FEDERAL RESERVE BANK
PHILADELPHIA

Working Papers
RESEARCH DEPARTMENT

CECL Implementation and Model Risk in Uncertain Times
An Application to Consumer Finance

José J. Canals-Cerdá

Federal Reserve Bank of Philadelphia Supervision, Regulation, and Credit Department

WP 24-03

PUBLISHED
February 2024

FRB CFI PHILADELPHIA

ISSN: 1962-5361

Disclaimer: This Philadelphia Fed working paper represents preliminary research that is being circulated for discussion purposes. The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. Philadelphia Fed working papers are free to download at: https://philadelphiafed.org/research-and-data/publications/working-papers.

DOI: https://doi.org/10.21799/frbp.wp.2024.03

CECL Implementation and Model Risk in Uncertain Times: An Application to Consumer Finance

By José J. Canals-Cerdá[1]

Abstract

I examine the challenges of economic forecasting and model misspecification errors confronted by financial institutions implementing the novel current expected credit loss (CECL) allowance methodology and its impact on model risk and bias in CECL projections. We document the increased sensitivity to model and macroeconomic forecasting error of the CECL framework with respect to the incurred loss framework that it replaces. An empirical application illustrates how to leverage simple machine learning (ML) strategies and statistical principles in the design of a nimble and flexible CECL modeling framework. We show that, even in consumer loan portfolios with tens of millions of loans, like mortgage, auto, or credit card portfolios, one can develop, estimate, and deploy an array of models quickly and efficiently, and without a forecasting performance penalty. Drawing on more than 20 years of auto loans data and the experience from the Great Recession and the COVID-19 pandemic, we leverage basic econometric principles to identify strategies to deal with biased model projections in times of high economic uncertainty. We advocate for a focus on resiliency and adaptability of models and model infrastructures to novel shocks and uncertain economic conditions.

JEL Codes: G01, G21, G28, G50, M41

Keywords: CECL, Allowance for Loan and Lease Losses, Accounting Regulations, Model Risk

[1] The views expressed here are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Earlier versions of this research were presented at the Interagency Risk Quantification Conference of financial regulators (2022), RiskMinds America (2022), RiskMinds International (2022), RiskMinds Edge (2022), the Second Model Risk Management International Association Best Practices Virtual Summit (2022), and the Credit Scoring and Credit Control Conference XVIII (2023). Comments from conference participants and colleagues, and of Gerald Rama in particular, are greatly appreciated. All remaining errors are my own.

I. INTRODUCTION

The Allowance for Loan and Lease Losses (ALLL) is an estimate of credit losses used to reduce the book value of loans and leases to the amount that a bank expects to collect. The ALLL is of great importance to bank management, investors, and regulators. In the aftermath of the 2007–09 Great Recession, the incurred loss methodology was criticized for its "failure to fully recognize existing credit losses earlier in the credit cycle."[2] In an attempt to address identified shortcomings with the existing approach, in 2016, the Financial Accounting Standards Board (FASB) introduced the current expected credit loss (CECL) framework, a novel approach for computing the ALLL.

CECL enhances the loan loss provisioning methodology by considering lifetime loan losses and by incorporating forward-looking forecasts of economic conditions.[3] The novel CECL methodology became effective for most U.S. Securities and Exchange Commission (SEC) filers after December 15, 2019. The group of initial CECL adopters included the most complex financial institutions in the United States.[4]

The focus of the CECL framework on the projection of lifetime loan losses — and its reliance on economic forecasts — increases the sensitivity of the allowance to economic forecasting and model error. We document the experience of CECL adopter and nonadopter financial institutions during the initial adoption phase, which coincided with the early months of the COVID-19 crisis. Allowances from CECL adopters increased faster early in the pandemic and reached a much higher peak when compared with nonadopters. In contrast, charge-off rates during this period decreased with respect to the already record-low levels of recent years, generating an historically unparalleled gap between allowances and charge-offs. Financial institutions faced with highly unusual macroeconomic conditions and underperforming models resorted to judgment-based adjustments to their provisioning projections. This experience underscores the significant challenges to the CECL framework in times of highly uncertain economic environments typical of

[2] See the Financial Stability Forum (2009) report.

[3] Additional information is available at www.federalreserve.gov/supervisionreg/topics/faq-new-accounting-standards-on-financial-instruments-credit-losses.htm.

