The Human Capital Behind Accounting Estimates: Evidence from CECL Adoption

Haozheng Wang*

Zicklin School of Business, Baruch College, CUNY

August 2026

Abstract

I examine whether an information-intensive accounting standard burdens the employees who implement it and whether their working conditions are associated with estimate quality. I study CECL adoption by U.S. bank holding companies, a setting in which the standard's incremental working demands concentrate in a specific function, risk management. Using regulatory filings and Glassdoor reviews classified by job function, I find that risk-function work-life balance declines relative to non-risk-function in the same bank following adoption, equivalent to roughly three-quarters of the positive differential that risk-function employees reported before the standard. I then examine whether the allowance for credit losses maps more strongly into subsequent charge-offs when risk-function work conditions are better. This association varies little under the incurred-loss model but is substantially stronger under CECL, with the sensitivity of allowance informativeness to workforce conditions roughly five times larger under the new regime. The findings indicate that the employees who bear a standard's implementation costs are also a critical production input into the informational benefit the standard is intended to deliver.

JEL classification: G21; G28; J28; M41

Keywords: Current Expected Credit Losses (CECL); accounting estimates; human capital; work-life balance; allowance informativeness

*Haozheng Wang: haozheng.wang1@baruch.cuny.edu. I sincerely thank my dissertation committee, Kalin Kolev (Chair), Heedong Kim, Dexin Zhou, and Edgar Rodriguez-Vazquez for their guidance and support. I am also grateful for the comments provided by faculty members and PhD students at Baruch College. I gratefully acknowledge the financial support of the Zicklin School of Business, Baruch College, and the Weinstein Professorship Fellowship. All errors are my own.

1. Introduction

Over the past two decades, accounting standards have expanded the informational demands of financial reporting by requiring estimates that rely on inputs generated outside the traditional accounting function (Dyer, Lang and Stice-Lawrence, 2017)[1]. The shift creates a regulatory tradeoff. Information-intensive standards can improve reporting relevance and timeliness (Beatty and Liao, 2014; Kim, Kim, Kleymenova and Li, 2026), but the inputs they require are costly to produce and noisy by nature, and implementation frictions can erode the intended benefits (Watts, 2003; Mahieux, Sapra and Zhang, 2023; Bonsall, Schmidt and Xie, 2025). While this debate has been conducted in terms of informativeness, relevance, and comparability, the evidence rarely reaches the floor where the numbers are made.

Accounting rules do not apply by themselves, and the preparer workforce is the operational margin on which the tradeoff plays out. When the workforce can absorb the new demands, the information production a standard requires is more likely to be performed at the depth the standard envisions. When capacity is constrained, frictions surface as errors and quality losses (Doyle, Ge and McVay, 2007; Khavis and Krishnan, 2021). Whether a standard delivers its intended benefits therefore depends on the operating state of the workforce that implements it, and no financial statement or regulatory filing records that state. This paper asks whether an information-intensive standard strains the specialized workforce that implements it, and whether the quality of the standard's central estimate depends on the operating condition of that workforce.

I examine these questions in the setting of U.S. banks' adoption of the current expected credit loss (CECL) standard. CECL requires banks to estimate lifetime expected credit losses using current conditions and reasonable and supportable forecasts, linking financial reporting directly to credit-risk modeling, macroeconomic forecasting, and model validation

[1] Examples include fair value measurement under ASC 820 and IFRS 13, revenue recognition under ASC 606 and IFRS 15, lease accounting under ASC 842 and IFRS 16, hedge accounting under ASC 815, and expected credit loss estimation under ASC 326 and IFRS 9.

(FASB, 2016). The setting has three features that make the questions tractable. First, the standard's incremental demands concentrate in an identifiable internal function, risk management, rather than dispersing across many units, which permits a within-bank comparison of the employees most exposed to the standard against other employees of the same institution. Second, CECL adoption followed a staggered schedule, with large banks adopting primarily in 2020 and smaller institutions in 2023 under the CARES Act deferral (CSBS, 2021), which provides adoption cohorts separated in calendar time. Third, the standard's central output, the allowance for credit losses, is an estimate whose quality is measurable, since an allowance that fulfills its objective should be associated with the credit losses the bank subsequently realizes (Altamuro and Beatty, 2010; Beatty and Liao, 2014).

Answering these questions requires observing the workload of the employees who implement the standard. CECL's incremental demands, lifetime loss modeling, macroeconomic forecasting, model validation, and documentation, fall on the bank's risk-function employees, so the relevant measure must capture the operating condition of that specific workforce rather than the sentiment of the bank at large. I construct such a measure from work-life balance (*WLB*) ratings posted by bank employees on Glassdoor. *WLB* is the rating dimension tied most directly to workload, deteriorating when demands on employees' time exceed sustainable supply, and prior work indicates that it captures working conditions rather than firm fundamentals (Green, Huang, Wen and Zhou, 2019). Exploiting function identification from self-reported job titles (Huang, Li and Markov, 2020), I construct *WLB_Gap*, the difference between the average *WLB* rating of risk-function reviews and that of non-risk reviews in the same bank-quarter. The within-bank, between-function design differences out bank-wide culture, aggregate shocks, and stable differences in how functions rate their employers, so a decline in the gap isolates deterioration in the working conditions of precisely the employees on whom the demands fall.

The sample joins two data sources. The bank panel comes from FR Y-9C regulatory filings and covers U.S. bank holding companies from 2015 through 2024, a period providing at least four years of pre-adoption data for both CECL adoption cohorts. Because Glassdoor coverage concentrates among larger institutions, review histories are collected for BHCs reaching five billion dollars in total assets during the sample period, using a matching pipeline that maps each parent BHC and its subsidiary commercial banks to their Glassdoor pages. Of the BHCs meeting the threshold, 279 have function-classifiable reviews and 143 are CECL adopters. The tests of workforce strain draw on the 1,960 bank-quarters across 128 BHCs meeting the review requirements for both function groups. The tests of allowance informativeness draw on the CECL-adopting subset, relating each adopter's allowance to its subsequently realized charge-offs in both the incurred-loss and CECL periods. At the four-quarter charge-off horizon, these tests cover 1,009 bank-quarters across 80 adopters, narrowing to 660 bank-quarters across 34 banks in the strictest sample, which holds the set of banks fixed across the two regimes.

I begin by examining whether CECL adoption strains the workforce that implements it. Using a difference-in-differences design around staggered adoption, I find that the within-bank work-life balance gap between risk-function and non-risk employees reduces significantly following adoption, a decline equivalent to roughly three-quarters of the positive differential that risk-function employees reported before the standard took effect. Decomposing the gap shows that the decline is entirely attributable to deteriorating conditions among risk-function employees, with the non-risk benchmark unchanged. Dynamic estimates show no differential trend before adoption and no reversion through the third post-adoption year, consistent with a persistent change in workload rather than a transitory implementation spike.

A natural concern is that the decline reflects something other than CECL's demands, perhaps due to general discontent among risk-function employees, or the disruption of the pandemic that coincided with the first adoption cohort. The evidence supports neither explanation. Across the five remaining Glassdoor rating dimensions, covering pay, advancement, culture, and management, I find no post-adoption deterioration, with the effect appearing only in the dimension tied to workload. Moreover, banks adopting in 2022 or later exhibit a gap decline more than three times the full-sample estimate, consistent with heavier per-employee implementation burdens at smaller institutions with thinner risk infrastructure and inconsistent with pandemic disruption driving the baseline. These estimates are robust to influence diagnostics, alternative inference methods, and estimators designed for staggered adoption.

Having established that CECL strains the risk function, I then ask whether the quality of the standard's central output depends on the condition of that same workforce. To this end, I relate each bank's allowance to the credit losses it subsequently realizes, and examine whether this association varies with risk-function work conditions measured over the preceding year. Under the incurred-loss regime, the informativeness of the allowance varies little with risk-function conditions, consistent with an estimate anchored to observable loss events and historical experience. Under CECL, allowances established under better risk-function conditions map substantially more strongly into subsequently realized losses, and the difference between the two regimes is statistically and economically significant. To gauge the magnitude, a dollar of allowance at a bank whose risk-function conditions are one standard deviation above the mean predicts 3.7 cents more of charge-offs realized over the following year under CECL, compared with 0.7 cents under the incurred-loss regime, a roughly fivefold increase in the sensitivity of allowance informativeness to workforce condition. The pattern strengthens as realizations accumulate across horizons, as the measurement logic of a lifetime estimate predicts, and it holds across control specifications, within the fixed set of banks observed in both regimes, and under inference methods suited to the sample's cluster counts.

Together, the findings stress a tension in information-intensive standard setting. CECL increases the demands placed on the risk function, and under CECL the predictive validity of the allowance is more sensitive to the operating condition of that same function. The employees who bear the standard's implementation costs are therefore also a critical production input into the informational benefit the standard is intended to deliver.

This study contributes to two streams of literature. First, it extends research on how changes in accounting requirements affect firms' internal information production. Prior studies find that new accounting requirements can change managers' information sets and operating decisions (Shroff, 2017; Cheng, Cho and Yang, 2017). Recent work also documents that firms respond to new standards by investing in accounting personnel, information systems, and specialized human capital (Roh, 2025; Kim et al., 2026). These studies provide evidence on firms' organizational responses to reporting requirements and the benefits associated with those responses. Less is known about how the additional demands created by a new standard affect the employees involved in implementation. I extend this literature by examining the work conditions of the internal function most directly involved in producing the information required by CECL. Whereas hiring and systems investments measure organizational inputs, work-life balance captures the operating conditions experienced by employees. The decline in risk-function work-life balance provides evidence that implementation demands are reflected in employee working conditions, a dimension that is not captured by observed investments in personnel and technology. This finding broadens the analysis of implementation costs to include the conditions under which the required accounting information is produced.

Second, this study contributes to research on the production of accounting estimates. Prior studies link financial reporting outcomes to workforce education and accounting employee compensation and link audit quality to employee work-life balance (Call, Campbell, Dhaliwal and Moon Jr, 2017; Armstrong, Kepler, Larcker and Shi, 2024; Khavis and Krishnan, 2021). Research on loan loss accounting further shows that the adequacy and informativeness of credit-loss estimates vary with borrower information, manager discretion, and errors in banks' forecasts of future losses (Yang, 2022; Bischof and Rudolf, 2025; Bhojraj, Kleymenova, Liu, Lu and Sengupta, 2025). These studies establish that information and the individuals involved in reporting affect accounting outcomes. I extend this literature by examining the time-varying work conditions of the internal function that produces a specific accounting estimate and by testing whether the relevance of those conditions changes with the task prescribed by the accounting regime. I find that the association between the allowance and future charge-offs becomes more sensitive to prior risk-function work conditions under CECL than under the incurred-loss model. This result is consistent with work conditions becoming more important when an accounting estimate requires greater use of forecasts and judgment. Taken together, the findings indicate that CECL places greater demands on the risk function and that the predictive validity of the allowance under CECL is more closely associated with the condition of that function.

The remainder of this study is organized as follows. Section 2 develops the hypotheses. Section 3 describes the data and sample. Section 4 presents the empirical results. Section 5 concludes.

2. Background and Hypothesis Development

Standard-setters argue that incorporating current economic conditions and up-to-date expectations of the future will increase the decision-usefulness of the accounting numbers (Barth, 2006). As a result, accounting standard-setting over the past two decades has steadily increased reliance on forward-looking, model-based, or forecast-driven inputs in recognition and measurement. Fair value measurement under ASC 820 and IFRS 13 requires valuation models calibrated to current market conditions. Revenue recognition under ASC 606 and IFRS 15 requires estimates of transaction price, allocation across performance obligations, and variable consideration. Lease accounting under ASC 842 and IFRS 16 requires present-value computations using firm-specific discount rates. Hedge accounting under ASC 815 requires ongoing effectiveness assessment of model-derived hedge relationships. Expected credit loss estimation under ASC 326 and IFRS 9 requires lifetime loss forecasts that incorporate macroeconomic projections. Taken together, these developments point to a substantive shift in the inputs that flow into recognition, with contemporary financial statements depending more heavily on judgment-intensive estimates than before.

