Workers Compensation Experience Rating and the Incidence of Workplace Injuries

Jonathan Gruber, December 2018

One potentially fruitful avenue for addressing the high costs of Disability Insurance (DI) is to be engage employers further in deterring employees from leaving work to join the program. One possible means of increasing employer engagement is by following private insurance products in introducing experience rating.

This project studies the effects of experience rating within the Workers Compensation (WC) program, the nation's largest social insurance program. I use unique data from a major WC insurer to model the impact of experience rating on firm behavior. These data allow me to measure the marginal cost of additional injuries, and to assess how higher marginal costs affect the incidence and duration of injuries. This allows me to address the question of whether experience rating programs such as DI would impact worker injury probabilities and costs.

In this report I detail the empirical strategy used in the AIG data to assess whether experience rating impacts utilization of the WC program. Despite detailed data construction and multiple careful empirical strategies, I am unable to draw a firm conclusion.

1. Background

Workers' compensation (WC) policies insure employers against medical and other (e.g. wage replacement) costs associated with on-the-job injuries. Most states mandate WC coverage for most or all employers. Except in states that run monopolistic systems (Ohio, Washington, North Dakota, and Wyoming), firms have a choice of insurance carriers for WC policies. However, unlike with other types of insurance (e.g. health insurance), firms typically do not have a choice of policies within a particular carrier. For example, AIG will write a single policy that is tailored to the characteristics of a firm, and the firm will then decide whether to accept AIG's policy or use a different carrier.

WC premiums are calculated using a complex, multi-step process. Many of the steps are mandated by state regulations, while others depend on factors specific to the insurance carrier (e.g. expense constants). One mandatory step involves applying a multiplicative adjustment to the premium faced by a firm called the EMOD which is based on the firm's prior claim experience. The process of calculating and applying the EMOD, thereby adjusting the premium on the basis of prior claims experience, is called **experience rating** and is the topic of this research. Specifically, the goal of this research is to determine whether or not firm behavior is influenced by experience rating.

Fundamentally, an EMOD is calculated by comparing actual losses to expected losses over an experience period. Actual losses refer to total payments by insurance carriers as a result of claims incurred by the insured firm. Expected losses refer to the total payments that would be expected for a firm of a similar size whose employees are engaged in similar occupations as the insured firm. It is important to keep in mind that, in this context, the term "expected losses" refers to an actuarially assigned value used in experience rating, *not* to the firm's subjective expectations as to its future losses.

A policy's experience period refers to the historical period over which the firm's loss history enters into a given year's EMOD. In most (though not all) cases, the experience period for a policy that is effective in year t will be years t – 4 through t – 2. For example, the EMOD for a 2015 policy will be based on

This research was supported by the U.S. Social Security Administration through grant #5 DRC12000002-06 to the National Bureau of Economic Research as part of the SSA Disability Research Consortium. The findings and conclusions expressed are solely those of the author(s) and do not represent the views of SSA, any agency of the Federal Government, or the NBER.

expected and actual losses from that firm's 2011 through 2013 policies. Every policy year within the experience period is included in the year t EMOD even if the firm was covered by a different carrier at the time. For most firms, therefore, a claim occurring in year t will first affect the EMOD in t + 2, even if the firm changes carriers.

Not all firms are eligible for experience rating. Eligibility depends on whether the average or total subject premium over the experience period exceeds a state-specific threshold. This is an important point that is relevant to our identification strategies—eligibility for experience rating depends on historical premiums, not current premiums. Note that this implies that a firm must exist for several years before it is eligible for experience rating. Generally, a firm with an EMOD below 1 has actual losses below their expected losses during the experience period and receives a "discount", while a firm with an EMOD above 1 has actual losses above their expected losses and receives a "penalty". For firms that are not eligible for experience rating, the EMOD is effectively 1 regardless of loss history. Experience rating is mandatory for firms that do meet the eligibility criteria. Note the EMOD is calculated by quasi-regulatory bodies known as rating bureaus, not specific insurance carriers like AIG.

For nearly all states, the applicable rating bureau is a national agency called the National Council on Compensation Insurance (NCCI). Five states (Massachusetts, Minnesota, New York, North Carolina, and Wisconsin) have independent bureaus that rate single-state policies (policies that only have coverage in that state), using formulas identical or similar to the NCCI formula. However, multistate policies with coverage in these states are still rated by NCCI. For multistate NCCI policies, a single EMOD, calculated based on the sum of actual and expected losses from all states covered by the policy, is applied to the entire policy.

Five other states (Pennsylvania, Delaware, California, Michigan, and New Jersey[1]) have independent bureaus that are responsible for all policies with coverage in those states. Therefore, a multistate policy with coverage in one or more of these states will have different EMODs applied to different states. For example, a policy with coverage in Connecticut (an NCCI state), New York (a non-NCCI state that partners with NCCI for multi-state policies), and Pennsylvania (a "purely" non-NCCI state) will have one EMOD for its Connecticut and New York coverage and a second for its Pennsylvania coverage. If this policy also had coverage in California, it would have a third separate EMOD for its California coverage. Most of these purely non-NCCI states use formulas that differ meaningfully from the NCCI formula. These formulas and associated parameters, which vary by state, year, and firm characteristics, are described in the appendix.

2. Introduction to methodology

For most firms, a claim occurring in year *t* will first affect the EMOD (and, through it, the final premium) in year *t + 2*. Our strategy is based on a "forward-looking" model of firm behavior. It assumes that, on average, firms have a sufficient understanding of the EMOD formula, and of the specific parameters that apply to them, to project the change in the EMOD (and the resulting change in the final premium) that would occur in year *t + 2* if an additional claim were incurred in year *t*. The hypothesis is that, all else equal, firms for which the marginal impact of a new claim is larger (that is, firms that are "more heavily" experience-rated) will have less frequent and/or less severe claims in year *t* as the firm will adjust practices in order to reduce claims.

This implies the following "core" regression specification:

[1] The New Jersey bureau reports very limited rating information and, therefore, our experience-rated sample currently excludes New Jersey policies.

$$(1)\ y_t = \beta \frac{\partial M_{t+2}}{\partial C_t} + \boldsymbol{\gamma X}$$

where y is a measure of claim frequency or severity, M is the EMOD, C is claim dollars, and $\mathbf{X}$ is a vector of controls. If the final premium (π_f) is the independent variable of interest, then the specification is:

$$(2)\ y_t = \beta \frac{\partial \pi_{f,t+2}}{\partial C_t} + \boldsymbol{\gamma X}$$

In both specifications, we hypothesize a negative β. For brevity, in the remainder of this report, the $t + 2$ and t subscripts are omitted from the numerator and denominator, respectively, of $\frac{\partial M}{\partial C}$ and $\frac{\partial \pi_f}{\partial C}$. However, it is crucial to keep in mind our central assumption: that a firm is making decisions on safety practices in year t based on the impact that a claim in year t would have on its EMOD in $t + 2$.

The key identification challenge for estimating Equations 1 and 2 is that the derivatives of the EMOD and final premium are driven by endogenous firm characteristics that are likely to be independently correlated with claims frequency/severity. Sections 2.1 and 2.2 demonstrate this by deriving the equations for $\frac{\partial M}{\partial C}$ and $\frac{\partial \pi_f}{\partial C}$. Sections 2.3 through 2.5 then discuss our three primary approaches to estimating β. Section 2.5 provides a brief overview of two other strategies we considered and our reasons for shifting away form them.

2.1. Derivative formulas for the EMOD

The derivative of the EMOD with respect to the marginal claim dollar is straightforward to derive from the applicable EMOD formulas. For all rating bureaus except California (post-2011), Pennsylvania, and Delaware, the formula implies that this derivative is given by:

$$(3)\ \frac{\partial M}{\partial C} = \begin{cases} \dfrac{m_{st}}{E + b_s(E)} & if\ C \leq p_s \\ \dfrac{w_s(E) m_{st}}{E + b_s(E)} & if\ C > p_s \end{cases}$$

In Equation 3, m_{st} is 0.3 if the claim is medical-only and occurs in a state that observes the medical-only adjustment, and 1 otherwise; E is expected losses (calculated using Equation 6, below); $b_s(E)$ is the ballast value, which varies by expected losses according to a function determined by the state(s) in which the firm has payroll; $w_s(E)$ is the credibility weight, which also varies by a state-determined function of expected losses; and p_s is the state-determined split point.

The derivative implied by the post-2011 California formula is:

$$(4)\ \frac{\partial M}{\partial C} = \begin{cases} \dfrac{w_p(E)}{E} & if\ C \leq p_s \\ \dfrac{w_e(E)}{E} & if\ C > p_s \end{cases}$$

where $w_p(E)$ and $w_e(E)$ are the credibility weights applied to primary and excess losses, respectively; both of which are functions of expected losses.

The derivative calculated from the Pennsylvania and Delaware formulas is:

$$(5)\ \frac{\partial M}{\partial C} = \begin{cases} \frac{w(E)}{E} & if\ C \leq p_s \\ 0 & if\ C > p_s \end{cases}$$

Expected losses are calculated as:

$$(6)\ E_t = \sum_{h=t-4}^{t-2} \sum_{s} \sum_{c} R_{csh} l_{cst}$$

where R_{csh} is the payroll in class code *c*, state *s*, and historical year *h*; and l_{cst} is the expected loss rate (ELR) applied to class code *c* by state *s* in year *t*. Note that Pennsylvania and Delaware use a slight variation on Equation 6, in that ELRs are allowed to vary with respect to the historical policy year (i.e. *l* is subscripted with *h* rather than *t*).

In all cases, therefore, the derivative of the EMOD with respect to the marginal claim dollar depends on: (1) whether the marginal claim dollar is from a medical-only or lost-time claim; and (2) whether the marginal claim dollar is above or below the split point. See Section 2.2 for a discussion of our method of accounting for the fact that the derivative depends on the source of the marginal claim dollar.

Note that Equations 3 through 5 do not account for the state accident limitation. If the marginal claim dollar is above the state accident limitation, the derivative will be 0. Because claims above the accident limitation are rare, we do not account for it in our calculations, thereby implicitly assuming that it does not enter into firms' decision-making.