[4] SEC filers, with the exception of smaller reporting companies, were required to adopt CECL on January 1, 2020, and other companies were required to adopt CECL on January 1, 2023.

crisis episodes. It is important to draw lessons from past crises and to take appropriate steps to strengthen the important allowance framework.

In order to analyze potential CECL challenges in times of high economic uncertainty, we construct a simple modeling framework based on sound statistical principles. Our approach is deployed in two steps. In the first, we utilize simple ML techniques to segment a loan portfolio into sets of loans with broadly homogeneous risk profiles, and in the second, we employ standard statistical methods across segments to model lifetime CECL projections, conditional on macroeconomic forecasts. Our framework is simple without compromising performance. It allows for quick and easy development, redesign, and deployment of models, irrespective of the size of the portfolio considered and, because of its simplicity, it can easily accommodate multiple models. For these reasons, the approach is particularly valuable in consumer finance portfolios, like personal loans, student loans, mortgages, or credit card loans, where the typical loan portfolio can comprise many millions of loans. We consider an application for auto loans, which have not previously received the same level attention as other types of loans in the consumer finance literature.

Leveraging the simplicity of our modeling framework, we analyze potential problems of forecasting bias and model misspecification that can impact CECL implementation during periods of high economic uncertainty. We analyze more than 20 years of portfolio performance, encompassing the Great Recession and the COVID-19 pandemic. We observe that model performance deteriorates significantly in periods of crisis with associated uncharted economic environments. However, model performance can improve significantly when models are re-estimated with additional data that include some exposure to the novel macroeconomic environment. We also observe that model performance may not deteriorate homogeneously across risk segments, so certain segments can potentially act as an early warning for more widespread underperformance. We observe that learning from a variety of model specifications can be fruitful, particularly in times of crisis. Guidance from econometric theory can also offer insights that can help alleviate the impact of model misspecification error. Furthermore, CECL long-run projections by design average out economic cycles to a certain extent, although short-term economic conditions are a key driver of CECL allowances.

The next section introduces the CECL framework in greater detail and provides a brief overview of the relevant literature. Section three analyzes the initial impact of CECL implementation on the allowances of financial institutions as well as the differential impact of the pandemic on the allowances across CECL adopters and nonadopters. Section four analyzes conceptually the impact of economic forecasting error and model misspecification error on CECL allowances. Section five introduces a simple empirical framework for CECL implementation with an application for auto loans as a particular example of a consumer finance portfolio. Section six discusses empirical findings and lessons learned on how to mitigate potential CECL projection bias in times of high economic uncertainty. Section seven concludes. An appendix provides some additional background on regulatory guidance regarding CECL implementation.

II. The CECL Framework: A Brief Introduction

ALLL is an estimate of credit losses within a bank's portfolio of loans and leases used to reduce the book value of the portfolio to the amount that the bank expects to collect. Over the last 40 years, the standard ALLL approach under U.S. generally accepted accounting principles has been the incurred loss methodology. Under this approach, the allowance is a valuation reserve established and maintained to cover losses that are probable and estimable as of the reserve calculation date.[5] Thus, potential future losses that are not deemed probable should not to be incorporated, even if it is reasonable to expect that losses will be realized that are not viewed as probable at this time, perhaps as a result of future credit risk deterioration or other factors, like a change in expected future economic conditions.

By delaying the recognition of loan losses during the Great Recession, the incurred loss framework contributed to the buildup of allowances amid the stress period. As a result, it could have contributed to a decrease in bank lending and to the overall procyclicality of the financial system. These concerns were identified by the Financial Stability Forum (FSF) in its 2009 report on procyclicality in the financial system.[6] The FSF indicated that earlier recognition of loan losses could help lessen procyclicality while enhancing the consistency of information provided to

[5] See Statement of Financial Accounting Standards 114.

[6] See the Financial Stability Forum (2009) report.

investors. Various stakeholders requested that accounting standard-setters work to enhance the loan loss provisioning methodology to incorporate forward-looking information.[7] In June 2016, the Financial Accounting Standards Board (FASB) issued an accounting standard update (ASU 2016-13), which introduced the new ALLL methodology, the CECL framework.

