adjusted payroll covered)				
Mandate	1.000	1.000	1.000	1.000
Optimal	0.673	0.688	0.679	0.698
Welfare (relative to status quo)				
Mandate				
per $100 of risk-adjusted payroll	-0.4118	-0.3513	-0.1992	-0.1522
scaled by $50,000	-205.91	-175.63	-99.59	-76.10
% of mean premium	-18.8%	-16.0%	-9.1%	-6.9%
Optimal				
per $100 of risk-adjusted payroll	0.0078	0.0136	0.0085	0.0149
scaled by $50,000	3.91	6.81	4.24	7.45
% of mean premium	0.4%	0.6%	0.4%	0.7%
Subsidy to support optimal allocation--25% MDWL of taxation				
per $100 of risk-adjusted payroll	-0.0511	-0.0676	-0.0509	-0.0659
scaled by $50,000	-25.55	-33.82	-25.47	-32.93
% of mean premium	-2.3%	-3.1%	-2.3%	-3.0%

Notes: The table above presents alternative welfare calculations that use the implied empirical cost curves based on the elasticities in Table 4 Panel A. Columns 1 and 3 display the baseline welfare estimates for reference. The table reports welfare measured in dollars per $100 of risk-adjusted payroll. In addition, the table reports two scaled measures of welfare to ease interpretation: (i) welfare measures scaled by $50K, approximately the mean annual earnings for this population and (ii) welfare as a percent of mean premiums observed in the status quo (one measure of the size of the market). See Table 5 and Appendix Section C for further details on the welfare calculations.