Also note that Equations 3 through 5 imply that the derivative of the EMOD with respect to the marginal claim dollar does *not* depend on the number or size of claims that the firm has already incurred. From the firm's perspective in year *t*, the *t + 2* EMOD will depend on losses incurred in years *t – 2* through *t*; however, losses incurred in *t – 2* and *t – 1* are, in effect, sunk costs that do not change the impact of a new claim. The only exception to this is for firms that have already incurred sufficient losses to place their *t + 2* EMOD at the applicable cap. If the *t + 2* EMOD is already at the cap, then an additional claim in year *t* will not increase it, and the derivative is therefore 0. Additional discussion of how we account for the cap is provided in Section 2.

2.2. Derivative formulas for the final premium

Based on AIG's internal "Introduction to Workers Compensation" manual and other documentation, the major steps in calculating workers' compensation premiums are:

1. **Manual premium**: Calculated by dividing current year payrolls into class-state cells, multiplying them by the manual rate for each class-state, and then summing to the policy level.
2. **Subject premium**: Calculated by applying several adjustments to the manual premium, which may vary by state. These include the waiver of subrogation factor and the employer's liability increased limits factor. Based on AIG's documentation, most of these adjustments appear to be multiplicative (i.e. are calculated as a percentage of the manual premium), but there may be some that are applied as flat charges.
3. **Modified premium**: Calculated by multiplying the subject premium by the EMOD.
4. **Standard premium**: For most policies, this equals the modified premium multiplied by a schedule modification factor. The schedule modification is an adjustment (which, in nearly all

states, is applied at the insurer's discretion) based on observable firm characteristics that may increase or decrease risk—for example, credits may be applied for firms that have drug-free workplace policies. Like the EMOD, it is a multiplicative adjustment. According to the documentation, the standard premium sometimes incorporates other adjustments, which can be additive (e.g. a charge for nonratable catastrophe loading) and/or multiplicative (e.g. a CPAP factor).

5. **Gross premium**: For our purposes, this is the final premium. It is equal to the standard premium multiplied by a discount factor, plus additional adjustments. The other adjustments include an expense constant, which, according to the documentation, should be additive in nature; however, according to AIG, it is sometimes calculated as a percentage of the total premium. Others include disease charges for coal mines (it is unclear whether this is additive or multiplicative) and a terrorism risk charge (which appears to be multiplicative).

Therefore, a generalized representation of the full premium formula is:

$$(7)\ \pi_f = [(\pi_m f_1 + a_1)MSf_2 + a_2]f_3 + a_3$$

In Equation 7, π_m is the manual premium, f_1 represents any multiplicative adjustments incorporated into the subject premium, a_1 represents any additive adjustments incorporated into the subject premium, M represents the EMOD, S represents the schedule modification factor, f_2 represents any other multiplicative adjustments incorporated into the standard premium, a_2 represents any additive adjustments incorporated into the standard premium, f_3 represents any multiplicative adjustments incorporated into the gross premium, and a_3 represents any additive adjustments incorporated into the gross premium. Note that f_1, a_1, f_2, a_2, f_3, and a_3 are composites that may each incorporate multiple individual adjustments.

The manual premium is given by:

$$(8)\ \pi_{m,t} = \sum_s \sum_c R_{cst} r_{cst}$$

where R_{cst} is the payroll in state s and class c in year t, and r_{cst} is the manual rate applicable to state s and class c in year t. Note that this is similar to the calculation of the expected losses used in the denominator of the EMOD (Equation 6), but not identical. The manual rate differs from the ELR, and the manual premium is based on payrolls in year t, whereas expected losses are based on payrolls summed over the experience period. The manual rate is set at the state level, but is often adjusted by a deviation factor that reflects expenses, tax, and profits specific to the AIG subsidiary ("paper company") that holds the policy.

Equation 7 implies that the derivative of the final $t + 2$ premium with respect to the marginal claim dollar in year t is given by:

$$(9)\ \frac{\partial \pi_f}{\partial C} = [(\pi_m f_1 + a_1)f_2 f_3 S]\frac{\partial M}{\partial C}$$

Note that the key assumption embedded in Equation 9 is that, of the adjustment factors incorporated into the $t + 2$ final premium (Equation 7), the EMOD is the only one that is a function of claims in year t. This is clearly true for the manual premium, which depends only on payrolls and manual rates. Based on AIG's documentation, it also appears to be a reasonable assumption for the various miscellaneous adjustment factors that enter into the subject, standard, and gross premiums (the various adjustments that make up f_1, a_1, f_2, a_2, f_3, and a_3).

For the schedule modification factor (*S*), this may be a stronger assumption. In most cases, *S* reflects observable firm behaviors or policies (e.g. the existence of certain safety equipment) that are believed by the insurer to mitigate or amplify risk. There are two scenarios in which the value of *S* in year $t + 2$ could become a function of claims in year *t*. First, it is possible that a claim in year *t* could lead the firm to install new safety equipment that subsequently leads to a schedule credit in $t + 2$. We do not have the data that would be needed to model this channel; in particular, we lack information on the actual safety features or other characteristics that led to the firm receiving a schedule credit or debit in a given year. For this reason, we cannot account for this channel in our calculation of the derivative and must assume that it is not operative. Second, for larger firms, AIG actuaries will sometimes implement manual adjustments to the premium that are based in part on the firm's claims history—in essence, implementing an internal experience modification process to improve upon the EMOD calculated by the rating bureau. When this occurs, any resulting debits or credits should be reflected in *S*, which means that *S* may become a direct function of *C*. However, because this manual adjustment process is a "black box," we again do not have the data necessary to model it. Furthermore, according to one of our contacts at AIG, AIG generally only implements this process for firms with expected losses above $750,000, although this is not an exact threshold. For this reason, we have excluded such firms from our master sample (and therefore from all analyses). To the extent that this threshold is accurate, this exclusion should break any relationship between *S* and *C* caused by this internal experience modification process.

In order to calculate $\frac{\partial \pi_f}{\partial C}$ using Equation 9, we would need to observe each of the individual adjustment factors that enter into the subject, standard, and gross premiums. In practice, while we observe the major steps in the calculation, we do not observe the individual adjustment factors. For example, while we can observe the manual and subject premium, and thus the difference between them, we cannot determine what portion of that difference is made up of multiplicative adjustments and what portion is made up of additive adjustments. It is clear from Equation 9 that the distinction between additive and multiplicative adjustments is crucial for accurately projecting the impact of a marginal claim on the final premium. For this reason, in order to obtain an estimate of the derivative from our available data, we need to make simplifying assumptions about the nature of the adjustments. These are discussed further in Section 2.3. The reliance on assumptions implies that our estimates of the derivative of the premium are likely to incorporate greater measurement error than our estimates of the derivative of the EMOD.

2.3. Identification using controls: projected EMOD

As shown in Equations 3 through 5, expected losses enter into the denominator of $\frac{\partial M}{\partial C}$. As shown in Equation 6, expected losses are themselves a function of firm size (reflected by total payrolls) and the riskiness of the firm's operations (reflected by the classes and possibly states across which those payrolls are distributed). This introduces a negative relationship between $\frac{\partial M}{\partial C}$ and firm size and riskiness. Since larger and/or riskier firms are likely to have more claims, this implies a negative bias in β when the derivative of the EMOD is the independent variable of interest. Additionally, other parameters that enter into the derivative (i.e. credibility weights and ballast values) are themselves stepwise functions of expected losses. This adds further complexity and nonlinearity to the relationship between $\frac{\partial M}{\partial C}$ and firm size and riskiness.

When the derivative of the final premium is the independent variable of interest, there are additional sources of possible bias. Most obviously, the manual premium, like expected losses, is a function of firm size and riskiness. Because the manual premium is in the numerator of $\frac{\partial \pi_f}{\partial C}$, this introduces a *positive* bias

into β. We have no *a priori* reason to know whether or to what extent this offsets the negative bias driven by expected losses, so the direction of the net bias is theoretically ambiguous. Furthermore, to the extent that the other adjustment factors incorporated into Equation 9 are independently correlated with claims, they may introduce additional bias in unknown directions.

Of course, the standard approach to addressing omitted variable bias is to include controls. Therefore, in one version of this strategy (which we refer to as the "projected EMOD" approach), we use the actual value of the derivative as the independent variable of interest and a variety of controls which are described in Section 3.4. We exclude multistate policies from the analytic sample for the projected EMOD approach. This is because we are unable to control for variations in the distribution of payrolls by state, conditional on the governing state.

2.4. Identification using simulation: simulated EMOD/premium

Given the highly nonlinear relationships between firm size and risk and the derivatives of the EMOD and premium, it is difficult to be confident that even relatively rich controls for these characteristics can fully remove the omitted variable bias from the estimate of β. An alternative approach is to create a version of the derivative that reflects only the exogenous components of formula (i.e. the state-, year-, and class-level parameters). This "simulated derivative" acts, in effect, as an instrument for the actual derivative. Equations 1 and 2 are estimated using the simulated derivative as the independent variable of interest, which implies that they become reduced-form instrumental variable specifications.

The major steps to calculate the simulated derivatives are as follows:

1. Calculate the median historical payroll (summed over the experience period) by class code.
2. Replace each policy's total historical payroll (again, summed over the experience period) with the median calculated in step 1 for its governing class code.
3. Assign all of each policy's simulated payrolls to its governing class code.
4. For multistate policies, assign all of each policy's simulated payrolls to its governing state.
5. For each of the adjustment factors that enter into the derivative of the premium (Equation 9), replace each firm's actual adjustment factor with the median within its governing class code. Note that this requires some simplifying assumptions to be made about the premium calculation model, as discussed in Section 2.3.
6. Using Equations 3 through 6, 8, and 9, calculate the simulated derivative of the EMOD and premium based on the simulated payrolls created in steps 1 through 4.

Step 2 removes variation in firm size conditional on the governing class code. Step 3 removes variation in the distribution of payrolls across class codes, again conditional on the governing class code. Step 4 removes variation in the distribution of payrolls across states, conditional on the governing state. Steps 2 through 4 together remove all sources of endogenous variation in the derivative of the EMOD and in the manual premium. Step 5 further removes any additional endogenous variation in the other adjustment factors that enter into the derivative of the final premium.

Therefore, the simulated derivatives calculated in step 6, after controlling for the governing state, governing class code, year, and carrier code[2], should vary only with respect to the exogenous parameters (weights, ballast values, and ELRs) that enter into the derivative. For this reason, we would expect this

[2] Controlling for the carrier code is necessary because the manual rates may differ between AIG subsidiaries.

specification to be less sensitive to controls, although this is not guaranteed and is evaluated as part of the empirical work. Because step 4 of the simulation accounts for the distribution of payrolls across states, multistate policies can be included in the estimation sample for this specification.