CECL represents a significant departure from the incurred loss standard that it replaces. It is built on the notion of forward-looking estimates of expected loan credit loss based on relevant information about past events, including historical experience, current conditions, and reasonable and supportable forecasts that affect the collectability of loans. Under CECL, institutions are expected to reserve for lifetime losses on loans at the time the loans are originated.[8] It also requires enhanced disclosures.[9] CECL is nonprescriptive about the models and loss projection methodology that should be employed, and about the economic projections that should be considered. However, it prescribes reasonable and supportable forecasts over a reasonable time frame, which can be less than the life of the loan, and convergence to long-run economic conditions after that. At a very high level, CECL considers lifetime losses on a static portfolio. However, portfolio-specific characteristics can impact the analysis of losses.[10] Specifically, in the case of credit card portfolios or similar portfolios where the bank can unconditionally cancel future line draws, CECL does not consider future drawdowns when accounting for potential future losses (Canals-Cerda, 2020). Intuitively, both the incurred loss approach and the CECL framework impose specific restrictions on the ALLL. Specifically, the incurred loss framework bars the recognition of losses beyond incurred losses, while the CECL framework requires the recognition of expected future lifetime losses under some additional assumptions. Methodological or regulatory constraints on the ALLL do not directly impact the amount or timing of realized losses.[11]

[7] Additional information can be found at www.federalreserve.gov/supervisionreg/topics/faq-new-accounting-standards-on-financial-instruments-credit-losses.htm.

[8] Banking regulators have issued Implementation and transition guidance. See the Board of Governors of the Federal Reserve System (BOG), www.federalreserve.gov/supervisionreg/topics/accounting.htm or www.federalreserve.gov/newsevents/pressreleases/files/bcreg20200826a2.pdf for recent guidance.

[9] CECL applies to every organization required to issue financial statements in compliance with U.S. GAAP. Following U.S. GAAP is required by the Federal Deposit Insurance Act, which notes that all insured depository institutions are required to be uniform and consistent with GAAP. FDI Act – SEC 37(a)(2)(A). Banks are likely to experience the largest implementation burden.

[10] See www.federalreserve.gov/supervisionreg/topics/faq-new-accounting-standards-on-financial-instruments-credit-losses.htm.

[11] See Question 3 in www.federalreserve.gov/supervisionreg/topics/faq-new-accounting-standards-on-financial-instruments-credit-losses.htm.

A. Research on CECL

The primary challenges highlighted in the relevant literature before CECL implementation were on procyclicality, first day impact, and CECL's impact on lending. Other practical concerns were related to the complexity of the CECL framework and implementation burdens. Regulators have attempted to address some of these practical concerns with concrete, and simple, methodological frameworks suitable for less complex financial institutions, as we discuss in an appendix.

A recent Bank of International Settlements (BIS) working paper (WP-39) conducted a review of the literature on CECL and the International Financial Reporting Standard 9 (IFRS 9). It reviewed 37 papers, with a special focus on the topic of the procyclicality of loss provision. It differentiated between two main forms of procyclicality. The concept of procyclicality of more pressing interest to policymakers considers a causal feedback loop between the allowance framework and the economic cycle. An alternative interpretation of procyclicality is the statistical comovement between allowances and the business cycle. BIS WP-39 refers to this second interpretation as cyclicality to differentiate it from the causal interpretation. Research studies before CECL's implementation most often analyze *cyclicality* rather than *causal feedback*. In general, it is very difficult to identify the causal feedback because of complex interactions among banking regulations, economic policy, and economic activity, as well as data limitations.

From the existing research, it seems clear that CECL is subject to cyclicality quasi-by-design, as expectations about the severity of credit loss are likely to move in tandem with a deterioration of economic conditions. However, the degree of cyclicality will be conditional on the level of forecasting accuracy in anticipation of a downturn, as Loudis and Ranish (2019) show. Specifically, they consider three different scenarios that attempt to represent different levels of economic foresight that be reflected in banks estimates. Under perfect foresight of economic conditions, financial institutions will be able to adjust their CECL allowances in anticipation of a downturn. A myopic forecast, by contrast, will necessitate a significant increase in allowances over the unanticipated downturn. An alternative low-foresight scenario will result in a level of allowance adjustment somewhat in between the perfect foresight and the myopic case.