Producing these inputs typically requires expertise that resides outside the accounting function. It relies on specialized internal units with credit-risk modeling, valuation, treasury, or contract-analysis capabilities. Financial reporting under information-intensive standards is therefore best characterized as a coordinated production process across multiple internal functions, in which the quality of recognized amounts depends jointly on the quality and timeliness of upstream inputs and on the accounting function's ability to convert them into recognized and disclosed values.

Standard-setters, as a result, face a tradeoff between the informational benefits of forward-looking, model-based reporting and the costs of producing the inputs those standards require (Watts, 2003). Related work has documented several cost-side consequences of information-intensive standards. The standards can affect firm behavior and decisions (Plantin, Sapra and Shin, 2008; Allen and Carletti, 2008; Laux and Leuz, 2010; Beatty and Liao, 2011; Mahieux et al., 2023), change the external information environment and information production (Daske, Hail, Leuz and Verdi, 2008; Christensen, Hail and Leuz, 2013; Bonsall et al., 2025), create implementation difficulties in internal control (Doyle et al., 2007), and alter managers' discretion and incentives over financial reporting (Beaver and Engel, 1996; Bushman and Williams, 2012). However, the human-capital strain imposed on the internal workforce that implements these standards has received less attention, despite being one of the most direct margins through which implementation costs arise.

Human capital strain is the operational margin on which information-intensive standards translate, or fail to translate, into reported numbers. When the implementing workforce has slack to absorb new demands, the information production activities the standard requires are more likely to be performed at the depth and precision the standard envisions. When the workforce is constrained, the same activities are performed under pressure, and recognized amounts can inherit the noise, errors, and shortcuts that pressure introduces. Consistent with this view, Khavis and Krishnan (2021) find that workforce strain transmits to the quality of reporting output. A standard can prescribe a particular methodology, but whether that methodology is implemented as intended depends on the workforce that implements it, and the operating state of that workforce is not directly observable in financial statements or in standard regulatory databases. A firm may appear well-resourced on its balance sheet yet be operationally constrained in the specific internal functions responsible for producing the inputs a new standard requires. Studying human capital strain directly therefore opens a margin of analysis that observable firm characteristics cannot reach.

This paper studies the implementation channel in the specific setting of U.S. banks' adoption of CECL. Codified in ASC 326, CECL requires entities to estimate lifetime expected credit losses on financial assets using current conditions, historical experience, and reasonable and supportable forecasts (FASB, 2016). For banks, the allowance for credit losses must therefore reflect expected losses on each loan and held-to-maturity security across its full remaining life, rather than the previously required incurred-loss estimates triggered by observable deterioration. Several features make CECL a particularly clean setting for the implementation-channel question. First, the standard's incremental demands, including lifetime expected loss modeling, macroeconomic forecasting, model validation, and documentation, concentrate in risk management, a distinct and significant internal function of banks (Beatty and Liao, 2014). This functional concentration is unusual among recent information-intensive standards, which more typically disperse incremental workload across sales, treasury, legal, valuation, and accounting functions simultaneously, diluting any function-level signal. Second, the allowance for credit losses is among the most economically significant accounting estimates banks produce. Although regulatory capital effects were subject to transition relief during CECL implementation, changes in expected credit-loss estimates can affect reported earnings, retained earnings, and capital-related measures. Third, bank holding companies file Y-9C reports quarterly, providing standardized measurement of the allowance, realized charge-offs, and bank fundamentals, which makes both the strain and its reporting consequences measurable in one panel. Fourth, the FASB schedule combined with the CARES Act deferral generates two adoption cohorts separated in calendar time, with large banks adopting primarily in 2020 and smaller filers adopting primarily in 2023 (FASB, 2016; CSBS, 2021). This staggered structure provides useful variation in adoption timing, although the interpretation still requires attention to contemporaneous shocks affecting each adoption cohort.

CECL's incremental demands did not exist in equivalent form under the prior incurred-loss model (Mahieux et al., 2023). Because these activities are housed in the risk-management infrastructure of the bank, the operational pressure created by CECL adoption should fall disproportionately on risk-function employees relative to employees in non-risk functions within the same bank. Other internal functions are affected indirectly through coordination, documentation, and reporting interfaces, but they are not the primary operational locus of the standard's incremental workload.

However, this prediction is not a foregone conclusion. Banks can respond to anticipated demand shocks by adjusting headcount, restructuring teams, or outsourcing portions of the workload. Kim et al. (2026) provide evidence that CECL adoption is associated with such workforce adjustments, suggesting that banks do not passively absorb the new demands. If hiring and reorganization proceed at sufficient speed and scale, the observed operational pressure on existing risk-function employees could remain unchanged, or even decline, as incremental work is absorbed by newly added staff. Whether the workforce adjustment is fast enough and large enough to fully offset the demand shock is therefore an empirical question. This discussion leads to the first hypothesis:

H1: Following CECL adoption, risk-function employees experience a relative deterioration in work-life balance compared with non-risk employees in the same bank.

The allowance for credit losses is an estimate of future losses, so its quality is measurable. An allowance that fulfills its reporting objective should be associated with the credit losses the bank subsequently realizes, and the strength of this association is the standard validity benchmark in the loan loss provisioning literature (Altamuro and Beatty, 2010; Beatty and

Liao, 2014). Altamuro and Beatty (2010) document that the provision's association with subsequent charge-offs strengthened after internal control regulation raised the quality of the reporting process, indicating that the provision-to-loss mapping is not fixed by the accounting rules alone but varies with the quality of the process that produces the estimate. This paper examines a production input that prior work has not observed, the operating condition of the workforce performing the estimation.

Producing a valid allowance is an information-production task. Loss forecasting, model maintenance, scenario selection, and validation are performed by the bank's risk function, and the resulting estimate can be no better than the process that generates it. Earlier work documents that capacity constraints in the reporting process degrade reporting outcomes. Doyle et al. (2007) show that firms with constrained resources are more likely to disclose material weaknesses in internal control, Ashbaugh-Skaife, Collins, Kinney Jr and LaFond (2008) show that internal control deficiencies are associated with lower accrual quality, and evidence from audit production indicates that workforce strain transmits to the quality of reporting output (Khavis and Krishnan, 2021). A constrained workforce might meet incremental demands by economizing on the tasks that determine estimate quality, including model maintenance, input verification, and review. The allowance then contains more estimation noise, which weakens the association between an estimate and the outcome it forecasts. Therefore, the association between the allowance and future losses should be increasing in the operating slack of the risk-function workforce that produces it.

The dependence of estimate quality on workforce condition should differ across the two regimes. Under the incurred-loss model, banks recognized credit losses only after available evidence indicated that a loss had probably been incurred. Although the model still required judgment, observable loss indicators and historical loss experience provided relatively strong anchors for the allowance. CECL retains these inputs but requires the risk-function workforce to translate current conditions and reasonable and supportable forecasts into estimates of lifetime expected losses (FASB, 2016). The resulting allowance therefore depends more heavily on the workforce's ability to process uncertain information, evaluate model outputs, and exercise judgment about losses that have not yet been observed. The argument is also consistent with recent evidence on the forecasting demands created by CECL. Bonaldi, Liang and Yang (2023) document greater accounting noise and reporting bias among CECL adopters during the uncertainty surrounding the COVID-19 pandemic. Canals-Cerdá (2026) shows that CECL is more sensitive than the incurred-loss framework to macroeconomic forecast error and model misspecification. Using confidential supervisory forecasts, Bhojraj et al. (2025) document systematic errors in banks' forecasts of future charge-offs and find that the effect of those errors on loan loss provisions is concentrated among CECL adopters. These findings suggest that the quality of forward-looking inputs becomes more important when the accounting regime requires lifetime expected-loss estimates. Risk-function work conditions should therefore be more consequential under CECL than under the incurred-loss model.

However, several mechanisms could mute or eliminate this strengthening. Banks increasingly estimate expected losses on vendor platforms and pooled industry models, which could decouple the quality of the recognized allowance from the operating condition of the local risk function. Supervisory review and model-validation requirements could hold the allowance to a quality floor that binds regardless of internal strain. Banks can also hire or reallocate staff toward the estimation task, as Kim et al. (2026) document workforce adjustments around CECL adoption. Each mechanism, operating at sufficient scale, may lead to no change in the dependence of allowance validity on work conditions across regimes. Whether the dependence strengthens under CECL is therefore an empirical question. This discussion leads to the second hypothesis:

H2: The association between the allowance for credit losses and subsequently realized credit losses is increasing in the work conditions of the bank's risk function, and this dependence is stronger under CECL than under the incurred-loss regime.

The two hypotheses are connected. *H1* predicts that CECL adoption strains the risk-function workforce, and *H2* predicts that under CECL the quality of the standard's central output depends more heavily on the condition of that same workforce. Together, the two hypotheses examine whether the standard concentrates incremental demands on the function whose operating condition its output quality has come to depend on.

3. Data and Sample

3.1. Employee Glassdoor Reviews and the Workforce Strain Measure

3.1.1. Glassdoor employee reviews

I obtain employee reviews from Glassdoor, an online platform on which current and former employees of public and private companies anonymously rate their employers. Glassdoor was founded in 2007 and launched its employer-review functionality in 2008. The platform requires reviewers to sign in using a permanent email address or a valid social-network account before they can browse content posted by other users, and the company maintains a content-review team that flags inauthentic submissions, providing baseline data quality (Hilary, Tian and Yu, 2025). Glassdoor data have been used extensively in recent accounting and finance research, including Huang et al. (2020) on the information content of employee outlook, Green et al. (2019) on the relation between employee ratings and stock returns, Hilary et al. (2025) on employee responses to earnings news, Dehaan, Li and Zhou (2023) on employee job search around earnings announcements, and Khavis and Krishnan (2021) on workforce conditions at audit firms.

Each review contains an overall employer rating and five sub-dimension ratings, each on a five-point scale. The sub-dimensions cover Work/Life Balance, Senior Management, Compensation and Benefits, Career Opportunities, and Culture and Values. Each review also records a self-reported job title, a posting date, an indicator for current or former employment status, and free-text comments. Two of these fields carry the empirical design. The job title permits classification of reviewers into internal functions (Section 3.1.3), and the sub-dimension structure permits the work-life balance rating to be examined separately from the other dimensions of the employment experience, a feature the placebo analysis in Section 4.1.3 exploits.

I retain reviews submitted by both current and former employees during the sample period. This choice differs from some prior Glassdoor-based studies that restrict attention to current employees. The construct of interest is the operating state of the function rather than the reviewer's current employment relationship. Former employees report on conditions they directly experienced, and those reports are informative about workforce strain regardless of the reviewer's status on the posting date.

3.1.2. Work-life balance as the proxy for workforce strain

I use the Work/Life Balance (*WLB*) rating as the empirical proxy for workforce strain. The construct underlying *H1* is the intensity of demands placed on a workforce relative to the time and personal resources its members have available. *WLB* is the rating dimension tied most directly to that construct. When human resource demand exceeds sustainable supply, employees work longer hours, take on more concurrent tasks, and report worse work-life balance. Using *WLB* rather than the overall rating or an average across the five sub-dimensions isolates the operational construct without diluting the signal with dimensions of the employment experience less closely tied to workforce utilization.

Two findings in prior research support this choice. First, *WLB* tracks workforce operating conditions that carry through to reporting outcomes. Khavis and Krishnan (2021) find that audit firms with poorer *WLB* ratings produce lower-quality audits, which indicates that the rating captures variation in workforce conditions consequential enough to affect measurable output quality. Second, *WLB* is comparatively free of firm-performance signal. Green et al. (2019) find that, among Glassdoor rating sub-dimensions, changes in *WLB* ratings are unrelated to stock returns, in contrast to changes in the career-opportunities and senior-management dimensions. If *WLB* were impounding shifting firm fundamentals, function-level variation in the rating would partly capture economic conditions rather than working conditions. Taken together, the two findings indicate that *WLB* measures the operating state of a workforce while remaining insulated from the general firm-level signal that other rating dimensions carry.

The remaining sub-dimensions are not discarded. Because every review rates Senior Management, Compensation and Benefits, Career Opportunities, and Culture and Values on the same scale, these dimensions provide natural comparison outcomes. If CECL adoption strains the risk-function workforce rather than altering the general standing of risk-function employment, the effect should concentrate in *WLB* and be absent from the other dimensions. Section 4.1.3 implements this comparison.