2.5. Natural Experiments

This section provides a brief overview on two of the natural experiments we tried to use in order to evaluate the relationship between firm behavior and experience rating. Unlike the projected EMOD and simulated EMOD/premium approaches, do not use EMOD and premium derivatives. We do not supply results associated with these strategies in this report as most of our latter energy focused on the projected and simulated EMOD strategies. Nevertheless we include this section which briefly describe these strategies and the sources of their limitations in order to demonstrate our efforts and highlight challenges presented by the nature of the data.

Recall from Section 1 that eligibility for experience rating is based on historical premiums. A firm is only experience rated if the firm's historical premiums or average premium over the experience period exceed a specific threshold. Thus, firms immediately above the threshold are experience rated while those immediately below the threshold are not. While this sounds like the ideal setting for a regression discontinuity design, data limitations prevent us from implementing this strategy.

In our data we are only able to observe historical premiums for experience rated firms. That is to say the data allow us to see how close firms eligible for experience rating are to the threshold, but not how close ineligible firms are. Our attempts to use current premiums to predict historical premiums to skirt this showed current premiums are a poor substitute for predicting eligibility. Furthermore, conversations with AIG indicate some of the ineligible firms in our sample are ineligible for reasons unrelated to the threshold (e.g. they are too new or spun off a larger company); data limitations prevent us from knowing why an ineligible firm is ineligible. For these reasons, this regression discontinuity was abandoned.

There are a number of policy changes that occur during the sample period which would allow a difference-in-differences strategy to evaluate the relationship between firm behavior and experience rating. Most notably, in 2013 a parameter of the EMOD known as the split point was increased from $5,000 to $10,000. A change which affected all NCCI rates states (except Texas). More details on the split point are offered in the appendix, but in short, a higher split point implies that claims over $5,000 will lead to a larger increase in a firm's EMOD than they would have before the change. If experience rating affects firm behavior as hypothesized, the split point change should be associated with a reduction in the rate of subsequent claims for affected firms. It is important to note that firms which had a claim above $5,000 in their experience rating period at the time of the split point change were immediately effected by the change based on a claim prior to the change, thereby having a sudden increase in their EMOD.

Results associated with our difference-in-differences strategies were mixed. We confirmed the split point change resulted in an increased EMOD for firms with a claim above $5,000, but the effects on WCs were not robust. Investigating this further revealed the composition of AIG's policies during this time period changed greatly. Discussing this with individuals at AIG indicated this was likely associated with changes in AIG's business strategies and practices during the period. Furthermore, other policy changes coincided with the split point change casting further doubt on this approach. For these reasons we moved focused on the projected EMOD and simulated EMOD identification strategies.

3. Technical Details on Sample Construction and Implementation of Methods

3.1. Sample construction

From the raw data provided by AIG, we identified 465,180 policies from 2010 to 2015 associated with experience rated firms and non-experience rated firms. After data cleaning our master analytic sample of policies associated with experience rated firms consisted of 73,332 policies. These were policies for experience rated firms that were consistent among all the raw data sets, not associated with Professional Employer Organizations (PEOs), yearlong which were also not updated during the course of the policy, and have expected losses under $750,000[3]. From here there are two different approaches to constructing the analytic sample for our strategies. We refer to them as the "more restrictive" and "less restrictive" approach.

To understand the distinction between the approaches, and why it is potentially important, it is important to think about how the data required to calculate the derivative interact with the structure of our files. Recall that the denominator of the $t + 2$ EMOD (and thus its derivative) is calculated by payrolls (broken down by class code and state) from $t - 2$, $t - 1$, and t. Thus, when calculating the derivative applicable to the firm's behavior in year t, we need to link the year t policy to payrolls from $t - 2$, $t - 1$, and t. However, the policy records for year t instead contain payrolls for $t - 4$ through $t - 2$ (i.e. the payrolls relevant to the year t EMOD, not the $t + 2$ EMOD). The payrolls that we need to calculate the derivative applicable to year t policies would be contained in the policy record for year $t + 2$.

For this reason, to calculate the derivative applicable to a year t policy, we need to link the year t policy to the policy records for $t + 2$. This implies that, strictly speaking, we can calculate the derivative applicable to a year t policy if and only if the same firm's $t + 2$ policy is also observed in the master sample.

With this in mind, the more restrictive sample construction approach incorporates the following steps[4]:

1. Restrict to policies whose experience period applicable to year t ends in $t - 2$; this ensures that, for all policies, a claim in year t will first affect the EMOD in $t + 2$ (n = 72,685)
2. Drop any firms with policies that are effective on different dates in the same year, which avoids complications with linking to the appropriate policy record (n = 72,088)
3. Keep policies for which there is an available policy record for $t + 2$, for reasons described above (n = 18,136)
4. Drop merit-rated policies, as these do not use the full EMOD formula and the formula for the derivative is therefore not applicable (n = 18,105)
5. Drop policies that have class codes that are rated on a per capita basis (n = 17,803)
6. Drop policies with certain premium-related inconsistencies (n = 17,444)
7. Drop NCCI policies that have a 0 EMOD in some states, as we do not know what causes this (n = 17,396)

Note that, because our master sample includes policy years 2010 – 2015, step 3 does not allow any policies effective after 2013 to be in the analytic sample.

In the less restrictive sample construction approach, we omit step 3, which allows a much larger analytic sample (n = 69,833). Under this approach, instead of linking the year t policy to the policy records for $t + 2$ (which contains the data that we need to precisely calculate the derivative), we calculate the derivative

[3] Various other restrictions were imposed, but those listed are the most significant.

[4] Note that this procedure begins with the master sample of 74,007 experience-rated policies, which is itself a subset of the full universe of AIG policies. A detailed description of the construction of the master sample is available in the memo dated February 26, 2018, entitled "Updated Sample Selection."

using the policy records for year *t* (which contains the historical data relevant to the year *t* EMOD rather than the *t + 2* EMOD). In order to assume that this provides a reasonable proxy for the applicable derivative, we need to assume that payrolls (and their state/class distribution) from *t – 4* through *t – 2* are the same as those from *t – 2* through *t*.

3.2. Calculating derivative of EMOD

Parameter effective year: The *t + 2* EMOD is calculated using state parameters (weights, ballast values, and ELRs) effective in year *t + 2*. Therefore, the most accurate calculation of the derivative would be obtained by applying the parameters effective in *t + 2* to Equations 3 through 5. However, recall that we are assuming that firms are projecting the derivative from the perspective of year *t*; therefore, to calculate the derivatives using the *t + 2* parameters would require an assumption of perfect foresight as to the parameters that would apply to the firm in the future. As this assumption is likely unrealistic, we calculate the derivative of the *t + 2* EMOD with respect to year *t* claims using the parameters in effect as of year *t*. In essence, we assume that firms use the parameters in effect as of year *t* as a proxy for the parameters that will be in effect as of year *t + 2*.

Claim types: Equations 3 through 5 imply that the derivative of the EMOD with respect to the marginal claim dollar depends on two characteristics of that marginal dollar: (1) whether it is from a lost-time or medical-only claim; and (2) whether it is above or below the split point.[5] To account for this, we calculate four versions of the derivative for each policy:

- *delta1*: assuming that the marginal dollar is from a lost-time claim and is below the split point
- *delta2*: assuming that the marginal dollar is from a medical-only claim and is below the split point
- *delta3*: assuming that the marginal dollar is from a lost-time claim and is above the split point
- *delta4*: assuming that the marginal dollar is from a medical-only claim and is above the split point

We can run separate regressions using each of these four versions of the derivative. We also calculate a weighted average of these four versions:

$$(12)\ delta_5 = \sum_{i=1}^{4} w_i delta_i$$

The weights are calculated by dividing policies into cells defined by state, 20 expected loss quantiles, and an indicator for policy years after 2012[6]; and, within each cell, calculating the proportion of claims that fall into each of the four categories. Each *delta* is weighted by the corresponding proportion from that policy's cell.

Multistate policies: For multistate policies, *delta2* and *delta4* (and therefore *delta5*) can vary depending on the state in which the claim occurs.[7] This is because, while most states observe the medical-only adjustment, some do not. For multistate policies, we account for this by calculating the derivatives

[5] As noted above, this ignores the state accident limitation.

[6] This accounts for the change in the split point that occurred in 2013 for most states.

[7] *delta1* and *delta3*, in contrast, do not depend on the state in which the claim occurs because they reflect lost-time claims, for which m_{st} in Equation 3 is 1 regardless of the location of the claim. Therefore, conditional on being above or below the split point, the derivative with respect to a lost-time claim dollar depends only on expected losses, ballast, and the credibility weight. For multistate policies, these parameters are determined by distribution of payrolls across states, not by the state in which the claim occurs.

separately for each state. We then take a weighted average across states, where the weights are equal to each state's share of the policy's total expected losses.

EMOD cap: A final detail relevant to the calculation of $\frac{\partial M}{\partial C}$ concerns the handling of the cap. As noted in Section 1, if, as of year *t*, the firm has already incurred sufficient losses to place its *t + 2* EMOD at the cap, then the derivative with respect to an additional dollar of claims is 0.

In the projected EMOD strategy, in which we use the actual policy-specific value of the derivative as the independent variable, it is relatively straightforward to identify policies whose *t + 2* EMOD would be affected by the cap on the basis of *t – 2* and *t – 1* losses alone. For any such policies, the calculated derivatives are set to 0.[8]

For the simulated EMOD strategy, accounting for the cap is less straightforward. By replacing actual with simulated payrolls, we change each firm's expected losses, meaning that we change the denominator of the EMOD. Based on prior experimentation, if actual losses in *t – 2* and *t – 1* are left "as is," this seems to substantially increase the percentage of policies at the cap and therefore the percentage of policies for which the simulated derivative is set to 0. This suggests that, if we want to account for the cap in the simulated EMOD strategy, we need a way to adjust actual losses to correspond to the adjustment of expected losses. While we developed a way to implement such an adjustment, we decided against incorporating the cap in this analysis. Our rational is twofold. First, from a substantive perspective, the simulated derivatives are meant to vary only by exogenous factors (i.e. the state-determined parameters), whereas the fact of being at the cap is by nature a firm-specific characteristic. Second, since the cap affects relatively few policies and the simulated derivative is only meant to be an instrument for the actual derivative, accounting for the cap in the simulated derivative should be unnecessary; the simulated derivative is still likely to be a strong instrument for the actual derivative even if the former does not incorporate the cap.