The CECL framework recognizes expected future losses beyond incurred losses, in contrast with the incurred loss framework. Because of this, there was broad agreement among studies on the assertion that if CECL had been adopted before 2007, the banking industry would have accumulated higher reserves in the early days of the Great Recession. There was also broad agreement that peak levels of allowances during downturns would have been higher under CECL, as the allowance in that case is determined over the life of the loan. These views are also broadly consistent with the experience of CECL adopters in the U.S. during the recent COVID-19-induced economic downturn, as we discuss in the next section. There also seems to be agreement that CECL adds flexibility to the ALLL when compared with the incurred loss framework and that it may add transparency to financial statements through enhanced disclosures. Studies before CECL's implementation generally also agreed on a relatively modest average "day one" impact of CECL, unless the economy was in the early stages of a recession.[12]

There was no broad agreement on the impact of CECL adoption on lending. Some authors argued that lending would be impacted if financial institutions are required to significantly increase their allowances during downturns (Covas and Nelson, 2018). Others, meanwhile, argued that with enough hindsight, the added flexibility of CECL would allow lenders to build additional allowances before the downturn or early in the downturn, and this could limit the impact on lending (DeRitis and Zandi, 2018). Loudis and Ranish (2021) find no significant evidence of a direct impact of CECL on lending during the COVID-19 crisis, although this particular downturn was unusual by its level of government support to consumers and businesses, and because banking regulations were temporarily adjusted to accommodate concerns about CECL's implementation (Wall, 2020). Research that directly addresses the challenges of the novel CECL regulation is still in its early stages. Our research fills a gap on the existing literature by analyzing CECL sensitivity to model and forecasting error.

III. CECL Implementation in the Time of COVID-19

[12] Covas and Nelson (2018), DeRitis and Zandi (2018), Loudis and Ranish (2019), and Chae, Sarama, Vojtech, and Wang (2018).

The onset of the pandemic created significant unanticipated challenges for CECL adopters. The pandemic's impact on the economy and credit markets was significant, and the unprecedented policy responses in the form of lockdowns and monetary and fiscal policy were significant as well.[13] Unemployment rose in the U.S. in April 2020 to 14.7 percent from a reported 4.4 percent in the prior month, and then it decreased rapidly over the remainder of the year at about 1 percentage point per month in the first few months, reaching a rate below 4 percent by the end of 2021. Credit supply was initially impacted but quickly recovered, in good part because of the implementation of government and central bank stimulus programs in the United States. Credit demand shifted significantly over time in synchrony with changes in consumer behavior, lockdowns, and pandemic waves. Other macroeconomic indicators were also significantly impacted. Economic and financial forecasts deteriorated significantly during this period.

A. CECL Allowances During COVID-19

While FASB issued the CECL standard in June 2016, it didn't become effective until the fiscal year beginning after December 15, 2019, for most SEC filers, including complex bank holding companies (BHCs). Other BHCs not included in the first round of adoptions were required to implement the CECL standard starting with the fiscal year beginning after December 15, 2022. Information on loan allowances under CECL were first reported by adopters in the March 2020 quarterly FR Y-9C consolidated financial statements for BHCs. Using publicly available consolidated financial statements for holding companies reporting form FR Y-9C, we construct a panel data set for the years 2017 to 2022 and use these data to analyze the performance of the allowance for CECL adopters and nonadopters.[14]

Figures 1 and 2, and table 1, employ the panel of FR Y-9C disclosures to analyze the evolution of the allowances across BHCs before and after CECL implementation.[15] We restrict our sample to BHCs with more than $5 billion in reported consumer loans in their balance sheet as of the end

[13] Pinello and Puschaver (2020) provide a financial account of the challenges faced by CECL adopters in the first quarter of 2020. Wall (2020) provides additional information about regulatory efforts to minimize the impact of CECL in the early days of the pandemic.
[14] www.chicagofed.org/banking/financial-institution-reports/bhc-data.
[15] Loudis, Pechenik, Ranish, Vojtech, and Xu (2021) conduct a similar analysis also using FR Y-9C, while Rosenblum and Lai (2020) employ other sources of financial disclosure. The focus of these studies is not the analysis of model risk.

of 2019, in order to have a more homogeneous group. The data include institutions that reported CECL allowances in the FR Y-9C for the first time in the first quarter of 2020 and institutions that didn't implement CECL before the end of 2022. There are also a small number of institutions that reported CECL allowances for the first time at some point after the first quarter of 2020 that are not included in our analysis.