3.1.3. Function classification

H1 predicts a differential effect across internal functions within the same bank, which requires assigning each review to a function. Huang et al. (2020) classify Glassdoor reviewers into functional areas using self-reported job descriptions and find that outlook from employees in a given function is incrementally useful in predicting the component of future operating performance most closely related to that function. Their evidence indicates that employee reviews carry function-specific information, and that job titles are sufficient to recover it. I build on this insight by classifying each review through the reviewer's reported job title. Employees in credit-risk, modeling, forecasting, and model-validation roles are positioned to experience the operational pressure created by CECL's expected-loss estimation requirements directly, and their ratings are correspondingly positioned to capture it.

The classifier operates in two stages. In the first stage, I normalize each self-reported job title by lowercasing, collapsing whitespace, standardizing punctuation, folding unicode dash variants, and stripping pure seniority modifiers that do not carry role-defining meaning[2]. Role-defining nouns that function as the head of the title are preserved[3]. The output of this stage is a base role for each review (e.g., *risk manager*, *credit analyst*, *operations director*). Working from base roles rather than raw job titles reduces classification noise from minor title variants.

In the second stage, I manually label the base-role catalog into a binary risk-function indicator. Risk-function reviewers are identified by base roles containing terms such as *risk*, *credit risk*, *model risk*, *model validation*, *credit analyst*, and related variants. Non-risk reviewers are all reviewers in the bank-quarter not classified into the risk function. Manual labeling at the base-role level, rather than keyword matching on raw titles, ensures that classification decisions are made once per role and applied consistently across every review carrying that role.

3.1.4. Matching bank holding companies to Glassdoor pages

A central data-construction challenge is mapping bank holding companies in the regulatory panel to their Glassdoor pages. BHCs file regulatory reports under their legal parent names, while Glassdoor organizes pages under brand names, and a parent may hold multiple commercial-bank subsidiaries whose separate pages must be attributed to the parent for consistency with the regulatory data. I address the challenge through a multi-step matching pipeline that generates cleaned name variants for each parent-subsidiary pair, scores candidate Glassdoor pages with fuzzy-matching metrics, applies conservative acceptance filters, and routes unmatched residuals to manual review. Collected review histories pass integrity

[2] The stripped tokens include *senior*, *sr*, *jr*, *junior*, *principal*, *staff*, *lead*, *assistant*, *associate*, *intern*, *temporary*, *contract*, *contractor*, *global*, *regional*, *corporate*, and numeric level suffixes (*I*, *II*, *III*, *IV*, *V*, *level 1*, *level 2*, and so on).

[3] For example, *manager*, *director*, *vice president* (and its variants *VP*, *AVP*, *SVP*, *EVP*), *chief*, *head of*, and *supervisor*.

checks covering existence, valid parsing, and non-empty content. Appendix A documents each step of the pipeline.

3.1.5. The WLB Gap

Individual reviews are aggregated to the BHC-quarter level using a four-quarter window. For a BHC-quarter (i, t), the aggregation pool consists of all reviews posted between the beginning of quarter t and the end of quarter $t + 3$[4]. The forward-looking window reflects how employee reviews arrive. Reviews are unlikely to capture only conditions on the exact posting date. Rather, they capture accumulated workplace experience and employee responses to recent firm events. Huang et al. (2020) include predictions from both current employees and recently departed former employees, consistent with employee knowledge depreciating gradually rather than disappearing immediately. Hilary et al. (2025) measure changes in Glassdoor ratings from the event quarter to the following quarter, allowing employee responses to accounting-related events to appear with a lag. Guided by this logic, the window assigns to quarter t the reviews that describe conditions experienced during and shortly after t, including reviews posted with a delay.

For each bank-quarter and function group, I compute the number of reviews and the mean *WLB* rating across reviews in the group. The measure of differential operating conditions is defined for BHC i in quarter t as

$$\text{WLB\_gap}_{i,t} \;=\; \overline{\text{WLB}}_{i,t}^{\text{risk}} \;-\; \overline{\text{WLB}}_{i,t}^{\text{non-risk}}. \tag{1}$$

Because higher ratings indicate better work-life balance, a decline in WLB_gap indicates that risk-function work-life balance deteriorates relative to non-risk work-life balance in the same bank at the same time. The within-bank, between-function structure differences out three

[4]For example, for a BHC-quarter observation ending March 31, 2020, the pool draws reviews posted between January 1, 2020 and December 31, 2020.

sources of variation that would otherwise confound the interpretation. These are bank-wide cultural baselines that shape how reviewers from any function rate their employer, time-invariant differences across reviewers from different functions, and aggregate shocks to the extent that they affect all functions of the bank similarly.

The bank-wide component is not a hypothetical concern. Appendix Table A1 examines the cross-sectional determinants of pre-adoption risk-function *WLB* across the 67 BHCs with sufficient pre-adoption coverage. Bank size, capital, earnings capacity, deposit funding, and employment scale have no explanatory power for the level of risk-function *WLB*, and the adjusted R^2 is near zero. Adding the bank's non-risk *WLB* average raises the adjusted R^2 to 0.24, with a coefficient of 0.90 that is statistically indistinguishable from one. The level of risk-function *WLB* therefore loads close to one-for-one on a bank-wide component shared with every other function, and any analysis of the level would capture bank-wide culture rather than conditions specific to the risk function. The same table shows that the pre-adoption gap is unrelated to all of the same bank characteristics. Differencing removes the shared component and leaves a measure that observable bank fundamentals do not explain.

The design does not, by itself, rule out contemporaneous shocks that differentially affect risk-function employees. I address this concern directly in the results, where placebo tests on the remaining rating dimensions (Section 4.1.3) and a separate analysis of the late-adopter cohort (Section 4.1.3) indicate that the post-adoption decline reflects workload concentrated in the risk function rather than function-wide dissatisfaction or pandemic-period disruption.

I require a bank-quarter cell to contain at least three risk-function reviews and at least three non-risk reviews for the gap to be computed. The restriction balances measurement precision against sample coverage, and Section 3.2 reports the resulting sample. The tests of *H2* use a backward-looking variant of the same measure. For those tests, the aggregation pool for BHC-quarter (i, t) consists of reviews posted during quarters $t - 4$ through $t - 1$, with the same function classification, the same review-count requirement, and the same gap construction. Section 4.2.1 describes the role this variant plays in the research design and the reason the window direction reverses.

3.2. Sample Construction and CECL Adoption

The bank panel is constructed from FR Y-9C regulatory filings and covers 96,991 bank-quarter observations across 5,527 U.S. bank holding companies from 2015Q1 through 2024Q4. Bank identifiers are constructed using the WRDS Bank Structure database. For each commercial bank in the panel, identified by RSSD ID with charter type corresponding to a commercial bank (charter codes 200, 300, 310, 320, 330, and 340), I trace the ultimate parent BHC at each point in time by walking up the ownership chain in the *wrds_struct_relationships* file. The time-aware procedure preserves changes in ownership over the sample period. When a subsidiary bank changes parents during the sample, the period of association with each parent is recorded separately and used for matching to the regulatory and review data.

Because Glassdoor coverage is concentrated among larger institutions, employee reviews are collected for BHCs whose total consolidated assets reach at least five billion dollars at some point during the sample period. The threshold serves two purposes. It focuses the sample on BHCs of meaningful economic size, where CECL adoption has first-order regulatory implications, and it concentrates collection on institutions likely to generate sufficient review volume to support function-level analysis. Of the BHCs meeting the threshold, 279 have at least one function-classifiable review in at least one sample quarter through the matching procedure described in Appendix A, of which 143 are CECL adopters. Applying the three-review cell requirement of Section 3.1.5 yields the H1 estimation sample of 1,960 bank-quarter observations across 128 BHCs. The tests of H2 draw on subsamples of these banks that satisfy additional data requirements described in Section 4.2.1. The 2015 through 2024 sample period provides at least four full years of pre-adoption data for both CECL adoption cohorts and at least one full year of post-adoption data for the later cohort.

CECL adoption is identified at the BHC-quarter level using three Y-9C fields introduced specifically to track ASC 326 implementation. These are BHCKJJ26 (CECL adoption indicator), BHCKJJ27 (initial allowance for purchased credit-deteriorated assets), and BHCKJJ28 (the day-one CECL impact). For each BHC, the treatment quarter τ_i is defined as the first quarter in which any of these three fields is reported as non-missing. This convention follows recent CECL studies that use the same Y-9C fields to identify adoption timing (Kim et al., 2026; Bonsall et al., 2025). I define an indicator $Post_{i,t}$ that equals one for BHC i in quarter t if $t \geq \tau_i$ and zero otherwise.

The treatment-quarter distribution exhibits the two-cohort structure that follows from the FASB effective dates combined with the CARES Act adoption deferral (FASB, 2016; CSBS, 2021). The first cohort, comprising SEC filers that are not smaller reporting companies, adopts CECL beginning in fiscal year 2020. The second cohort, comprising smaller reporting companies, private companies, and not-for-profit entities, adopts CECL beginning in fiscal year 2023, with a small number of institutions electing earlier adoption. Throughout the paper, I refer to BHCs with τ_i in 2022Q1 or later as late adopters. The inclusive cutoff isolates the adopters whose treatment postdates the acute pandemic period, and Section 4.1.3 uses this cohort to separate the effects of CECL adoption from pandemic-period disruption. The staggered structure provides identifying variation in adoption timing that is not driven by any single contemporaneous shock.

3.3. Descriptive Statistics

Table 1 reports summary statistics for the two analysis samples. Panel A describes the *H1* estimation sample of 1,960 bank-quarter observations across 128 BHCs, meeting the within-cell review requirements of Section 3.1.5. The mean WLB_gap is 0.25 with a standard deviation of 0.63, and the component means sit at 3.76 for risk-function *WLB* and 3.51 for non-risk *WLB*, so risk-function reviewers on average rate their work-life balance above non-risk reviewers in the same bank-quarter over the sample period. The advantage is not an artifact of pooling across adoption regimes. Treated banks enter CECL with a pre-adoption mean gap of 0.22, with risk-function *WLB* averaging 3.67 against 3.45 for non-risk employees, and the late-adopter cohort enters with a mean gap of 0.19. Risk-function employees therefore approach adoption from a relatively favorable position within the bank rather than from pre-existing strain, and the estimates that follow capture the erosion of that favorable position rather than the deepening of an already-strained one. The 10th-to-90th percentile range spans -0.45 to 1.00, providing substantial variation for the difference-in-differences analysis.

Risk-function cells are considerably thinner than non-risk cells, with a mean (median) of 21.6 (7) risk-function reviews per bank-quarter against 590 (139) non-risk reviews, reflecting the smaller share of the bank workforce employed in risk and modeling roles. Two-thirds of the bank-quarter observations fall in the post-adoption period, and the remaining rows indicate a sample of large institutions, with mean log assets of 18.1, corresponding to roughly $75 billion in total assets, and a mean capital ratio of 11.3 percent.

Panel B describes the sample for the *H2* predictive-validity tests at the four-quarter horizon. The tests relate the allowance to subsequently realized credit losses, and the panel reports the three variables at the center of that design. *Allowance* is the allowance for credit losses scaled by gross loans, measured at quarter end. *Future NCO* is net charge-offs accumulated over quarters $t + 1$ through $t + 4$, scaled by gross loans. The prior-four-quarter WLB_gap is the backward-looking variant defined in Section 3.1.5. Section 4.2.1 develops the design in full, including the shorter accumulation horizons and the sample variants.

The panel covers 660 bank-quarter observations with a valid prior-four-quarter gap and the full four-quarter charge-off window. The allowance averages 1.6 percent of gross loans, cumulative four-quarter net charge-offs average 0.6 percent with a 90th percentile of 2.3 percent, and the prior-four-quarter gap has a mean of 0.25 and a standard deviation of 0.48, closely matching the contemporaneous measure in Panel A. Panel B banks are larger than Panel A banks, with mean log assets of 19.1 against 18.1, roughly $197 billion against $75 billion.