3.3. Calculating derivative of premium

Manual premium: As is clear from Equation 9, the manual premium π_m (the first step in the premium calculation process) enters into the formula for the derivative of the premium with respect to the marginal claim dollar. Recall that the manual premium is calculated similarly to expected losses; however, whereas expected losses are calculated by applying ELRs to historical payrolls (i.e. payrolls summed over years *t – 4* to *t - 2*), the manual premium is calculated by applying manual rates to current (i.e. year *t*) payrolls. A stylized expression of this difference is given by:

$$(13)\ E = f(R_0)$$

$$(14)\ \pi_m = g(R_1)$$

where R_0 represents the historical distribution of payrolls, R_1 represents the current distribution of payrolls, *f* represents the function that translates historical payrolls into expected losses, and *g* represents the function that translates current payrolls into manual premiums.

[8] Pennsylvania is unique in that it imposes a minimum as well as a maximum EMOD. In particular, the EMOD is not permitted to change more than 25% in either direction from the prior year. For Pennsylvania policies, we set the derivative to 0 if the *t + 2* EMOD on the basis of *t – 2* and *t – 1* losses is either above the maximum or below the minimum. This implicitly assumes that, if a policy is below the minimum, the marginal claim dollar would not put it above the minimum.

While we have access to the actual value of the manual premium for each policy in each year that is not the value we use for π_m when calculating the derivative. Rather, we recalculate the manual premium (call this π_m^*) by applying manual rates to the historical payrolls used to calculate expected losses. Using the stylized notation from Equations 13 and 14:

$$(15)\ \pi_m^* = g(R_0)$$

The primary reason for this is that the payrolls used to calculate expected losses (R_0, using this stylized notation) represent payrolls summed over all years of the firm's experience period (usually three years). This is useful because a claim occurring in year *t* will affect the EMOD not just in *t + 2*, but also in all subsequent years in which it remains in the experience period (typically, *t + 3* and *t + 4*). Therefore, π_m^* provides an approximation of the total manual premium over all of the years in which the marginal claim will affect the EMOD. This, in turn, means that the resulting estimate of $\frac{\partial \pi_f}{\partial C}$ should better reflect the *total* cost of the marginal claim to the firm (ignoring discounting). For this reason, we primarily use π_m^* to translate the derivative of the EMOD into the derivative of the final premium, although we do also calculate a version based on π_m that is used as a robustness check.

Premium calculation model: As discussed in Section 1, the generalized formula for calculating the final premium is given by Equation 7, which implies the derivative given by Equation 9. However, a challenge in implementing these formulas is that our data files do not allow us to observe all of the individual adjustment factors (besides the EMOD and schedule modification factor) that enter into the premium. We do observe each major step of the process (i.e. the manual premium, subject premium, modified premium, standard premium, and gross premium); and, therefore, we observe the difference between each step. However, the difference between two steps may incorporate a number of individual adjustments, some of which are additive and others of which are multiplicative. Therefore, we do not what share of the difference between two steps is attributable to multiplicative adjustments (and therefore should be modeled as a constant percentage of the prior step) and what share is attributable to additive adjustments (and therefore should be modeled as a constant dollar charge).

In practice, therefore, we do not have sufficient data to implement the generalized formulas given in Equations 7 and 9. To estimate the relevant adjustment factors, and thereby calculate the derivative of the final premium with respect to the marginal claim dollar, we need to make simplifying assumptions about the nature of the adjustments.

In our results, the derivative of the final premium has been calculated using a "multiplicative" model, which assumes that all intermediate adjustments between the major steps are multiplicative in nature. For example, this model assumes that if the modified premium increases by 10%, the standard premium and gross premium will each also increase by 10%. Formally, this implies:

$$(16)\ \pi_f = \pi_m f_1 M S f_2$$

$$(17)\ \frac{\partial \pi_f}{\partial C} = \pi_m f_1 S f_2 \frac{\partial M}{\partial C}$$

In Equations 16 and 17, f_1 represents the adjustments that convert the manual premium into the subject premium; *S* represents the schedule modification factor; and f_2 represents the adjustments that convert the standard premium to the gross premium. Because this model assumes that each step is a constant ratio of the previous step, these can be estimated from variables available in the databases:

- $f_1 = \frac{Subject\ Premium}{Manual\ Premium}$

- $S = \frac{Modified\ Premium + Standard\ Premium}{Modified\ Premium}$
- $f_2 = \frac{Gross\ Premium}{Modified\ Premium + Standard\ Premium}$

We prefer the multiplicative model because, based on the documentation provided by AIG and our discussions with them, it appears that the majority of adjustments are multiplicative, as opposed to additive, in nature; therefore, the multiplicative model is likely the most realistic of the approaches we have used.

When using the derivatives of the premium (either non-simulated or simulated), it is crucial to keep in mind that they reflect simplifying assumptions about the nature of the premium calculation process. Thus, there is a greater risk of measurement error in our estimates of the premium derivative than of the EMOD derivative, implying that results using the premium derivatives are likely to be less reliable. This is an unavoidable byproduct of the fundamental limitations of our available data on premiums.

3.4. Calculating controls

Governing state: For a single-state policy, the governing state is simply the state in which the policy is located. For a multistate policy, the governing state is the state in which the policy has the largest payrolls. We identify the governing state using total historical payrolls.

Governing class code: Analogously, the governing class code is the class code in which the policy has the largest payrolls, again based on the historical payrolls.

Carrier: This is the AIG subsidiary that wrote the policy.

Payrolls: We control for payrolls by including dummies for centiles (100 quantiles). Our payroll control is based on total historical payrolls. Note that, in the regressions with simulated derivatives, the payroll variable used as a control reflects the firm's *actual* historical payrolls, *not* class-level medians.

Firm risk index: The firm risk index is a summary measure of the riskiness of the firm's operations, based on (1) the class codes in which the firm has payrolls and (2) the proportion of the total payroll in each class code. Again, this is based on the actual historical payrolls not the simulated values. To calculate the firm risk index, we first calculate a nationwide ELR for each class code, by averaging the state-specific ELRs. This allows us to calculate the firm risk index as follows:

$$(21)\ RI = \frac{\sum_c R_c l_c^*}{\sum_c R_c}$$

where R_c is the total historical payroll in class c and l_c^* is the average ELR in class c.

Expected losses: As with payrolls, we use dummies for centiles of expected losses. Again, in the simulated derivative regressions, we control for the firm's *actual* expected losses, not the simulated expected losses used in calculating the simulated derivative.

4. Results

4.1. Dependent Variables

All results are reported using the following set of dependent variables:

- **has_current_claim**: indicator for at least one claim in policy year

- **claim_count**: number of claims in policy year
- **claim_count_no0**: number of claims in policy year (excluding 0s)
- **claim_count_log**: logged number of claims in policy year
- **pay_indem_pol**: indemnity payments on claims incurred during policy year[9]
- **pay_indem_pol_no0**: indemnity payments on claims incurred during policy year (excluding 0s)
- **pay_indem_pol_log**: logged indemnity payments on claims incurred during policy year
- **pay_indem_pol_avg**: average indemnity payment per claim incurred during policy year (conditional on having at least one claim)
- **pay_medic_pol, pay_medic_pol_no0, pay_medic_pol_log, pay_medic_pol_avg**: medical payments on claims in policy year, and analogous transformations

Table 1 shows the means of these dependent variables (based on the less restrictive version of the analytic sample). Unless otherwise stated, all results in this report are based on *delta5*, the version of the derivative that is calculated as a weighted average of the derivatives corresponding to each specific claim type.

4.2. Projected EMOD Results

The following are the key takeaways from analyses using this specification:

- Estimates are highly sensitive to controls, as expected
- Estimates for some dependent variables are sensitive to whether the EMOD derivative or premium derivative is used as the independent variable of interest
- Estimates are sensitive to the choice between a more restrictive and a less restrictive approach to defining the analytic sample
- When the less restrictive sample is used, estimates are heavily driven by California policies

All results using this specification exclude multistate policies.

4.2.1. Sensitivity to controls and to choice of derivatives

Table 2 shows regression results using the derivative of the EMOD (Equation 1) and derivative of the final premium (Equation 2) as the independent variables of interest, with two tiers of controls[10]. Because the derivative of the EMOD is extremely small for all policies, the raw regression coefficients are difficult to interpret. Therefore, in this and in subsequent tables, the standardized regression coefficient (interpreted as the effect of a standard deviation increase in the independent variable) is also shown.

[9] Because workers' compensation claims may stay open for variable periods of time, we have been using payments as of 12 months after the injury date.
[10] Implemented on the more restrictive sample.

For all dependent variables, the magnitude of the estimated effect is highly sensitive to controls. This is consistent with expectations, because both derivatives are correlated with firm characteristics (specifically, firm size and riskiness) that are independently correlated with claim frequency.

When the derivative of the EMOD is the independent variable, coefficients are negative in the least rich specification, and move in a positive direction with progressively richer controls. This suggests that, in the absence of controls, the coefficient on the EMOD derivative is negatively biased; this is likely driven by the fact that expected losses are in the denominator of the derivative. In the richest specification (with controls for state, effective year and month, governing class code, carrier, payroll, the firm risk index, and expected losses), we observe a significantly negative effect on claim counts (though not robust to the exclusion of 0s or to logging); a significantly positive effect on the average indemnity payment per claim[11]; and significantly positive effects on various transformations of medical payments. The positive effects on the payment variables are inconsistent with expectations.

When the derivative of the final premium is the independent variable, the opposite pattern is observed: coefficients are positive in the least rich specification, and move in a negative direction with progressively richer controls. This suggests a positive bias on the coefficient of the premium derivative, which likely reflects the fact that manual premiums (which, like expected losses, are driven by firm size and riskiness) are in the numerator of this derivative. In the richest specification, we observe a negative effect on claim counts and a marginally significant positive effect on logged indemnity payments.