We analyze the 2017–22 period, with particular attention to the 2019–20 period, which encompasses the first mandated CECL transition and the COVID-19 pandemic. The COVID-19 pandemic represented a unique economic shock that was difficult to forecast, and this contributed to an increase in allowances that is consistent with the myopic forecasting case discussed in Loudis and Ranish (2019). Figure 1.a depicts the behavior of allowances for CECL 2020 adopters and nonadopters, with allowances reported as a percentage of the allowances in the fourth quarter of 2019, which is selected as the reference point. The dotted line represents the first day CECL transition amount, resulting in an increase in allowances of about 30 percent on average. The dash line denotes CECL allowances in the last quarter of 2021, which is representative of a period with low unemployment and low charge-off rates. Allowances remained broadly unchanged during the 2017–19 period of stable economic conditions. CECL allowances increased significantly early in the pandemic, as a result of a combination of worsening economic forecasts, the added flexibility of the CECL framework, and more expansive provisioning requirements. CECL allowances reached their peak in 2020:Q2, while allowances for nonadopters reached their peak in 2020:Q3. Allowances decreased significantly during 2021, as economic conditions improved, and remained relatively stable in 2022.[16]

Figures 1.b to 1.d depict the behavior of allowances for CECL adopters and nonadopters across consumer loan portfolios: residential, credit cards, and autos. First, we observe that while both adopters and nonadopters responded to the pandemic by increasing allowances, CECL adopters increased allowances by a much larger margin across portfolios, and at a faster pace. The main differences are observed in residential loans, consistent with the intuition that the CECL

[16] Beck and Beck (2022) report the same performance of provisions across CECL adopters and nonadopters and suggest that this represents preliminary evidence that ASU 2016–13 has achieved its objective of making allowances more sensitives to changing economic conditions. Chen, Dou, Ryan, and Zou (2022) argue that the observation that CECL adopters in 2020 increased provisions more than nonadopters is consistent with the CECL approach increasing cyclicality.

methodology has a larger impact for portfolios of long-duration loans, and also consistent with models trained with data encompassing the Great Recession, which resulted in particularly severe losses for mortgage portfolios.

Table 1 provides disaggregated information about the quarterly evolution of allowances for CECL adopters and nonadopters for the period 2020–22 with respect to the reference allowance in the last quarter of 2019. The table highlights that CECL allowances picked up between the second and third quarters of 2020 and subsequently experienced a steep decline until reaching relative stability toward the end of 2021. The values in the last quarter of 2021 suggest that under the mild economic conditions at the time, allowances for residential loan portfolios would be about 50 percent higher under CECL, while allowances for credit cards and auto portfolios would be around 70 percent and 90 percent higher under CECL, respectively.

Figure 2 depicts charge-off rates over the period 2017–22 across different consumer loan portfolios. Charge-off rates during the COVID-19 pandemic decreased with respect to the already record-low levels of recent years, both at the aggregated level and across consumer portfolios. Figure 3 depicts allowances and charge-offs for all commercial banks over the period 2000–22 at the aggregate level. As Figure 3 indicates, the relationship between ALLL and charge-offs was particularly strong during the Great Recession, with charge-off rates increasing significantly and more rapidly initially than the ALLL, which did not peak until early 2010. In contrast, charge-offs decreased to record-low levels during the 2020–22 period, while allowances increased significantly during the initial phase of the pandemic and then decreased significantly over the next few quarters, until they stabilized in 2022. This difference in charge-off and allowances performance across two stress periods is difficult to reconcile without considering the unprecedented fiscal and monetary policy responses experienced during the pandemic, and those responses' impact on the performance of consumer finance portfolio.

Comparing charge-offs across portfolios with allowances in 2020, it is apparent in hindsight that both the incurred and the CECL framework provisioned for significant losses that didn't materialize. The projections of losses were clearly impaired by the effects of a one-in-100-year pandemic and the associated government response. Next, we summarize public information on the response of financial institutions to the challenges to the provisioning framework emerging from the pandemic.