4. Research Design and Results

4.1. Workforce Strain Following CECL Adoption

This section tests *H1*. Section 4.1.1 presents the baseline difference-in-differences estimates and the dynamic pattern around adoption. The subsequent subsections decompose the estimate into its function-level components, address concurrent events and pandemic-period disruption, examine the influence of individual institutions, and assess robustness to staggered-adoption estimation.

4.1.1. Difference-in-differences estimates

The baseline specification is

$$\text{WLB\_Gap}_{i,t} \;=\; \beta\, Post_{i,t} \;+\; \Gamma' X_{i,t} \;+\; \alpha_i \;+\; \gamma_t \;+\; \varepsilon_{i,t}, \tag{2}$$

where i and t index BHC and calendar quarter, respectively. The dependent variable, $WLB\_Gap_{i,t}$, is the mean work-life balance rating of risk-function reviews minus the mean rating of non-risk reviews, as defined in Section 3.1.5. The variable of interest is $Post_{i,t}$, an indicator equal to one in and after the quarter in which BHC i adopts CECL and zero otherwise. The coefficient of interest, β, captures the average within-bank change in the risk-minus-non-risk work-life balance gap following CECL adoption, relative to contemporaneous changes at other banks in the sample. If CECL's incremental estimation demands fall disproportionately on the risk function, as *H1* predicts, β would be negative. $X_{i,t}$ is a vector of bank-level controls similar to those in Kim et al. (2026), comprising $Size_{i,t}$, the natural logarithm of total consolidated assets, $CapRatio_{i,t}$, total equity capital divided by total assets, $EBLLP_{i,t}$, earnings before loan loss provisions divided by total assets, and $Deposit_{i,t}$, total deposits divided by total assets, which absorb time-varying bank fundamentals associated with both workforce conditions and adoption timing. α_i denotes BHC fixed effects, which absorb time-invariant differences across banks in the level of the gap, including stable differences in how each bank's functions rate their employer. γ_t denotes calendar-quarter fixed effects, which absorb aggregate conditions common to all banks, including the credit-cycle and labor-market movements of the sample period. Standard errors are clustered by BHC to allow for serial correlation in the errors within banks.

Column (1) of Table 2 reports the baseline estimate. Consistent with *H1*, the coefficient on *Post* is negative and significant (-0.168, $p < 0.05$), indicating that risk-function work-life balance deteriorates relative to non-risk work-life balance within the same bank following CECL adoption. The magnitude is economically meaningful. The decline equals 27 percent of the sample standard deviation of the gap and is equivalent to roughly three-quarters of the treated banks' pre-adoption mean gap of 0.22. The remaining columns of Table 2 decompose the estimate and re-estimate it on the late-adopter cohort, and Sections 4.1.2 and 4.1.3 take up those results.

Figure 1 traces the dynamic pattern, replacing $Post_{i,t}$ in Equation (2) with event-time indicators spanning eight quarters before through twelve quarters after adoption. Because the *WLB* measure for quarter t pools reviews posted through quarter $t + 3$, measurements dated within three quarters of adoption already contain post-adoption reviews. The reference period is therefore $k = -4$, the last measurement constructed entirely from pre-adoption reviews, and the figure marks $k = -3$ through $k = -1$ as an overlapping window, separately from the clean pre-period and post-adoption estimates.

The figure indicates three patterns. First, the clean pre-period coefficients at $k = -8$

through $k = -5$ are small, with absolute magnitudes below 0.04, and statistically indistinguishable from zero ($p > 0.78$ in every quarter), consistent with the absence of a pre-existing trend in the gap. Second, the gap turns negative within the overlapping window and declines further at adoption. Two explanations can account for the early decline. The first is the structure of the measurement itself. The *WLB* measure for quarter t pools reviews posted through quarter $t + 3$, so measurements dated within three quarters of adoption already contain post-adoption reviews, and the effect materializing at adoption appears in these earlier-dated measurements mechanically. The second is preparatory workload. Firms can begin implementation work before the formal effective date (Fiechter, Hitz and Lehmann, 2022), and substantial CECL-related activity plausibly accumulates in the year preceding adoption[5]. Both explanations tie the early *WLB* gap decline to CECL, implying a potentially underestimated pooled point estimate in Table 2. Third, the post-adoption coefficients are negative in all thirteen event quarters, fluctuating between -0.10 and -0.33 with no reversion toward zero through the third post-adoption year. Individual quarterly coefficients are estimated with limited precision, reaching five percent significance in event quarters two and three, and the pooled estimate captures the average of the persistently negative post-adoption path. The pattern captured by the figure is a level shift that begins at adoption and persists, rather than a transitory implementation spike that dissipates once day-one adoption work concludes.

4.1.2. Decomposition of the WLB Gap

A decline in WLB_Gap can be consistent with two distinct patterns. Risk-function work-life balance can deteriorate, or non-risk work-life balance can improve, and the two carry different interpretations. An improvement in the non-risk benchmark would indicate that CECL adoption coincides with better conditions elsewhere in the bank, for example through

[5] As context, Goldman Sachs reported CECL-related disclosures in its 2018 annual report and disclosed a quantified day-one allowance impact in its 2019 annual report, indicating material preparatory activity in the year before adoption.

implementation resources or hiring that ease the workload of functions outside the risk area. In contrast, *H1* attributes the gap decline to rising demands on the risk function and a strained risk workforce, which predicts that the effect loads on the risk component.

Columns (2) and (3) of Table 2 estimate Equation (2) separately for the two components. Because the three columns share an identical estimation sample, the column (1) estimate equals the difference between the column (2) and column (3) estimates by construction. The coefficient on *Post* for risk-function *WLB* is negative (-0.168, $p < 0.10$), while the coefficient for non-risk *WLB* is precisely zero (-0.0004). The entire gap decline reflects deterioration in risk-function work-life balance, and the non-risk component is unchanged by adoption. The coefficient on *Post* in column (2) carries a larger standard error than the coefficient in column (1) (0.092 versus 0.082), and its statistical significance is accordingly weaker ($p < 0.10$ versus $p < 0.05$). This pattern follows from the design. The gap removes the bank-level conditions that move both functions together, while the component regressions retain them, so the common component adds noise to columns (2) and (3) that column (1) has already eliminated.

The flat non-risk benchmark carries two implications. First, the baseline estimate captures deterioration where *H1* predicts it, rather than improvement where *H1* is silent. Second, the stability of the non-risk component indicates that CECL adoption leaves the work-life balance of the average non-risk employee unaffected, which is difficult to reconcile with explanations in which adoption coincides with bank-wide workload shocks, since such shocks would move both components. The same pattern holds in the late-adopter cohort, where columns (5) and (6) show the decline loading on the risk component with the non-risk benchmark slightly positive and insignificant.

4.1.3. Is the Decline Attributable to CECL?

Two alternative explanations could account for the baseline estimate. First, risk-function employees may have become generally dissatisfied around adoption for reasons unrelated to workload, such as compensation decisions, organizational disturbance, or the standing of the function within the bank. Under this explanation, the gap decline reflects discontent rather than strain. Second, the first adoption cohort implemented CECL in the opening quarter of the COVID-19 pandemic, a period in which loan workouts, forbearance processing, and pandemic-related stress modeling plausibly raised risk-function workload independently of the standard. Under this explanation, the estimate reflects pandemic disruption rather than CECL. Each explanation implies a distinct empirical prediction, and I test each in turn.

If risk-function employees became generally dissatisfied, the deterioration should appear across the dimensions of the employment experience rather than concentrating in work-life balance. Following Khavis and Krishnan (2021), who examine the remaining Glassdoor rating dimensions alongside work-life balance, I construct risk-minus-non-risk gaps for the overall rating and the four sub-ratings (career-opportunities, compensation-and-benefits, culture-and-values, and senior-leadership) and estimate Equation (2) for each dimension. Table 3 reports the results. The coefficient on *Post* is statistically insignificant for every alternative dimension, with estimates ranging from -0.079 to 0.032, compared with -0.168 ($p < 0.05$) for work-life balance. The average coefficient across the five placebo dimensions is -0.024, and the difference between the *WLB* coefficient and the average placebo coefficient is -0.144 ($p = 0.094$). The deterioration is concentrated in the dimension tied to workload and absent from the dimensions tied to pay, advancement, culture, and management. This pattern is consistent with rising demands on risk-function employees and inconsistent with function-wide discontent.

The compensation-and-benefits dimension warrants separate interpretation, because its expected response to a workload shock differs from that of the other dimensions. The rating measures perceived level of pay rather than the pay itself, and employees evaluate compensation against outside options and perceived effort. A rise in demands with rigid pay would leave risk employees underpaid relative to effort and predicts a declining compensation gap, while a rise in demands met by compensating pay adjustments, consistent with the salary and employee benefit expenses increase that Kim et al. (2026) document, predicts a flat rating even as pay increases. The null coefficient is consistent with either a partial pay response or pay benchmarked to the external market, and it does not indicate that pay was unchanged. The contrast with work-life balance is the informative feature. The work-life balance rating tracks the time demands of the job directly, and a pay response can restore parity in the compensation dimension but cannot return the hours the additional work absorbs. A workload shock therefore must appear in work-life balance and need not appear in compensation.

The pandemic explanation implies a prediction about timing. If pandemic-period disruption rather than CECL produced the baseline estimate, banks adopting well after the acute pandemic period should exhibit little or no gap decline. Panel B of Table 2 tests this prediction by estimating Equation (2) on the late-adopter cohort, the 35 BHCs comprising adopters with treatment quarters in 2022Q1 or later and their never-treated controls, with early adopters excluded from the comparison set. These banks adopted CECL two to three years after the pandemic onset, and their post-adoption periods fall in 2022 through 2024.

The results are inconsistent with the pandemic explanation. The coefficient on *Post* in column (4) is negative and significant (-0.606, $p < 0.01$), roughly three and a half times the full-sample estimate. The decomposition in columns (5) and (6) indicates that the decline again loads on the risk component (-0.543, $p < 0.05$), while the non-risk benchmark is positive and insignificant (0.063). The estimate is robust to inference and influence diagnostics appropriate for a 35-bank cohort. It remains significant at the five percent level under wild-cluster bootstrap inference with Webb weights ($p = 0.025$, 9,999 replications), and in a leave-one-out analysis omitting each of the 35 BHCs in turn, every iteration produces a negative estimate significant at the five percent level, with coefficients ranging from -0.708 to -0.468.

The magnitude is economically large. A decline of 0.61 rating points equals 96 percent of the sample standard deviation of WLB_Gap and roughly three times the cohort's pre-adoption mean gap of 0.19, so late-adopter risk functions move from rating their work-life balance above the non-risk benchmark to rating it substantially below. The larger effect among late adopters is consistent with the mechanism underlying *H1*. Late adopters are smaller institutions with thinner risk-management infrastructure, and the same implementation demands spread across fewer specialized employees imply a heavier burden per risk-function employee. Taken together, the late-adopter evidence is inconsistent with the pandemic explanation, which predicts a smaller effect in this cohort, and consistent with the workload mechanism, which predicts a larger effect where implementation capacity is thinner.

4.1.4. Robustness of Estimation and Inference

The review volume underlying the gap measure concentrates in the largest institutions. Table 1 shows mean review counts well above their medians in both function groups, so a small number of heavily reviewed BHCs could contribute a disproportionate share of the identifying variation, and the baseline estimate could reflect these institutions rather than a pattern across the sample. Figure 2 addresses this concern by re-estimating Equation (2) 128 times, omitting one BHC per iteration, and plotting the resulting coefficients in sorted order. Every iteration produces a negative estimate, with coefficients ranging from -0.199 to -0.133 around the full-sample estimate of -0.168. All 128 estimates are significant at the ten percent level, and 88 percent are significant at the five percent level. The iterations with p-values between 0.05 and 0.10 are expected, as the full-sample p-value (0.042) sits close to the five percent threshold. Omitting a bank that contributes identifying variation raises the standard error, potentially driving the p-value above the 5 percent threshold. Overall, the leave-one-out test estimates are stable in sign and magnitude to the exclusion of any individual BHC.