Thus, while the estimated effect on claim counts (using the richest set of controls) is roughly consistent for the EMOD and premium derivatives, observed effects on other dependent variables are less robust. The discrepancy is likely attributable to the fact that the derivative of the premium incorporates multiple additional factors (including the manual premium and various adjustment factors) that may vary endogenously in ways that we do not fully understand. Additionally, data limitations imply that simplifying assumptions must be made in order to translate the derivative of the EMOD into a derivative of the final premium, so the latter likely incorporates additional measurement error.

4.2.2. Sensitivity to sample definition

As explained in previous sections, given the structure of our available data, the calculation of the true derivative for a year t policy requires it to be linked to the policy record for the $t + 2$ policy from the same firm. This requires the analytic sample to be restricted to policies for which the $t + 2$ policy exists in our master sample. This requirement restricts the number of policies available for this analysis to 17,396 of the 74,007 policies in our master sample.

If this requirement is relaxed, the available sample for this analysis expands to 69,833 (incorporating several other, less significant restrictions). The tradeoff is that this requires us to calculate the derivative of the $t + 2$ EMOD using payrolls from $t – 4$ through $t – 2$, whereas the derivative is actually based on $t – 2$ through t payrolls. Thus, to use the larger sample with this approach, we must assume that payrolls from $t – 4$ through $t – 2$ are a reasonable proxy for those from $t – 2$ through t.

Table 3 compares OLS results between the more restrictive sample of 17,396 policies and the less restrictive sample of 69,833 policies, using the derivative of the EMOD as the independent variable of

[11] A possible explanation for an increase in average payments per claim is that, when the EMOD is more sensitive to the marginal claim, firms may pay for smaller injuries out of pocket, thus increasing the average size of claims that are reported to the insurer and thus observed in our data. This channel is posited in an empirical study by Neuhauser et al. 2013 ("Increasing Experience Rating for Small Employers: Would Lowering the Threshold for Experience Rating Improve Safety").

interest. Table 4 does the same using the derivative of the premium as the independent variable of interest. Both tables use the richest specification including all controls.

In both cases, there are substantive changes in several estimates when the larger sample is used. In particular, note that the sign of the effect on claim counts flips from negative to positive. This suggests that estimates using the less restrictive sample may not be comparable to those using the more restrictive sample. In particular, the assumptions required to use the less restrictive sample may not be valid.

4.2.3 Sensitivity to California policies

An additional effect of using the less restrictive sample is that the OLS results become somewhat sensitive to the removal of California policies. Table 5 compares results from a sample including California to those excluding California, using the derivative of the EMOD as the independent variable and all available controls. Table 6 does the same using the derivative of the premium as the independent variable.

In both cases, the majority of estimates that exclude CA are not within the 95% confidence interval of the estimates that include CA. With the derivative of the EMOD (Table 5), the apparent positive effect on claim counts disappears. The signs of the estimated effects on total medical payments also flip from positive to negative. Similarly, with the derivative of the premium (Table 6), the positive effect on raw claim counts disappears (though it remains when 0s are excluded and when the count is logged), as do the apparent positive effects on several payment-related variables.

In contrast, OLS results using the more restrictive sample do not appear to be as sensitive to CA policies; nearly all estimates without CA policies are within the 95% confidence interval of estimates with CA policies.

Our hypothesis is that the sensitivity to CA is driven by the fact that, after 2011, CA dropped the ballast value from its EMOD formula. Since the purpose of ballast is to reduce between-firm variability in the derivative of the EMOD, this formula change would have increased the variability of derivatives among CA policies, relative to policies from states that continued to use ballast (all except Pennsylvania and Delaware). The more restrictive sample only includes policies from 2010 through 2013, whereas the less restrictive sample includes policies from 2010 through 2015 and therefore includes a larger number of post-change CA policies.

4.3. Simulated Derivative Approach

The specification for this approach is similar to Equations 1 and 2, but uses simulated values of the derivatives that act as instruments for the actual derivatives. Thus, this specification can be understood as a reduced-form instrumental variable (IV) specification. The simulation approach is described in detail in the methodology memo. In brief, however, it entails: (1) replacing every firm's total payroll with the median payroll from its governing class code; (2) assigning the firm's entire simulated payroll to its governing class code and governing state; (3) replacing the various adjustment factors that enter into the final premium with their medians within the governing class code; and (4) re-calculating the derivatives of the EMOD and premium based on these simulated values. The purpose of doing so is to obtain an estimate of the derivative that (after controlling for the governing state, governing class code, effective period, and carrier) varies only due to exogenous parameters.

The takeaways from this strategy are similar to those from the projected EMOD approach:

- Although the simulation was intended to reduce the reliance on parametric controls, some estimates remain sensitive to controls

- Estimates for some dependent variables remain sensitive to whether the simulated EMOD derivative or the simulated premium derivative is used

- To a greater extent than in the projected EMOD approach, results in the simulated derivative approach are almost entirely driven by CA policies; there is very little variation in the simulated derivatives among non-CA policies

Results for the simulated derivative approach (unlike the projected EMOD approach) appear to be relatively insensitive to the choice between the less restrictive and more restrictive samples. This is likely because the assumptions required to use the less restrictive sample are weaker with the simulated derivative approach. Because we are replacing total payrolls with the median by class code, and assigning all simulated payrolls to the governing state and class code, we no longer need to assume that payrolls from *t – 4* through *t – 2* are the same as those from *t – 2* through *t*; we need only assume that the governing state and class code identified based on the *t – 4* through *t – 2* payrolls are identical to those that would be identified based on the *t – 2* through *t* payrolls. For this reason, all results in this section are based on the less restrictive sample.

4.3.1 Sensitivity to controls and to choice of derivatives

Table 7 shows reduced-form regression results using the simulated derivative of the EMOD and the final premium. Contrary to expectations, many results are sensitive to controls.

With the EMOD derivative, the addition of the risk and expected loss controls leads to a positive estimated effect on the probability of having a claim. In contrast, the negative effect on claim counts is relatively robust to controls, though not when the count is logged. Several payment-related variables show the expected negative effects in the least rich specification (including only controls for state, effective period, class, and carrier), but these become indistinguishable from 0 as controls are added.

More substantive sensitivity is observed with the premium derivative. In particular, the effect on claim counts is negative in the least rich specification but positive in the richest specification. Several variables related to medical payments also flip signs from negative to positive as controls are added.

Notably, the EMOD and premium results are generally consistent with each other (at least in terms of sign and significance) in the least rich specification, but diverge as controls are added. In particular, in the richest specification, the estimated effect on claim counts is negative (as expected) for the EMOD derivative but positive for the premium derivative. As noted previously, the greater sensitivity of the premium results may be driven by measurement error introduced by the assumptions we make in translating the EMOD derivative into a premium derivative.

In general, these results suggest that our simulation approach does not fulfill the original goal of reducing the dependence of the estimates on controls. When the simulated derivatives are used with the least rich specification, Equations 1 and 2 are identified off of interactions between the governing state, governing class code, effective period, and carrier. Because we do not directly control for any interactions, the sensitivity to additional controls suggests that there are some interactions that are independently associated with firm characteristics and/or with claim frequency.

4.3.2. Sensitivity to California policies

Aside from the sensitivity to controls, we have also found that the results using the simulated derivatives are driven almost entirely by 2012 – 2015 CA policies.

Table 8 compares results with and without CA policies, using the simulated derivative of the EMOD and the richest specification with all available controls. Table 9 shows the same using the simulated derivative of the premium. All effects (except one) become indistinguishable from 0 when CA policies are dropped.[12] Of particular note is that the standard errors increase by a factor of about 15 for the EMOD derivatives and about 5 for the premium derivatives, which far exceeds the inflation that would be expected from the reduction in the sample size. This suggests that the bulk of the variation in the simulated derivatives is accounted for by CA policies.

Our hypothesis is that this pattern is attributable to the 2012 formula change implemented in CA. Prior exploratory work, not reported here, found that the pattern is largely driven by 2012 – 2015 CA policies; that is, excluding only 2010 – 2011 CA policies leads to results that look similar to those that include all CA policies, while excluding only 2012 – 2015 CA policies leads to results that look similar to those that exclude all CA policies. The most important effect of the 2012 formula change was to drop the ballast from the denominator of the EMOD (and thus from the denominator of the derivative). The ballast component of the formula is explicitly designed to reduce between-firm variation in the sensitivity of the EMOD to a new claim—that is, in the derivative. Thus, the elimination of this component would be expected to increase the variability of the derivative among CA policies. In contrast, in the states that continued to use ballast (all except PA and DE), this component would continue to fulfill its purpose of dampening variability. The heavy influence of 2012 – 2015 CA policies on the simulated derivative results—particularly on the standard errors—suggests that, once endogenous sources of variation are removed, states that use ballast show relatively little exogenous between-firm variation in the derivative.

5. Conclusion

We have attempted to explore the relationship between experience rating and the incidence of Workers Compensation using data provided by AIG. The primary approaches presented in this report are based on a "forward-looking" model of firm behavior which assumes firms have a sufficient understanding of experience rating and the EMOD formula. The hypothesis tested was that firms for which the marginal impact of a claim was higher would adjust their practices in order to reduce claim frequency and/or severity, thereby preventing a potential increase to their insurance premiums. Our analyses were inconclusive.

Observed estimates were highly sensitive to controls and sample composition. Like with the natural experiment strategies employed during our earlier estimation efforts, data limitations hindered our "forward-looking" strategies. For these models, necessary simplifying assumptions regarding the additive and multiplicative intermediate adjustments made while calculating the derivative of premium likely added measurement error, which may be the cause of the seemingly contradictory results between our EMOD derivative and premium derivative analyses. The lack of policy records in year $t+2$ for the majority of policies in our sample also forced us to either use a relatively small sample (more restrictive sample) or make additional assumptions regarding firms' payrolls (less restrictive sample). Ultimately, the assumptions necessary to test our hypothesis are suspect and our highly sensitive results prevent us from providing conclusive evidence for our hypothesis.

[12] While Tables 8 and 9 show the specification that includes all controls, it is worth noting that, with the EMOD derivative, there is little sensitivity to controls when CA policies are excluded. That is, nearly all estimates are non-significant regardless of controls. With the derivative of the premium, however, some estimates (including claim counts) remain sensitive to controls.

Table 1: Summary statistics for dependent variables

Variable	Mean	S.D.
has_current_claim	0.40	0.49
claim_count	1.65	4.43
claim_count_no0	4.13	6.23
claim_count_log	0.84	0.96
pay_indem_pol	$3,343.70	$11,158.37
pay_indem_pol_no0	$15,748.74	$19,775.96
pay_indem_pol_log	$8.74	$1.64
pay_indem_pol_avg	$2,864.56	$6,501.85
pay_medic_pol	$6,270.05	$19,267.11
pay_medic_pol_no0	$17,285.98	$28,862.14
pay_medic_pol_log	$8.49	$1.81
pay_medic_pol_avg	$5,324.34	$11,899.73

Table 2: Estimates from OLD projected EMOD specification , by controls and EMOD vs. premium derivative

		Using Derivative of EMOD				Using Derivative of Premium			
Dependent Variable	Controls	Coef.	S.E.	Standardized Coef.	N	Coef.	S.E.	Standardized Coef.	N
has_current_claim	Initial Controls[a]	-14475.14***	525.02	-0.16	14,965	0.13***	0.01	0.12	14,948
	Rich Set of Controls[b]	96.60	665.03	0.00	14,965	0.00	0.01	0.00	14,948
claim_count	Initial Controls	-109171.93***	3344.64	-1.21	14,929	0.46***	0.05	0.43	14,912
	Rich Set of Controls	-8104.32	3569.99	-0.09	14,929	-0.09**	0.05	-0.09	14,912
claim_count_no0	Initial Controls	-205897.81***	10255.38	-1.91	4,851	0.48***	0.13	0.51	4,851
	Rich Set of Controls	-10004.18	12631.68	-0.09	4,851	-0.17	0.11	-0.19	4,851
claim_count_log	Initial Controls	-42315.82***	1668.13	-0.39	4,851	0.18***	0.02	0.19	4,851
	Rich Set of Controls	-1159.73	2110.20	-0.01	4,851	0.02	0.02	0.02	4,851
pay_indem_pol	Initial Controls	-$176,337,976.57***	$8,521,591.36	-$1,957.99	14,949	$880.77***	$131.75	$822.08	14,932
	Rich Set of Controls	-$15,115,198.38	$10,440,369.82	-$167.83	14,949	$2.32	$134.02	$2.16	14,932
pay_indem_pol_no0	Initial Controls	-$497,423,295.46***	$67,885,779.94	-$4,249.93	2,095	$2,105.42***	$721.49	$2,341.95	2,095
	Rich Set of Controls	-$38,914,118.01	$121,425,359.63	-$332.48	2,095	$1,089.38	$783.40	$1,211.76	2,095
pay_indem_pol_log	Initial Controls	-$45,955.04***	$6,648.34	-$0.39	2,095	$0.27***	$0.07	$0.30	2,095
	Rich Set of Controls	-$2,137.21	$11,808.94	-$0.02	2,095	$0.14*	$0.08	$0.16	2,095
pay_indem_pol_avg	Initial Controls	-$8,565,296.97	$12,584,779.82	-$79.29	4,841	$173.44	$150.99	$187.44	4,841
	Rich Set of Controls	$46,134,935.97**	$18,292,198.05	$427.07	4,841	$38.80	$163.86	$41.94	4,841
pay_medic_pol	Initial Controls	-$301,403,340.20***	$13,369,535.13	-$3,346.15	14,937	$1,705.78***	$207.02	$1,592.51	14,920
	Rich Set of Controls	$33,516,825.82**	$15,902,465.08	$372.10	14,937	-$116.57	$203.88	-$108.83	14,920
pay_medic_pol_no0	Initial Controls	-$602,857,497.10***	$50,031,788.09	-$5,550.68	4,286	$1,999.98***	$615.07	$2,135.74	4,286
	Rich Set of Controls	$116,683,936.29*	$69,371,281.76	$1,074.34	4,286	-$516.23	$617.12	-$551.28	4,286
pay_medic_pol_log	Initial Controls	-$58,212.64***	$4,278.14	-$0.54	4,286	$0.29***	$0.05	$0.31	4,286
	Rich Set of Controls	$12,532.08**	$6,023.23	$0.12	4,286	$0.00	$0.05	$0.00	4,286
pay_medic_pol_avg	Initial Controls	-$15,438,566.16	$21,593,916.08	-$142.81	4,831	$294.15	$258.82	$317.98	4,831
	Rich Set of Controls	$46,717,908.28	$31,700,216.23	$432.14	4,831	$7.79	$283.09	$8.42	4,831

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Regressions are estimated on the more restrictive version of the analytic sample. Standardized coefficient shows the effect of a standard deviation increase in the independent variable. Multistate policies are excluded.

a - Initial set of controls include state, effective year/month, class, and carrier controls.

b - Rich Set of controls include all initial controls and quantiles for payroll (100), firm risk (25), and expected loss (100).

Table 3: Estimates from OLS Projected EMOD Specification (derivative of EMOD), by sample definition

	Using More Restrictive Sample				Using Less Restrictive Sample			
Dependent Variable	**Coef.**	**S.E.**	**Standardized Coef.**	**N**	**Coef.**	**S.E.**	**Standardized Coef.**	**N**
has_current_claim	96.60	665.03	0.00	14,965	-603.44**	273.93	-0.01	59,641
claim_count	-8104.32**	3569.99	-0.09	14,929	6469.43***	1885.92	0.08	59,476
claim_count_no0	-10004.18	12631.68	-0.09	4,851	23935.66***	5400.67	0.25	22,388
claim_count_log	-1159.73	2110.20	-0.01	4,851	2957.27***	839.63	0.03	22,388
pay_indem_pol	-$15,115,198.38	$10,440,369.82	-$167.83	14,949	-$15,738,840.35***	$5,767,100.85	-$192.42	59,553
pay_indem_pol_no0	-$38,914,118.01	$121,425,359.63	-$332.48	2,095	-$24,059,345.96	$32,014,989.20	-$240.10	11,680
pay_indem_pol_log	-$2,137.21	$11,808.94	-$0.02	2,095	-$2,181.99	$2,928.81	-$0.02	11,680
pay_indem_pol_avg	$46,134,935.97**	$18,292,198.05	$427.07	4,841	-$5,616,827.84	$7,883,789.29	-$58.38	22,329
pay_medic_pol	$33,516,825.82**	$15,902,465.08	$372.10	14,937	$42,639,800.13***	$9,835,697.85	$521.29	59,490
pay_medic_pol_no0	$116,683,936.29*	$69,371,281.76	$1,074.34	4,286	$87,278,037.17***	$32,446,426.21	$898.52	20,322
pay_medic_pol_log	$12,532.08**	$6,023.23	$0.12	4,286	-$23.34	$2,199.47	$0.00	20,318
pay_medic_pol_avg	$46,717,908.28	$31,700,216.23	$432.14	4,831	-$27,293,066.52*	$15,099,901.31	-$283.59	22,267

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Results are shown for the richest specification incorporating controls for state, carrier, class, effective year/month, payrolls, firm risk, and expected losses. Standardized coefficient shows the effect of a standard deviation increase in the independent variable. Multistate policies are excluded.

Table 4: Estimates from OLS Projected EMOD Specification (derivative of premium), by sample definition

	Using More Restrictive Sample				Using Less Restrictive Sample			
Dependent Variable	**Coef.**	**S.E.**	**Standardized Coef.**	**N**	**Coef.**	**S.E.**	**Standardized Coef.**	**N**
has_current_claim	0.00	0.01	0.00	14,948	0.00	0.00	0.00	59,260
claim_count	-0.09**	0.05	-0.09	14,912	0.02***	0.01	0.04	59,095
claim_count_no0	-0.17	0.11	-0.19	4,851	0.27***	0.04	0.35	22,382
claim_count_log	0.02	0.02	0.02	4,851	0.04***	0.01	0.05	22,382
pay_indem_pol	$2.32	$134.02	$2.16	14,932	$35.52	$23.49	$65.95	59,172
pay_indem_pol_no0	$1,089.38	$783.40	$1,211.76	2,095	$708.98***	$249.95	$895.47	11,677
pay_indem_pol_log	$0.14*	$0.08	$0.16	2,095	$0.04*	$0.02	$0.05	11,677
pay_indem_pol_avg	$38.80	$163.86	$41.94	4,841	$38.10	$61.93	$48.94	22,323
pay_medic_pol	-$116.57	$203.88	-$108.83	14,920	$71.38*	$40.02	$132.74	59,109
pay_medic_pol_no0	-$516.23	$617.12	-$551.28	4,286	$908.88***	$246.96	$1,174.54	20,319
pay_medic_pol_log	$0.00	$0.05	$0.00	4,286	$0.03*	$0.02	$0.04	20,315
pay_medic_pol_avg	$7.79	$283.09	$8.42	4,831	$17.90	$117.47	$23.11	22,261

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Results are shown for the richest specification incorporating controls for state, carrier, class, effective year/month, payrolls, firm risk, and expected losses. Standardized coefficient shows the effect of a standard deviation increase in the independent variable. Multistate policies are excluded.

Table 5: Estimates from OLS Projected EMOD Specification (derivative of EMOD), including vs. excluding CA

	Including CA				Excluding CA			
Dependent Variable	**Coef.**	**S.E.**	**Standardized Coef.**	**N**	**Coef.**	**S.E.**	**Standardized Coef.**	**N**
has_current_claim	-603.44**	273.93	-0.01	59,641	-1055.57***	330.04	-0.01	50,042
claim_count	6469.43***	1885.92	0.08	59,476	-2178.66	1953.99	-0.02	49,925
claim_count_no0	23935.66***	5400.67	0.25	22,388	619.68	6138.49	0.01	17,589
claim_count_log	2957.27***	839.63	0.03	22,388	-112.30	1030.94	0.00	17,589
pay_indem_pol	-$15,738,840.35***	$5,767,100.85	-$192.42	59,553	-$32,819,746.93***	$6,214,597.73	-$339.30	49,984
pay_indem_pol_no0	-$24,059,345.96	$32,014,989.20	-$240.10	11,680	-$150,143,235.68***	$38,582,714.51	-$1,230.99	8,835
pay_indem_pol_log	-$2,181.99	$2,928.81	-$0.02	11,680	-$18,331.79***	$3,752.25	-$0.15	8,835
pay_indem_pol_avg	-$5,616,827.84	$7,883,789.29	-$58.38	22,329	-$18,088,109.77*	$10,366,072.33	-$153.18	17,553
pay_medic_pol	$42,639,800.13***	$9,835,697.85	$521.29	59,490	-$21,342,467.64*	$11,590,379.58	-$220.60	49,918
pay_medic_pol_no0	$87,278,037.17***	$32,446,426.21	$898.52	20,322	-$56,846,009.13	$41,816,130.12	-$482.32	15,890
pay_medic_pol_log	-$23.34	$2,199.47	$0.00	20,318	-$8,379.95***	$2,861.09	-$0.07	15,886
pay_medic_pol_avg	-$27,293,066.52*	$15,099,901.31	-$283.59	22,267	-$39,571,113.97*	$21,066,446.69	-$334.59	17,488

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Results are shown for the richest specification incorporating controls for state, carrier, class, effective year/month, payrolls, firm risk, and expected losses. Standardized coefficient shows the effect of a standard deviation increase in the independent variable. Multistate policies are excluded. Less restricted sample used.