Another feature of the setting, the staggered adoption structure, raises a separate estimation concern. When treatment effects vary over time, two-way fixed effects estimates place weight on comparisons in which banks treated earlier serve as controls for banks treated later, and these comparisons can bias the estimate toward zero or reverse its sign (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021). Appendix Table A2 reports estimates from the Callaway and Sant'Anna (2021) estimator, which restricts comparisons to not-yet-treated and never-treated banks and, in this setting, shifts each bank's effective adoption date three quarters earlier so that no measurement containing post-adoption reviews enters a control group. The average treatment effect across event quarters zero through four is -0.287 ($p = 0.026$), consistent in sign and significance with the baseline estimate.

4.2. Allowance Informativeness and Risk-Function Work Conditions

This section tests *H2*. Section 4.2.1 develops the research design, including the measurement of allowance informativeness and the construction of the regime comparison. Section 4.2.2 presents the baseline estimates and their economic magnitude. Section 4.2.3 examines whether the estimates behave as the measurement logic of a lifetime estimate predicts, and assesses the robustness of inference and influence.

4.2.1. Research design

H2 concerns the quality of the allowance as an estimate of future credit losses. An allowance that fulfills its reporting objective should be associated with the losses the bank subsequently realizes, and this association is a well-established quality benchmark in the loan loss provisioning literature (Altamuro and Beatty, 2010; Beatty and Liao, 2021). I implement it with the allowance, for a reason specific to the setting. CECL defines its measurement objective on the allowance itself, which must equal lifetime expected credit losses on the portfolio at each reporting date (FASB, 2016), so the allowance-to-future-loss association tests directly whether the standard's central output achieves its stated objective. The provision mixes new expected-loss information with mechanical replenishment of charge-offs taken during the period, which makes it a noisier object for the validity question under an expected-loss regime.

Realized losses are measured as net charge-offs accumulated over the quarters following the allowance date, scaled by gross loans. Because the allowance is a lifetime estimate, realizations over any short window capture only part of the losses the estimate anticipates, and the informativeness of the allowance should accumulate with the measurement horizon. I therefore measure future net charge-offs over one, two, and four quarters, and report all three horizons.

The specification estimates, within each accounting regime $r \in \{Pre, Post\}$,

$$\begin{aligned}\text{FutNCO}^h_{i,t} \;=\;& \beta_{1r}\,\text{Allowance}_{i,t} \;+\; \beta_{2r}\,\text{WLB\_Gap}^{tr}_{i,t} \;+\; \beta_{3r}\,\text{Allowance}_{i,t} \times \text{WLB\_Gap}^{tr}_{i,t} \\ &+\; \Gamma'_r X_{i,t} \;+\; \alpha_{i,r} \;+\; \gamma_{t,r} \;+\; \varepsilon_{i,t},\end{aligned} \tag{3}$$

where $\text{FutNCO}^h_{i,t}$ is net charge-offs accumulated over quarters $t+1$ through $t+h$, scaled by gross loans at quarter t. $\text{Allowance}_{i,t}$ is the allowance for credit losses divided by gross loans at quarter end. $\text{WLB\_Gap}^{tr}_{i,t}$ is the trailing variant of the gap measure, constructed from reviews posted during quarters $t-4$ through $t-1$ as described in Section 3.1.5. $X_{i,t}$ is the control vector described below, and $\alpha_{i,r}$ and $\gamma_{t,r}$ are BHC and calendar-quarter fixed effects. Every regressor and fixed effect carries the regime subscript because the model is estimated in stacked form with all terms fully interacted with the regime indicator, so each regime receives its own slopes and its own fixed effects, and cross-regime restrictions are imposed nowhere.

I estimate the stacked form rather than a triple-interaction specification because it presents each regime's coefficients directly. The two parameterizations are equivalent, and untabulated estimates of the triple-interaction form produce an identical cross-regime difference. The coefficient of interest is β_{3r}, which captures how the association between the allowance and subsequently realized losses varies with risk-function work conditions. If allowance informativeness is increasing in the operating condition of the risk function, as the first part of *H2* predicts, β_{3r} is positive. If this dependence is stronger under CECL, as the second part predicts, $\beta_{3,Post} - \beta_{3,Pre}$ is positive, and the stacked form permits a direct test of this difference. Standard errors are clustered by BHC.

The gap measure enters Equation (3) in its trailing form for two reasons, both anticipated in Section 3.1.5. First, the moderator must be predetermined relative to the estimate it conditions. The forward window used in the *H1* tests pools reviews posted after quarter t, so it could absorb employee reactions to the very reporting outcomes under examination. The trailing window fixes the risk function's operating condition before the allowance is recorded. Second, the window is regime-pure. An observation enters the estimation only if all four trailing quarters and the full charge-off horizon fall within a single accounting regime, so no moderator mixes incurred-loss and CECL-period conditions and no outcome window spans the adoption date.

Two samples are used throughout. The all-eligible sample contains every bank-quarter satisfying the data requirements. The same-bank sample retains only BHCs with usable observations in both regimes, so the comparison of β_{3r} across regimes holds the set of banks fixed and cannot reflect changes in sample composition. Because pre-period review coverage concentrates among the largest institutions, the same-bank requirement selects toward large BHCs, and the all-eligible sample indicates whether the findings extend beyond the largest institutions. Table 1, Panel B, describes the same-bank sample at the four-quarter horizon.

4.2.2. *Allowance Informativeness Before and After CECL*

Table 4 reports estimates of Equation (3) at the four-quarter horizon. The controls enter in four cumulative blocks, reported as a specification ladder in the table. The baseline block follows the allowance-to-charge-off mapping specification of Bischof and Rudolf (2025) and contains size, capitalization, and loan growth. Size absorbs differences in portfolio composition and diversification that shape loss rates across institutions of different scale. Capitalization absorbs the incentive to manage the allowance toward capital objectives, which affects the level of the allowance independently of its informational content. Loan growth absorbs portfolio seasoning, since newly originated loans realize losses with a delay that mechanically alters the mapping from the allowance into near-term charge-offs. The second block adds the $EBLLP$ and deposit ratios, which absorb earnings capacity and funding structure. Banks with higher pre-provision earnings have greater scope to smooth income through the provision, and funding structure proxies for differences in lending models and credit-risk appetite. The third block adds current-quarter net charge-offs, which absorb serial correlation in realized losses, so that the interaction does not capture the persistence of current credit deterioration into future quarters. The fourth block adds the current change in non-performing loans, the observable deterioration measure central to the incurred-loss information set (Beatty and Liao, 2021), so that the allowance's incremental mapping into future losses is not proxying for delinquency information already observable at the allowance date.

Consistent with the first part of *H2*, the post-CECL interaction is positive and significant in every specification, at 0.0780 ($p < 0.01$) in the full specification on the all-eligible sample. The pre-CECL interaction is an order of magnitude smaller, at 0.0155 ($p < 0.10$), indicating that allowance informativeness varied little with risk-function conditions under the incurred-loss regime. Consistent with the second part of *H2*, the difference between the regime coefficients is positive and significant (0.0626, $p < 0.05$). The estimates are stable as controls enter. Across the four specifications and both samples, the difference ranges from 0.0624 to 0.0927, so the result does not depend on the choice of controls. The same-bank sample in Panel B produces a larger post-CECL interaction (0.1023, $p < 0.05$) and difference (0.0888, $p < 0.10$), indicating that the pattern is present within the fixed set of banks observed in both regimes and is not attributable to compositional change.

The magnitudes are economically meaningful. In the post-CECL regime, a one standard deviation increase in the trailing gap (0.48) raises the association between the allowance and four-quarter realized losses by $0.48 \times 0.0780 = 0.037$, so a dollar of allowance at a bank with risk-function conditions one standard deviation better corresponds to roughly 3.7 cents more of subsequently realized charge-offs. The same computation in the pre-CECL regime yields 0.7 cents. The sensitivity of allowance informativeness to risk-function conditions is therefore roughly five times larger under CECL than under the incurred-loss regime.

4.2.3. Does the Informativeness Pattern Withstand Scrutiny?

Table 5 traces the estimates across charge-off horizons. At the one-quarter horizon the interaction is indistinguishable from zero in both regimes and both samples. The pattern emerges at the two-quarter horizon, where the post-CECL interaction is 0.0388 ($p < 0.01$) and the difference is 0.0355 ($p < 0.05$) on the all-eligible sample, and strengthens further at four quarters. This profile is consistent with the measurement logic of a lifetime estimate. Losses anticipated by the allowance realize over an extended period, so a single quarter of charge-offs provides a weak benchmark against which even a well-constructed allowance shows little association, and the informativeness differential that *H2* predicts becomes visible only as realizations accumulate.

Table 6 reports inference and influence diagnostics for the same-bank estimates. At the two-quarter horizon, the results are robust on every dimension. The post-CECL interaction and the cross-regime difference remain significant under wild-cluster bootstrap inference with Webb weights ($p = 0.008$ and $p = 0.004$), every leave-one-out iteration produces a same-signed post-CECL coefficient, and 97 percent of iterations retain $p < 0.10$. At the four-quarter horizon, the leave-one-out distributions remain stable, with all 34 post-CECL iterations positive, but the wild bootstrap p-values rise to 0.217 and 0.376 against analytic values of 0.015 and 0.057. The four-quarter same-bank cells rest on 34 clusters, and inference at that cluster count is sensitive to the method used. The two-quarter results, which survive both inference approaches, therefore provide the appropriate anchor for the statistical conclusions, and the four-quarter estimates indicate that the economic pattern extends to the longer horizon.

Taken together, the tables show that the association between the allowance and subsequently realized credit losses depends on the operating condition of the risk function that produces the estimate, and that this dependence is substantially stronger under CECL than under the incurred-loss regime. Combined with the evidence in Section 4.1 that CECL adoption itself strains the risk-function workforce, the results indicate that the standard concentrates incremental demands on the function whose condition its output quality has come to depend on.

The regime comparison warrants a note on scope. The stronger dependence of allowance informativeness on risk-function conditions under CECL admits two readings. The dependence may reflect the measurement task the standard prescribes, since lifetime expected-loss estimation rests on forecasts and judgment and is therefore inherently more sensitive to the capacity of the workforce performing it than estimation anchored to observed deterioration. Alternatively, the dependence may reflect the particular way sample banks implemented that task over the sample period, with first-generation models, maturing validation routines, and staffing still adjusting to the new demands, conditions under which estimate quality would vary with workforce condition regardless of the task's inherent sensitivity to judgment. The design cannot separate the two. Every post-adoption observation carries both the prescribed task and its early implementation, so the regime indicator switches both at once, and the estimated difference is the composite of the two channels. The first reading implies a durable feature of expected-loss accounting, while the second implies a transition effect that may fade as models and processes mature, and the estimates in Tables 4 through 6 do not adjudicate between them. The estimates therefore describe CECL as banks have implemented it rather than the standard in an idealized form. This scope is sufficient for the question the paper asks as banks do not experience an accounting standard separately from its implementation.

5. Conclusions

This paper examines whether an information-intensive accounting standard affects the working conditions of the specialized employees who implement it and whether the condition of those employees is associated with the quality of the resulting estimate. CECL provides a useful setting because it assigns identifiable forecasting, modeling, and validation responsibilities to the risk function and produces an accounting estimate whose predictive validity can be evaluated using subsequently realized credit losses. By combining employee reviews classified by job function with bank regulatory filings, the study observes both the workforce involved in implementation and the accounting output that the workforce helps produce.

The first set of results indicates that CECL adoption places additional demands on risk-function employees. The within-bank work-life balance gap between risk and non-risk employees declines by 0.168 rating points following adoption, which equals 27 percent of the sample standard deviation and is equivalent to approximately three-quarters of the positive pre-adoption gap. The decomposition shows that the decline is attributable to risk-function work-life balance, while non-risk work-life balance remains unchanged. The dynamic estimates show little evidence of a differential trend during the clean pre-adoption period and persistently negative post-adoption estimates. The effect is also absent from the other Glassdoor rating dimensions and is not attributable to the omission of any individual BHC. Banks adopting in 2022 or later exhibit an even larger decline. This result reduces the concern that pandemic disruption accounts for the baseline estimate, although it does not eliminate the possibility that other cohort-specific conditions affect the magnitude for late adopters.

The second set of results indicates that risk-function work conditions are more closely associated with allowance informativeness under CECL. Under the incurred-loss model, the association between the allowance and future charge-offs varies little with prior risk-function work conditions. Under CECL, the allowance maps more strongly into subsequently realized losses when relative risk-function work-life balance is better. The difference between the two regimes is present across the control specifications and in both the all-eligible and same-bank samples. It is indistinguishable from zero at the one-quarter horizon, emerges after two quarters, and becomes economically larger as realized losses accumulate over four quarters. The two-quarter estimates remain significant under wild-cluster bootstrap inference and are stable when individual banks are excluded. The four-quarter estimates are also stable in sign and magnitude across leave-one-out iterations, although statistical inference in the same-bank sample is sensitive to the small number of clusters. The two-quarter evidence therefore provides the strongest basis for statistical inference, while the four-quarter estimates show that the economic pattern extends as additional losses are realized.

The findings contribute to research on the consequences of accounting standards by identifying employee working conditions as an implementation outcome that is not captured by observable investments in personnel and technology. They also contribute to research on the production of accounting estimates. Prior work shows that reporting outcomes depend on employee characteristics, information quality, and managerial judgment. This study shows that the time-varying condition of the specialized workforce producing an estimate is also associated with its predictive validity. The comparison across accounting regimes further indicates that the relevance of workforce conditions depends on the task assigned by the standard as banks implemented it. Work conditions are more closely associated with estimate validity when the accounting regime requires employees to convert forecasts and judgment into a recognized lifetime loss estimate. Whether this dependence reflects a durable feature of expected-loss accounting or the frictions of early implementation is a question the present sample cannot answer. Future studies with longer post-adoption panels offer a natural test, since a dependence rooted in implementation should fade as models and processes mature while a dependence rooted in the task itself should persist.

The findings carry a practical message for the assessment of information-intensive standards. The costs of such standards are typically measured in preparation expenditures, systems, and personnel, and the benefits in the informativeness of the resulting numbers. The evidence here indicates that part of the cost arrives as a deterioration in the working conditions of the employees who implement the standard, and that the benefit is not independent of that cost, since the predictive validity of the estimate is associated with the condition of the same workforce. A standard that strains its implementing workforce may therefore erode a portion of the informational benefit it was designed to produce.

References

Allen, F., Carletti, E., 2008. Mark-to-market accounting and liquidity pricing. Journal of accounting and economics 45, 358–378.

Altamuro, J., Beatty, A., 2010. How does internal control regulation affect financial reporting? Journal of accounting and Economics 49, 58–74.

Armstrong, C.S., Kepler, J.D., Larcker, D.F., Shi, S.X., 2024. Rank-and-file accounting employee compensation and financial reporting quality. Journal of Accounting and Economics 78.

Ashbaugh-Skaife, H., Collins, D.W., Kinney Jr, W.R., LaFond, R., 2008. The effect of sox internal control deficiencies and their remediation on accrual quality. The accounting review 83, 217–250.

Barth, M.E., 2006. Including estimates of the future in today's financial statements. Accounting Horizons 20, 271–285.

Beatty, A., Liao, S., 2011. Do delays in expected loss recognition affect banks' willingness to lend? Journal of accounting and economics 52, 1–20.

Beatty, A., Liao, S., 2014. Financial accounting in the banking industry: A review of the empirical literature. Journal of accounting and Economics 58, 339–383.

Beatty, A., Liao, S., 2021. What do analysts' provision forecasts tell us about expected credit loss recognition? The Accounting Review 96, 1–21.

Beaver, W.H., Engel, E.E., 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. Journal of accounting and economics 22, 177–206.

Bhojraj, S., Kleymenova, A., Liu, M., Lu, Y., Sengupta, P., 2025. Imperfect expectations in loan loss forecasts. Available at SSRN 4858012 .

Bischof, J., Rudolf, N., 2025. Manager characteristics and the informativeness of banks' loan loss provisioning: Bischof and rudolf. Review of Accounting Studies 30, 3677–3718.

Bonaldi, P., Liang, P.J., Yang, L., 2023. Did cecl improve banks' loan loss provisions and earnings quality during the covid-19 pandemic? Available at SSRN 4360356 .

Bonsall, S.B., Schmidt, B., Xie, B., 2025. Current expected credit loss (cecl) model and banks' external information environment. Available at SSRN 4236547 .

Bushman, R.M., Williams, C.D., 2012. Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking. Journal of accounting and economics 54, 1–18.

Call, A.C., Campbell, J.L., Dhaliwal, D.S., Moon Jr, J.R., 2017. Employee quality and financial reporting outcomes. Journal of Accounting and Economics 64, 123–149.

Callaway, B., Sant'Anna, P.H., 2021. Difference-in-differences with multiple time periods. Journal of Econometrics 225, 200–230.

Canals-Cerdá, J.J., 2026. Model risk under cecl: A consumer finance perspective. FRB of Philadelphia Working Paper .

Cheng, Q., Cho, Y.J., Yang, H., 2017. Financial reporting changes and the internal information environment: Evidence from sfas 142. Singapore Management University School of Accountancy Research Paper .

Christensen, H.B., Hail, L., Leuz, C., 2013. Mandatory ifrs reporting and changes in enforcement. Journal of accounting and economics 56, 147–177.

CSBS, 2021. Annual Survey of Community Banks. Survey Report. Conference of State Bank Supervisors. URL: `https://www.csbs.org/sites/default/files/2024-06/CBRC%20pubs/cb21publication_2021%20(1).pdf`. accessed 2025.

Daske, H., Hail, L., Leuz, C., Verdi, R., 2008. Mandatory ifrs reporting around the world: Early evidence on the economic consequences. Journal of accounting research 46, 1085–1142.

Dehaan, E., Li, N., Zhou, F.S., 2023. Financial reporting and employee job search. Journal of Accounting Research 61, 571–617.

Doyle, J., Ge, W., McVay, S., 2007. Determinants of weaknesses in internal control over financial reporting. Journal of accounting and Economics 44, 193–223.

Dyer, T., Lang, M., Stice-Lawrence, L., 2017. The evolution of 10-k textual disclosure: Evidence from latent dirichlet allocation. Journal of Accounting and Economics 64, 221–245.

FASB, 2016. Accounting Standards Update No. 2016-13: Financial Instruments—Credit Losses (Topic 326). Technical Report. Financial Accounting Standards Board. Norwalk, CT.

Fiechter, P., Hitz, J.M., Lehmann, N., 2022. Real effects of a widespread csr reporting mandate: Evidence from the european union's csr directive. Journal of Accounting Research 60, 1499–1549.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of econometrics 225, 254–277.

Green, T.C., Huang, R., Wen, Q., Zhou, D., 2019. Crowdsourced employer reviews and stock returns. Journal of Financial Economics 134, 236–251.

Hilary, G., Tian, X., Yu, M., 2025. Does meeting financial expectations boost employee satisfaction? The Accounting Review 100, 277–302.

Huang, K., Li, M., Markov, S., 2020. What do employees know? evidence from a social media platform. The Accounting Review 95, 199–226.

Khavis, J.A., Krishnan, J., 2021. Employee satisfaction and work-life balance in accounting firms and audit quality. Auditing: A Journal of Practice & Theory 40, 161–192.

Kim, S., Kim, S., Kleymenova, A., Li, R., 2026. Current expected credit losses (cecl) standard and banks' information production. The Accounting Review , 1–34.

Laux, C., Leuz, C., 2010. Did fair-value accounting contribute to the financial crisis? Journal of economic perspectives 24, 93–118.

Mahieux, L., Sapra, H., Zhang, G., 2023. Cecl: Timely loan loss provisioning and bank regulation. Journal of Accounting Research 61, 3–46.

Plantin, G., Sapra, H., Shin, H.S., 2008. Marking-to-market: panacea or pandora's box? Journal of accounting research 46, 435–460.

Roh, Y., 2025. Internal accounting hiring and operational efficiency: Evidence from the implementation of asc 842. The Accounting Review 100, 317–344.

Shroff, N., 2017. Corporate investment and changes in gaap. Review of Accounting Studies 22, 1–63.

Watts, R.L., 2003. Conservatism in accounting part i: Explanations and implications. Accounting horizons 17, 207–221.

Yang, L., 2022. An information quality-based explanation for loan loss allowance inadequacy during the 2008 financial crisis. Journal of Accounting and Economics 73, 101433.

Figures and Tables

Table 1: **Summary Statistics**

Panel A: H1 WLB Gap Sample

Variable	N	Mean	SD	p10	p50	p90
WLB gap	1,960	0.25	0.63	-0.45	0.27	1.00
WLB (risk function)	1,960	3.76	0.69	2.90	3.81	4.50
WLB (non-risk function)	1,960	3.51	0.38	3.04	3.53	3.96
Post	1,960	0.67	0.47	0	1	1
Log(Assets)	1,960	18.13	1.56	16.00	18.07	20.54
Capital ratio	1,960	0.11	0.03	0.08	0.11	0.15
EBLLP ratio	1,960	0.02	0.03	0.01	0.02	0.05
Deposit ratio	1,960	0.69	0.18	0.45	0.77	0.85
N (risk reviews)	1,960	21.56	37.10	3	7	60.5
N (non-risk reviews)	1,960	589.74	1,012.09	29	139	1,731

Panel B: H2 Predictive-Validity Sample (four-quarter horizon)

Variable	N	Mean	SD	p10	p50	p90
Future NCO, t+1–t+4 (%)	660	0.60	0.85	0.02	0.27	2.26
Allowance (%)	660	1.57	1.26	0.34	1.24	3.87
Prior-four-quarter WLB gap	660	0.25	0.48	-0.26	0.24	0.78
Current-quarter NCO (%)	660	0.14	0.22	0.00	0.06	0.52
Current change in NPL (%)	660	0.00	0.16	-0.12	0.00	0.13
Log(Assets)	660	19.10	1.34	16.90	19.06	20.88
Capital ratio	660	0.11	0.03	0.08	0.10	0.14
Loan growth (%)	660	1.48	4.47	-2.54	0.92	5.53
EBLLP ratio	660	0.03	0.02	0.01	0.02	0.06
Deposit ratio	660	0.65	0.18	0.38	0.72	0.82