Table 6: Estimates from OLS Project EMOD Specification (derivative of premium), including vs. excluding CA

	Including CA				Excluding CA			
Dependent Variable	**Coef.**	**S.E.**	**Standardized Coef.**	**N**	**Coef.**	**S.E.**	**Standardized Coef.**	**N**
has_current_claim	0.00	0.00	0.00	59,260	0.00*	0.00	0.00	49,714
claim_count	0.02***	0.01	0.04	59,095	0.00	0.01	0.00	49,597
claim_count_no0	0.27***	0.04	0.35	22,382	0.36***	0.09	0.20	17,585
claim_count_log	0.04***	0.01	0.05	22,382	0.04***	0.02	0.02	17,585
pay_indem_pol	$35.52	$23.49	$65.95	59,172	$4.33	$22.96	$7.04	49,656
pay_indem_pol_no0	$708.98***	$249.95	$895.47	11,677	$798.31	$538.58	$442.86	8,833
pay_indem_pol_log	$0.04*	$0.02	$0.05	11,677	$0.01	$0.05	$0.00	8,833
pay_indem_pol_avg	$38.10	$61.93	$48.94	22,323	-$67.33	$155.30	-$36.04	17,549
pay_medic_pol	$71.38*	$40.02	$132.74	59,109	-$28.96	$42.79	-$47.08	49,590
pay_medic_pol_no0	$908.88***	$246.96	$1,174.54	20,319	-$405.75	$617.00	-$220.41	15,888
pay_medic_pol_log	$0.03*	$0.02	$0.04	20,315	-$0.01	$0.04	-$0.01	15,884
pay_medic_pol_avg	$17.90	$117.47	$23.11	22,261	-$271.85	$314.80	-$145.51	17,484

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Results are shown for the less restrictive sample, and for the richest specification incorporating controls for state, carrier, class, effective year/month, payrolls, firm risk, and expected losses. Standardized coefficient shows the effect of a standard deviation increase in the independent variable. Multistate policies are excluded.

		Using Derivative of EMOD				Using Derivative of Premium			
Dependent Variable	Controls	Coef.	S.E.	Standardized Coef.	N	Coef.	S.E.	Standardized Coef.	N
has_current_claim	Initial Controls[a]	-3.37	9.31	0.00	69,833	0.00**	0.00	-0.01	66,068
	Rich Set of Controls[b]	27.02***	8.56	0.02	69,823	0.01***	0.00	0.01	66,059
claim_count	Initial Controls	-421.89***	81.57	-0.26	69,560	-0.10***	0.02	-0.12	65,809
	Rich Set of Controls	-224.22***	67.09	-0.14	69,550	0.03*	0.02	0.04	65,800
claim_count_no0	Initial Controls	-809.97***	158.79	-0.76	27,727	-0.17***	0.04	-0.23	26,460
	Rich Set of Controls	-340.79***	133.06	-0.32	27,722	0.07*	0.04	0.09	26,455
claim_count_log	Initial Controls	-117.34***	23.89	-0.11	27,727	-0.03***	0.01	-0.04	26,460
	Rich Set of Controls	-27.32	19.52	-0.03	27,722	0.02***	0.01	0.02	26,455
pay_indem_pol	Initial Controls	-$749,062.28***	$216,764.48	-$460.09	69,684	-$259.08***	$56.46	-$304.17	65,926
	Rich Set of Controls	$70,214.01	$199,862.72	$43.13	69,674	$40.70	$52.05	$47.67	65,917
pay_indem_pol_no0	Initial Controls	-$777,312.95	$737,689.94	-$728.59	14,795	-$532.13***	$201.99	-$725.38	14,104
	Rich Set of Controls	$753,581.32	$709,475.06	$706.41	14,792	-$27.87	$193.96	-$37.99	14,101
pay_indem_pol_log	Initial Controls	-$48.21	$61.30	-$0.05	14,794	-$0.02	$0.02	-$0.02	14,103
	Rich Set of Controls	$62.06	$60.07	$0.06	14,791	$0.02	$0.02	$0.03	14,100
pay_indem_pol_avg	Initial Controls	$44,192.57	$176,515.54	$41.74	27,620	-$21.30	$48.33	-$28.19	26,359
	Rich Set of Controls	$140,173.58	$177,470.82	$132.41	27,615	-$33.05	$48.68	-$43.53	26,354
pay_medic_pol	Initial Controls	-$1,123,274.24***	$374,154.17	-$690.50	69,579	-$334.24***	$97.39	-$393.21	65,827
	Rich Set of Controls	$151,235.25	$340,456.28	$92.97	69,569	$199.51**	$88.54	$234.09	65,818
pay_medic_pol_no0	Initial Controls	-$2,104,269.07**	$819,611.38	-$1,565.36	25,238	-$594.29***	$227.28	-$793.00	24,111
	Rich Set of Controls	$340,001.29	$767,043.09	$252.95	25,233	$268.59	$213.07	$356.49	24,106
pay_medic_pol_log	Initial Controls	-$126.32**	$51.27	-$0.09	25,233	-$0.03**	$0.01	-$0.04	24,106
	Rich Set of Controls	$21.45	$48.34	$0.02	25,228	$0.02*	$0.01	$0.03	24,101
pay_medic_pol_avg	Initial Controls	$256,826.91	$327,028.37	$242.99	27,531	$63.77	$89.93	$84.72	26,274
	Rich Set of Controls	$314,333.44	$328,972.62	$297.42	27,526	$33.02	$90.68	$43.65	26,269

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Regressions are estimated on the less restrictive version of the analytic sample. Standardized coefficient shows the effect of a standard deviation increase in the independent variable.

a - Initial set of controls include state, effective year/month, class, and carrier controls.

b – The rich set of controls include all initial controls and quantiles for payroll (100), firm risk (25), and expected loss (100).

Table 8: Estimates from Reduced-Form Simulated Derivative Specification (derivative of EMOD), including vs. excluding CA

	Including CA				Excluding CA			
Dependent Variable	**Coef.**	**S.E.**	**Standardized Coef.**	**N**	**Coef.**	**S.E.**	**Standardized Coef.**	**N**
has_current_claim	27.02***	8.56	0.02	69,823	-133.27	124.19	-0.01	60,224
claim_count	-224.22***	67.09	-0.14	69,550	-415.00	915.65	-0.02	59,999
claim_count_no0	-340.79**	133.06	-0.32	27,722	-1267.52	1873.98	-0.06	22,923
claim_count_log	-27.32	19.52	-0.03	27,722	58.78	281.44	0.00	22,923
pay_indem_pol	$70,214.01	$199,862.72	$43.13	69,674	-$533,282.44	$2,745,689.92	-$21.38	60,105
pay_indem_pol_no0	$753,581.32	$709,475.06	$706.41	14,792	$6,414,967.90	$11,081,837.08	$260.82	11,947
pay_indem_pol_log	$62.06	$60.07	$0.06	14,791	-$130.88	$954.82	-$0.01	11,946
pay_indem_pol_avg	$140,173.58	$177,470.82	$132.41	27,615	$1,391,235.81	$2,658,944.56	$64.17	22,839
pay_medic_pol	$151,235.25	$340,456.28	$92.97	69,569	-$648,252.77	$4,954,576.35	-$25.97	59,997
pay_medic_pol_no0	$340,001.29	$767,043.09	$252.95	25,233	$4,911,562.86	$11,429,519.30	$226.90	20,801
pay_medic_pol_log	$21.45	$48.34	$0.02	25,228	$893.88	$715.64	$0.04	20,796
pay_medic_pol_avg	$314,333.44	$328,972.62	$297.42	27,526	$3,143,616.09	$5,153,734.34	$144.85	22,747

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. NOTE: Results are shown for the less restrictive sample, and for the richest specification incorporating controls for state, carrier, class, effective year/month, payrolls, firm risk, and expected losses. Standardized coefficient shows the effect of a standard deviation increase in the independent variable.

Table 9: Estimates from Reduced-Form Simulated Derivative Specification (derivative of premium), including vs. excluding CA

	Including CA				Excluding CA			
Dependent Variable	**Coef.**	**S.E.**	**Standardized Coef.**	**N**	**Coef.**	**S.E.**	**Standardized Coef.**	**N**
has_current_claim	0.01***	0.00	0.01	66,059	-0.01	0.01	0.00	56,724
claim_count	0.03*	0.02	0.04	65,800	0.02	0.07	0.01	56,513
claim_count_no0	0.07*	0.04	0.09	26,455	0.30	0.20	0.10	21,756
claim_count_log	0.02***	0.01	0.02	26,455	0.06*	0.03	0.02	21,756
pay_indem_pol	$40.70	$52.05	$47.67	65,917	-$24.59	$204.94	-$8.40	56,612
pay_indem_pol_no0	-$27.87	$193.96	-$37.99	14,101	-$841.20	$1,033.47	-$290.09	11,315
pay_indem_pol_log	$0.02	$0.02	$0.03	14,100	-$0.08	$0.09	-$0.03	11,314
pay_indem_pol_avg	-$33.05	$48.68	-$43.53	26,354	-$458.17	$283.10	-$151.67	21,678
pay_medic_pol	$199.51**	$88.54	$234.09	65,818	-$131.65	$369.27	-$45.01	56,509
pay_medic_pol_no0	$268.59	$213.07	$356.49	24,106	-$52.94	$1,217.92	-$17.61	19,758
pay_medic_pol_log	$0.02*	$0.01	$0.03	24,101	-$0.04	$0.08	-$0.01	19,753
pay_medic_pol_avg	$33.02	$90.68	$43.65	26,269	-$319.92	$551.15	-$106.03	21,589

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$. Results are shown for the less restrictive sample, and for the richest specification incorporating controls for state, carrier, class, effective year/month, payrolls, firm risk, and expected losses. Standardized coefficient shows the effect of a standard deviation increase in the independent variable.