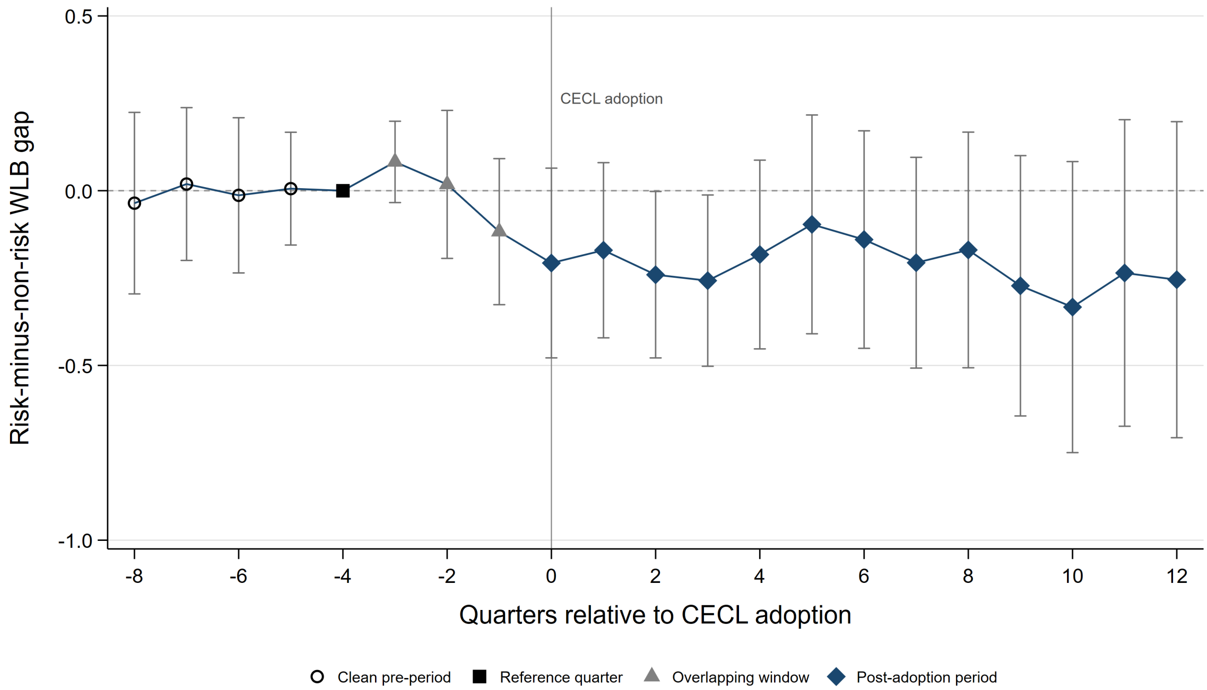
This table reports summary statistics for the two analysis samples. The unit of observation in both panels is the bank-quarter. Panel A reports descriptive statistics for the H1 estimation sample, defined as bank-quarters with at least three function-classifiable Glassdoor reviews in both the risk and non-risk categories. Panel B reports descriptive statistics for the H2 predictive-validity sample at the four-quarter horizon, defined as bank-quarters of same-bank sample BHCs for which the prior-four-quarter WLB gap and the full four-quarter charge-off window are available. *WLB_gap* is the difference between the average work-life balance rating of risk-function reviews and the average work-life balance rating of non-risk-function reviews. Reviews are classified by the position reported by the reviewer. Reviews without reported positions cannot be assigned to the risk function and are included in the non-risk benchmark. Because Glassdoor reviews are typically posted with a delay after the underlying employment experience, WLB measures for H1 aggregate all reviews posted in the focal quarter and the three calendar quarters immediately following it. For a bank-quarter ending March 31, 2020, this draws reviews posted between January 1, 2020 and December 31, 2020. The *prior-four-quarter WLB gap* is the backward-looking variant, aggregating reviews posted during the four quarters preceding the focal quarter. *Post* equals one on and after the BHC's CECL adoption quarter. *Future NCO, t+1–t+4* is net charge-offs accumulated over the four quarters following the focal quarter, scaled by gross loans. *Allowance* is the allowance for credit losses divided by gross loans at quarter end. *Current-quarter NCO* is net charge-offs in the focal quarter scaled by gross loans, and *current change in NPL* is the quarterly change in non-performing loans scaled by gross loans. *Loan growth* is the quarterly growth rate of gross loans. *Log(Assets)* is the natural logarithm of total consolidated assets. *EBLLP ratio* is earnings before loan loss provisions divided by total assets. *Deposit ratio* is total deposits divided by total assets. Variables marked (%) are expressed in percent. All financial variables are constructed from FR Y-9C filings. Financial variables are winsorized at the 1st and 99th percentiles.

Table 2: **CECL Adoption and Employee Work-Life Balance**

	Panel A: Full H1 sample			Panel B: Late adopters		
	(1) WLB Gap	(2) Risk WLB	(3) Non-risk WLB	(4) WLB Gap	(5) Risk WLB	(6) Non-risk WLB
Post	-0.1680**	-0.1684*	-0.0004	-0.6055***	-0.5428**	0.0627
	(0.0818)	(0.0915)	(0.0401)	(0.2115)	(0.2157)	(0.0613)
Log(Assets)	-0.3421**	-0.3506**	-0.0085	-0.5684*	-0.3306	0.2378
	(0.1720)	(0.1593)	(0.1256)	(0.2846)	(0.2474)	(0.1801)
Capital ratio	-1.9493	-2.2469	-0.2976	-2.8189	-1.6081	1.2108
	(1.8503)	(2.4270)	(1.5331)	(3.3380)	(3.7723)	(1.9767)
EBLLP ratio	-2.1321*	-2.0131*	0.1190	-3.0812**	-2.6488	0.4323
	(1.1377)	(1.1988)	(0.5709)	(1.3182)	(1.8879)	(1.0787)
Deposit ratio	0.8519	0.9557	0.1038	-2.9021*	-1.9845	0.9176
	(0.7369)	(0.8459)	(0.3565)	(1.6191)	(1.9065)	(0.6230)
Observations	1,960	1,960	1,960	378	378	378
BHCs	128	128	128	35	35	35
R^2	0.0203	0.0224	0.0007	0.0964	0.0581	0.0503
BHC FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports difference-in-differences estimates of the effect of CECL adoption on employee work-life balance. The unit of observation is the bank-quarter, and the sample requires at least three risk-function and three non-risk reviews per bank-quarter cell. Panel A uses the full H1 estimation sample. Panel B restricts the sample to late adopters, defined in Section 3.2 as BHCs with adoption quarter in 2022Q1 or later, and their never-treated controls. The dependent variable is *WLB_gap* in columns (1) and (4), the risk-function *WLB* component in columns (2) and (5), and the non-risk component in columns (3) and (6). *Post* equals one on and after the BHC's CECL adoption quarter. Within each panel, all three columns are estimated on an identical sample, so the estimate in the gap column equals the difference between the two component-column estimates by construction. The Panel B gap estimate remains significant at the five percent level under wild-cluster bootstrap inference with Webb weights (p = 0.025, 9,999 replications). Variable definitions are as in Table 1. Standard errors clustered by BHC are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Fig. 1. **Dynamic Effects of CECL Adoption on the WLB Gap**



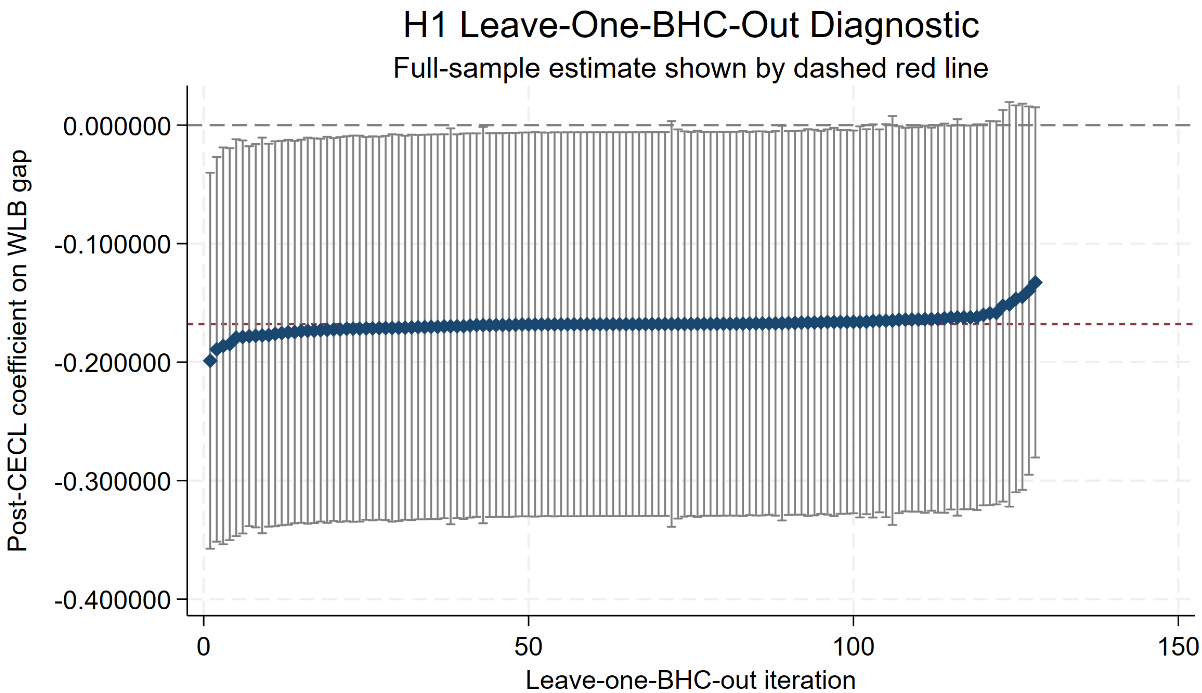
This figure plots event-study estimates of the effect of CECL adoption on WLB_Gap, obtained by replacing $Post_{i,t}$ in Equation (2) with indicators for event quarters $k = -8$ through $k = 12$ relative to each BHC's adoption quarter. The specification includes the bank-level controls, BHC fixed effects, and calendar-quarter fixed effects of Equation (2). Because the review window for quarter t pools reviews posted through quarter $t + 3$, the last measurement drawing exclusively on pre-adoption reviews is dated four quarters before adoption. The reference period is therefore $k = -4$ (solid square). Hollow circles mark the clean pre-period ($k = -8$ through $k = -5$), grey triangles mark the overlapping window ($k = -3$ through $k = -1$), in which measurements partially pool post-adoption reviews, and solid diamonds mark the post-adoption period. Vertical bars are 95 percent confidence intervals based on standard errors clustered by BHC.

Table 3: **CECL Adoption and Alternative Employee Rating Gaps**

	Dependent variable: risk-minus-non-risk rating gap					
	(1) WLB	(2) Overall	(3) Career	(4) Compensation	(5) Culture	(6) Leadership
Post	-0.1680**	-0.0266	-0.0189	-0.0787	0.0323	-0.0295
	(0.0818)	(0.0653)	(0.0967)	(0.0758)	(0.0879)	(0.1079)
Observations	1,960	1,960	1,960	1,960	1,960	1,960
BHCs	128	128	128	128	128	128
R^2	0.0203	0.0042	0.0159	0.0127	0.0056	0.0061
Bank controls	Yes	Yes	Yes	Yes	Yes	Yes
BHC FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Average placebo coefficient	−0.0243					
WLB − average placebo	−0.1437*					
p-value of difference	0.094					

This table reports difference-in-differences estimates of the effect of CECL adoption on within-bank, risk-minus-non-risk rating gaps across the six Glassdoor rating dimensions. Each column estimates Equation (2) with the dependent variable constructed as in Section 3.1.5, replacing the work-life balance rating with the rating dimension indicated in the column header. Column (1) uses the work-life balance gap and repeats the baseline estimate. Columns (2) through (6) use the overall rating, career opportunities, compensation and benefits, culture and values, and senior leadership, respectively. The sample consists of bank-quarters with at least three risk-function and three non-risk reviews reporting all six rating dimensions, which coincides with the H1 estimation sample. Bank controls are as defined in Table 2, and coefficient estimates on the controls are suppressed. The bottom rows are computed from a stacked regression that pools the six outcomes and interacts *Post*, the controls, and the fixed effects with outcome indicators, with standard errors clustered by BHC. The rows report the average of the five placebo coefficients from columns (2) through (6), the difference between the column (1) coefficient and that average, and the *p*-value from a test that the difference equals zero. Standard errors clustered by BHC are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Fig. 2. **Leave-One-Out Distribution of the Baseline Estimate**



This figure plots the coefficient on *Post* from Equation (2), re-estimated 128 times with one BHC omitted per iteration, sorted by estimate magnitude. Vertical bars are 95 percent confidence intervals based on standard errors clustered by BHC. The horizontal line marks the full-sample estimate of −0.168 from column (1) of Table 2. Estimates range from −0.199 to −0.133. All 128 estimates are negative and significant at the 10 percent level, and 88 percent are significant at the 5 percent level.