Appendix

NCCI formula

A number of parameters, which vary by state, year, and firm characteristics, enter into the EMOD. Most of our identification strategies rely on some type of exogenous variation in these parameters.

The best way to understand all of the parameters is to walk through the formula piece by piece. The formulas shown here apply to a policy effective in year *t* with an experience period from *t – 4* through *t – 2*. *h* indexes years in the experience period, *c* indexes employee class codes, and *s* indexes states.

The formula used by NCCI, as well as the independent bureaus in Michigan, Massachusetts, North Carolina, Wisconsin, and New York, is given by Equation 1:

$$(1)\ M_t = \frac{\sum_{h=t-4}^{t-2} A_{p,h} + w_{st}(E_t) \sum_{h=t-4}^{t-2} A_{e,h} + b_{st}(E_t) + [1 - w_{st}(E_t)]E_{e,t}}{E_t + b_{st}(E_t)}$$

In Equation 1, *A* denotes actual losses. Actual losses are divided into primary losses (A_p) and excess losses (A_e). The division is based on a state-determined parameter called the **split point**. Suppose that a $7,000 claim occurs in a state with a split point of $5,000. The first $5,000 of the claim is classified as primary losses and the remaining $2,000 is classified as excess losses. It is crucial to remember that the split point is applied at the claim level, not the aggregate level. Again assuming a split point of $5,000, a firm with five claims of $2,000 each would have primary losses of $10,000 and no excess losses; while a firm with one $7,000 claim and one $3,000 claim would have primary losses of $8,000 and excess losses of $2,000. It is also important to remember that the applicable split point is that in effect in the state in which the claim occurred, and the year *t* for which the EMOD is being calculated (*not* the historical year in which the claim took place). The rationale for splitting claims into primary and excess losses is discussed in more detail below.

E denotes expected losses. Expected losses are calculated as follows:

$$(2)\ E_t = \sum_{h=t-4}^{t-2} \sum_{s} \sum_{c} R_{csh} l_{cst}$$

where R_{csh} is the payroll in class code *c*, state *s*, and historical year *h*; and l_{cst} is the **expected loss rate (ELR)** applied to class code *c* by state *s* in year *t*. The ELR is a state-determined regulatory parameter that is typically updated each year. Note that, while historical (not current) payrolls are used to calculate expected losses, the applicable ELR is the one promulgated in year *t* (the policy year for which the EMOD is being calculated), *not* the one in effect during the experience period. In other words, the current ELR is applied to the sum of historical payrolls over the experience period. For multistate policies, the applicable ELR is that promulgated by the state in which the payroll is located.

As with actual losses, expected losses are also divided into primary and excess components. Expected excess losses appear in Equation 1 and are calculated as follows:

$$(3)\ E_{e,t} = \sum_{h=t-4}^{t-2} \sum_{s} \sum_{c} R_{csh} l_{cst} (1 - d_{cst})$$

where d_{cst} is another state regulatory parameter called the **d-ratio**. The d-ratio is the proportion of losses within a class-state cell that are expected to be primary losses. Again, the applicable d-ratio is the one promulgated in year *t* by the state in which the payroll is located.

Returning to Equation 1, $w_{st}(E)$ is the **credibility weight**, which ranges from 0 to 1. Therefore, the credibility weight discounts the contribution of excess losses to the EMOD. The rationale for this is that firms are assumed to have more control over claim frequency than severity. That is, the firm can affect the probability that a claim occurs by implementing various safety procedures; but, once a claim occurs, the size of the claim is more random. For this reason, primary losses (which reflect claim frequency) are not discounted while excess losses (which reflect claim severity) are. The credibility weight is an increasing function[13] of expected losses, so, all else equal, excess losses have more of an impact on larger firms' EMODs. The exact function used to translate expected losses into the weight varies by state and typically changes from year to year (hence the function is subscripted with both *s* and *t*). Note that, unlike the split point (which is applied at the claim level), the credibility weight is applied after summing primary and excess losses to the firm level. For a multistate policy, the credibility weight is an average of the state-specific weights, with the average weighted by the state's share of the firm's expected losses.

The final parameter in Equation 1 is the **ballast**, $b_{st}(E)$. The ballast is a value that is applied to reduce the potential impact of a single claim on the EMOD. As with the weight, the ballast is an increasing function of expected losses; the exact function is determined at the state and year level. Also as with the weight, the ballast for a multistate policy is determined by taking a weighted average of the state-specific ballast values.

Two other important steps are not shown in Equation 1. The first is the **experience rating adjustment (ERA)**, sometimes referred to as the **medical-only discount**. The majority of WC claims are medical-only and therefore have no wage replacement costs. Most NCCI states include medical-only claims in the EMOD at only 30% of their actual value. The rationale for this is to incentivize firms to report medical-only injuries; firms might otherwise choose to pay them out of pocket to avoid impacting their EMODs. Thus, for medical-only claims occurring in states that observe the ERA, both primary and excess losses are multiplied by 0.3 before summing to the policy level. Note that this is done *after* applying the split point. For example, in a state with a split point of $5,000, a $7,000 medical-only claim would translate to $1,500 in primary losses and $600 in excess losses. That is, there would still be excess losses even though the total claim multiplied by 0.3 would be below the split point.

The second is the **state accident limitation**, which is the cap on the amount of losses from a single claim that can be included in the EMOD. The accident limitation is above $100,000 for all states and thus is rarely relevant. Accident limitations are applied because claims this large are rare and random enough that they are not reliable predictors of a firm's future losses. Again, this is applied at the claim level, not the aggregate level. For example, if a firm had one $205,000 claim and one $1,000 claim, both occurring in a state with a $200,000 accident limitation, its total actual losses would be $201,000; the accident limitation would have no bearing on the second claim.

[13] Specifically, both the weight and ballast are stepwise functions of expected losses.

Non-NCCI formulas

As noted above, Equation 1 applies to policies rated by NCCI or by one of the following non-NCCI bureaus: Michigan, Massachusetts, North Carolina, Wisconsin, or New York. Other bureaus use different formulas, but incorporate most of the same parameters.

The Minnesota formula (used for single-state policies only, since Minnesota partners with NCCI for multistate policies) is:

$$(4)\ M_t = 1 + \frac{\left(\sum_{h=t-4}^{t-2} A_h - E_t\right) w_t(E_t) - \left(\sum_{h=t-4}^{t-2} A_{P,h} - E_{p,t}\right)[1 - w_t(E_t)]}{E_t + b_t(E_t)}$$

Expected losses *E* are calculated as in Equation 2. Expected primary losses E_p are calculated as in Equation 3, except that d_{cst} itself is substituted in place of (1 - d_{cst}). Note that the *s* subscripts are no longer needed, because Equation 4 applies only to single-state Minnesota policies.

California used the NCCI formula through 2012. Beginning in 2013, the California formula is:

$$(5)\ M_t = \frac{w_{p,t}(E_t) \sum_{h=t-4}^{t-2} A_{p,h} + \left[1 - w_{p,t}(E_t)\right] E_{p,t} + w_{e,t}(E_t) \sum_{h=t-4}^{t-2} A_{e,h} + [1 - w_{e,t}(E_t)] E_{e,t}}{E_t}$$

Again, the various expected loss values are calculated as in Equations 2 and 3. The *s* subscripts are also unnecessary for California, since California payrolls and losses are never pooled with other states. Note that California applies credibility weights both to primary losses (w_p) and excess losses (w_e). Therefore, unlike NCCI, the California bureau also discounts primary losses for some (mostly small) firms.

Pennsylvania and Delaware use the following formula:

$$(6)\ \frac{w_t(E_t) \sum_{h=t-4}^{t-2} A_{p,h} + c_t E_t + [1 - w_t(E_t)] E_t}{E_t}$$

In Equation 6, *c* is an additional state-determined parameter called the **limit charge amount**. Again, the *s* subscript is unnecessary because Pennsylvania and Delaware do not pool with other states (including with each other). Note that excess losses do not enter into the EMOD at all. To compensate for this, these states use higher split points than the others. Additionally, the Delaware split point is an increasing function of expected losses.

Pennsylvania and Delaware also use a slightly different method of calculating expected losses:

$$(7)\ E_t = \sum_{h=t-4}^{t-2} \sum_{c} R_{ch} l_{ch,t}$$

Note the subtle difference from Equation 2: rather than applying the same set of ELRs to all historical payrolls, Pennsylvania and Delaware use different ELRs for each historical year in the experience period. However, the exact ELRs used for each lagged year do change over time. For example, given a policy effective in 2012, different ELRs are used for each of the years in the experience period (2008, 2009, 2010). If the policy is renewed in 2013, different ELRs will again be used for 2009, 2010, and 2011; but the ELR that is used for, say, 2009 may not the same as what was previously used in calculating expected losses for the 2012 policy.

Also note that Equation 6 does not distinguish between primary and excess expected losses, so the d-ratio is not a relevant parameter for these states.

Caps and limits on the EMOD

Most states do not allow the EMOD to exceed a particular cap, which increases with expected losses. For all states except California, Delaware, and Pennsylvania, the following formula determines the cap for policies in effect *prior to 2013*:

$$(8)\ 1 + \ 0.00005 * \left(E_t + \frac{2E_t}{g_{st}}\right)$$

where g_{st} is the **g-value**, another state-determined parameter. Beginning in 2013, the formula applicable to all of these states (except Massachusetts, which continued to use Equation 8) is:

$$(9)\ 1.1 + \left(0.0004 * \frac{E_t}{g_{st}}\right)$$

Remember, multistate NCCI policies use a single EMOD for all states. For such policies, the cap is calculated separately for each state, and the cap from the state with the largest expected losses is then used as the overall cap.

California did not use a cap prior to 2015. Beginning in 2015, California applies a cap *only* to policies that have a single historical claim in their experience period. For such policies, the EMOD is capped at 125% of the EMOD that would be applied if the firm had not had any claims during the experience period. No cap applies to policies with more than one claim in the experience period.

Pennsylvania does not allow the EMOD to change by more than 25% in either direction from year to year. Thus, Pennsylvania is the only bureau that applies a minimum as well as a maximum limit on the EMOD.

It appears that Delaware does not apply any cap or limit.