Table 4: **Workforce Conditions and Allowance Informativeness**

	(1) Baseline Controls	(2) + Non-credit	(3) + Current NCO	(4) + Change in NPL
Panel A: All eligible banks				
Pre: Allowance × WLB_Gap	0.0126	0.0166*	0.0156*	0.0155*
	(0.0093)	(0.0086)	(0.0090)	(0.0087)
Post: Allowance × WLB_Gap	0.0856**	0.0790**	0.0792***	0.0780***
	(0.0382)	(0.0364)	(0.0293)	(0.0293)
Post − Pre	0.0730**	0.0624*	0.0635**	0.0626**
	(0.0360)	(0.0324)	(0.0296)	(0.0302)
Observations	1,009	1,009	1,009	1,009
BHCs	80	80	80	80
R^2	0.1218	0.1595	0.2560	0.2657
Panel B: Same-bank sample				
Pre: Allowance × WLB_Gap	0.0137	0.0136	0.0143	0.0136
	(0.0096)	(0.0096)	(0.0098)	(0.0090)
Post: Allowance × WLB_Gap	0.0897**	0.0935**	0.1070***	0.1023**
	(0.0394)	(0.0356)	(0.0364)	(0.0397)
Post − Pre	0.0760**	0.0799**	0.0927**	0.0888*
	(0.0326)	(0.0295)	(0.0418)	(0.0449)
Observations	660	660	660	660
BHCs	34	34	34	34
R^2	0.2104	0.2401	0.4000	0.4197
BHC FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes

This table reports estimates of the association between the allowance and subsequently realized credit losses, and its dependence on risk-function work conditions, before and after CECL adoption. The dependent variable is cumulative net charge-offs over quarters $t+1$ through $t+4$, scaled by gross loans at quarter t. *Allowance* is measured at quarter t, and *WLB_Gap* over quarters $t-4$ through $t-1$ using a regime-pure window. Each column reports the coefficients on Allowance × *WLB_Gap* in the pre-CECL and post-CECL regimes, and their difference, from a stacked model in which the allowance, *WLB_Gap*, their interaction, the controls, and the fixed effects are fully interacted with the regime indicator. Columns add controls cumulatively as indicated. Panel A includes all eligible bank-quarters, and Panel B restricts the sample to BHCs represented in both regimes. Within each panel, all four columns use the same complete-observation sample. Coefficients on the lower-order terms and controls are suppressed. Standard errors clustered by BHC are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: **Workforce Conditions and Allowance Informativeness across Future-Loss Horizons**

Future NCO horizon	(1) $t+1$	(2) $t+1$ to $t+2$	(3) $t+1$ to $t+4$
Panel A: All eligible banks			
Pre: Allowance × WLB_Gap	0.0010	0.0033	0.0155*
	(0.0033)	(0.0062)	(0.0087)
Post: Allowance × WLB_Gap	0.0079	0.0388***	0.0780***
	(0.0076)	(0.0112)	(0.0293)
Post − Pre	0.0070	0.0355**	0.0626**
	(0.0103)	(0.0141)	(0.0302)
Observations	1,193	1,122	1,009
BHCs	94	86	80
R^2	0.4248	0.3854	0.2657
Panel B: Same-bank sample			
Pre: Allowance × WLB_Gap	0.0004	0.0024	0.0136
	(0.0035)	(0.0066)	(0.0090)
Post: Allowance × WLB_Gap	0.0067	0.0365***	0.1023**
	(0.0114)	(0.0110)	(0.0397)
Post − Pre	0.0063	0.0341***	0.0888*
	(0.0143)	(0.0123)	(0.0449)
Observations	843	802	660
BHCs	41	39	34
R^2	0.4141	0.3986	0.4197
Full Table 4 controls	Yes	Yes	Yes
BHC FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes

This table reports the estimates of Table 4 across accumulation horizons for realized credit losses. The dependent variable in each column is cumulative net charge-offs over the indicated horizon, scaled by gross loans at quarter t. Column (3) reproduces Table 4, column (4), exactly. *Allowance* is measured at quarter t, and *WLB_Gap* over quarters $t-4$ through $t-1$ using a regime-pure window. Controls are size, capitalization, loan growth, EBLLP, deposits, current NCO, and the current change in NPL, each fully interacted with the pre- and post-CECL regimes. Sample sizes decline with the horizon because longer accumulation windows require additional quarters of charge-off data within a single regime. Panel A includes all eligible bank-quarters, and Panel B restricts the sample to BHCs represented in both regimes. Coefficients on the lower-order terms and controls are suppressed. Standard errors clustered by BHC are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: **Inference and Influence Diagnostics for Allowance Informativeness**

	(1)	(2)
Future NCO horizon	$t+1$ to $t+4$	$t+1$ to $t+2$
Panel A: Post-CECL Allowance × WLB_Gap coefficient		
Estimate	0.1023	0.0365
Clustered standard error	(0.0397)	(0.0110)
Clustered p-value	0.015	0.002
Wild-cluster bootstrap p-value	0.217	0.008
Leave-one-out coefficient range	[0.0571, 0.1467]	[0.0230, 0.0477]
Leave-one-out same-sign share	1.00	1.00
Leave-one-out share $p < 0.10$	0.97	0.97
Panel B: Post − Pre change in Allowance × WLB_Gap coefficient		
Estimate	0.0888	0.0341
Clustered standard error	(0.0449)	(0.0123)
Clustered p-value	0.057	0.009
Wild-cluster bootstrap p-value	0.376	0.004
Leave-one-out coefficient range	[−0.0023, 0.1397]	[0.0209, 0.0675]
Leave-one-out same-sign share	0.97	1.00
Leave-one-out share $p < 0.10$	0.94	0.97
Observations	660	802
BHCs	34	39
Full Table 4 controls	Yes	Yes
BHC FE	Yes	Yes
Quarter FE	Yes	Yes

This table reports inference and influence diagnostics for the same-bank specifications of Table 5 at the four-quarter and two-quarter charge-off horizons. Each displayed estimate and sample count matches Table 5 exactly. Clustered p-values use standard errors clustered by BHC. Wild-cluster bootstrap p-values impose the null and use Webb weights with 9,999 replications. Leave-one-out ranges and shares summarize estimates obtained by omitting one BHC at a time. Controls are the full set of Table 4, fully interacted with the pre- and post-CECL regimes.

Appendix A. Glassdoor Data Construction

A central data-construction challenge is mapping bank holding companies in the regulatory panel (Section 3.2) to their corresponding Glassdoor pages. The mapping is non-trivial for two reasons. First, BHCs file regulatory reports under their legal parent name (e.g., "JPMorgan Chase & Co."), while their Glassdoor pages are organized under brand names that may use the parent name, a subsidiary bank name, or both. Second, BHCs hold multiple commercial-bank subsidiaries with their own Glassdoor pages, and reviews posted to subsidiary pages must be attributed to the parent BHC for consistency with the regulatory data.

I address these challenges through a multi-step matching pipeline. For each BHC parent that meets the size threshold described in Section 3.2, I enumerate all subsidiary commercial banks identified through the WRDS Bank Structure relationships file. The resulting control sheet contains one row per parent-subsidiary pair, augmented with a self-row for each parent because some Glassdoor pages use the parent BHC's name directly. Each name is then cleaned in two passes to produce candidate search variants. The *light* pass strips standard legal-entity and geographic suffixes[6] while preserving brand-identifying terms[7]. The *aggressive* pass additionally strips the brand-identifying terms, subject to a minimum-length floor of four characters that prevents over-stripping of short bank names (e.g., *Ally*, *Citi*, *USAA*). Single-word stems that survive cleaning are augmented with the suffix "Bank" (e.g., "*Apple*" to "*Apple Bank*"), which provides the banking-context cue needed to disambiguate from non-financial entities of the same name. Abbreviation periods are preserved throughout so that "*U.S. Bank*" does not collapse to "*US Bank*".

For each parent-subsidiary pair, I query the cleaned variants in sequence, beginning with the light variant, followed by the original name, and ending with the aggressive variant. Each

[6] E.g., *N.A.*, *Inc.*, *Corp.*, *LLC*, *L.L.C.*, *Ltd.*, *USA*, *(USA)*, *FSB*, *SSB*, *Federal Savings Bank*, *National Association*, and similar terms.

[7] E.g., *Bank*, *Bancorp*, *National Bank*, *Trust*.

query retrieves the top three Glassdoor candidate pages, and each candidate is scored with three fuzzy-matching metrics from the `rapidfuzz` library, namely the Levenshtein ratio, the token-sort ratio, and the token-set ratio. Three conservative filters minimize false-positive matches. A *first-token gate* requires the first content-bearing token of the query to match the first content-bearing token of the candidate, with a prefix tolerance of four or more characters when the tokens are not exactly equal. The gate prevents spurious matches such as "*Citizens Bank*" matching "*First Citizens Bank*." Common stop words such as *The*, *A*, and *An* are skipped when identifying the first content-bearing token. A *length-ratio floor* of 0.50 rejects candidates more than twice as long as the query, protecting against short queries matching substantially longer candidates that share token overlap. A *score threshold* requires the best of the three fuzzy scores to exceed 95 on the standard 0–100 scale. For queries of five characters or fewer, I require a Levenshtein-ratio score of 95 or higher specifically, since token-based metrics are unreliable on short strings.

A candidate that survives all filters is accepted as the Glassdoor page for the pair. Pairs unmatched in the first round are re-queried using the original BHC parent name, and the remaining unmatched pairs are manually reviewed to identify pages the automated procedure missed due to substantially different brand naming conventions. Matched pairs are then deduplicated within parent BHC so that each parent-page pair appears at most once. Reviews from any page mapped to more than one parent BHC are attributed to all matched parents, which preserves data for the small number of shared-brand cases without arbitrarily assigning the page to a single parent. Finally, I collect the full review history for each matched page, retaining the rating sub-dimensions, job title, posting date, and employment status for each review. Each collected file is verified through three integrity checks covering existence, valid parsing, and non-empty content, and pages failing any check are flagged for re-collection until the full sample passes.

Table A1: **Pre-Adoption WLB Determinants and Gap Validation**

	Pre-adoption risk-function WLB			*Pre-adoption WLB_Gap*	
	(1)	(2)	(3)	(4)	(5)
Log(Assets)	−0.0430	−0.2199	−0.0751	−0.0184	−0.0584
	(0.0542)	(0.2271)	(0.2029)	(0.0409)	(0.1994)
Capital ratio		−4.6998	−2.7274		−2.4998
		(3.6681)	(2.8568)		(2.8804)
EBLLP ratio		1.8883	0.9917		0.8882
		(3.5873)	(3.4072)		(3.4732)
Deposit ratio		0.0787	0.2042		0.2187
		(0.6237)	(0.4727)		(0.4714)
Log(Employees)		0.1805	0.0633		0.0497
		(0.2056)	(0.1818)		(0.1825)
Pre-adoption non-risk WLB			0.8965***		
			(0.2607)		
BHCs	67	67	67	67	67
Adjusted R^2	−0.0046	−0.0164	0.2397	−0.0126	−0.0515

This table reports cross-sectional OLS regressions estimated at the BHC level using pre-adoption averages of all variables, across the 67 BHCs with sufficient pre-adoption review coverage. The dependent variable in columns (1) through (3) is the bank-level pre-adoption average of risk-function *WLB*, weighted by the number of risk-function reviews in each quarter. The dependent variable in columns (4) and (5) is the pre-adoption average of *WLB_Gap*, constructed analogously. *Log(Employees)* is the pre-adoption average of the natural logarithm of total employees, and *Pre-adoption non-risk WLB* is the bank-level pre-adoption weighted average of non-risk-function *WLB*. Remaining variables are pre-adoption averages of the controls defined in Table 1. Columns (1) and (4) include size only, columns (2) and (5) add the bank-level controls, and column (3) adds non-risk *WLB*. Standard errors in parentheses are heteroskedasticity-robust. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A2: **Staggered-Adoption Estimates of the Effect of CECL Adoption on the WLB Gap**

	Estimate	Std. error	p-value	95% CI
ATT, adoption-quarter measurement	-0.3478*	(0.1811)	0.055	$[-0.7028,\ 0.0072]$
Average ATT, event quarters 0 through +4	-0.2869**	(0.1286)	0.026	$[-0.5390,\ -0.0349]$
Observations			801	
BHCs			128	

This table reports estimates of the effect of CECL adoption on WLB_Gap from the Callaway and Sant'Anna (2021) estimator, which restricts comparisons to not-yet-treated and never-treated banks. Because the WLB measure for quarter t pools reviews posted through quarter $t+3$, each bank's effective adoption date is shifted three quarters earlier, so that no measurement containing post-adoption reviews enters a control group. Event quarters are stated relative to the actual adoption quarter. The first row reports the average treatment effect for the measurement dated at the adoption quarter, whose review window spans the first four post-adoption quarters. The second row reports the average treatment effect across event quarters zero through four. Estimates use the regression-based variant of the estimator without covariates, so identification relies on parallel trends in the gap with respect to the shifted adoption dates. The sample comprises bank-quarters meeting the three-review requirement of Section 3.1.5 with a defined adoption cohort. Standard errors clustered by BHC are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.