B. Firms' Responses to Errors in Economic Forecasts and Models

A recent BIS (2022) newsletter offers a window into the strategies leveraged by financial institutions to mitigate model risk and adapt their credit risk modeling policies and practices to the challenges of the pandemic.[17] As discussed, credit risk performance during the pandemic deviated considerably from historical patterns and trends. In response, banks applied sizeable judgment-based adjustments (overlays and overrides) to their provisioning models. This created challenges of monitoring controls and governance around model adjustments. Supervisors observed three main challenges in relation to banks' provisioning models: first, challenges around controls regarding model risk management and data; second, challenges capturing economic uncertainty; and third, challenges identifying credit deterioration in vulnerable sectors and borrowers.

Observed adopted approaches to model development challenges included: (1) exclusion of COVID-19-related data, primarily because of the observed disconnect between macroeconomic variables and default rates; (2) utilization of new data collected during the COVID-19 pandemic with the application of judgmental overlays to counteract any changes to existing relationships (e.g., macroeconomic variables versus defaults); (3) enhanced infrastructure and data feeds to ensure the relevant data are fully understood and properly integrated into analysis of decision-making systems. Thus, banks and supervisors are grappling with how to incorporate and reflect data over the COVID-19 period into the allowance framework going forward.

The challenges of incorporating COVID-19 information into the analysis and the reliance on overlays points to weaknesses in the allowance framework in times of crisis, when confidence in the framework matters most. This underscores the importance of drawing lessons from crisis episodes in order to improve the robustness of the framework in future crisis.

IV. Forecasting Pitfalls

[17] See the BIS (2022) newsletter on COVID-19-related credit risk issues (bis.org).

The reliance of CECL on reasonable and supportable forecasts increases the sensitivity of the allowance to economic forecasting errors, which can be particularly large during periods of economic stress. Another potential source of error less frequently discussed and possibly more detrimental is the problem of model misspecification error. Intuitively, model misspecification occurs when a model is a poor representation of the process that it intends to mimic. Model misspecification is a biproduct of the unique challenges that a new crisis usually brings. It differs from error in an economic forecast in that it applies to the core models of the allowance framework and will result in biased predictions, even in cases when economic forecasts are accurate.[18]

Model accuracy is desirable in principle, one cannot always aim for model projections that are conservatively inaccurate in periods of stress when the underlying framework has significant flaws. In addition, model accuracy impacts a second objective of CECL, which is balance sheet transparency.[19] It is important to recognize the roots of the bias in order to address model shortcomings and to implement model infrastructures that are resilient to shocks and less reliant on overlays. In this section, we formally analyze the challenges of economic forecast and model misspecification error and consider potential remedies.[20]

A. Economic Forecast Error

CECL allowances constitute forward-looking estimates of credit losses, with reasonable and supportable forecasts representing a critical input in its calculation. It should be of no surprise that the impact of economic forecasting error may have been substantial at times during the COVID-19 pandemic.[21] To illustrate the potential magnitude of forecasting error, we review the historical evidence on one-year-ahead forecast accuracy from the Philadelphia Fed's *Survey of Professional Forecasters* (*SPF*). For simplicity, we focus on the forecast of the unemployment rate,

---

[18] A popular quote among statisticians is that "all models are wrong, but some are useful."

[19] Pinello and Puschaver (2022) provide a financial account of the challenges of implementing CECL during the pandemic, including an overreliance on management's judgment in view of the challenges interpreting results from CECL models.

[20] In a recent speech, Governor Christopher J. Waller (2021) stresses the limitations of economic forecasting by highlighting that "forecasters need to approach this work with humility." He also emphasizes that "economic forecasting is a pretty hopeless endeavor. So why do we do it? Because of how much is riding on the outcome."

[21] For example, Canals-Cerdá (2020), looking at credit card portfolios, observed that the impact of forecasting error could have been substantial during the initial quarters of the Great Recession, with deviations from the baseline between 30 percent and 40 percent in most segments.

which is an important macroeconomic driver of CECL projections across consumer finance portfolios. Figure 4 displays historical realized unemployment rate for the period 1970–2022; the figure also displays the level of the one-year-ahead average forecasting error from the *SPF*. Before the COVID-19 pandemic, the largest one-year-ahead forecasting error was 4 percent in absolute value, which was achieved during the Great Recession. In contrast, during the initial days of the pandemic, partly as a result of lockdown mandates, the unemployment rate increased suddenly to above 14 percent, and the one-year-ahead forecasting error increased to a record 9 percent in absolute value.

CECL offers the flexibility to increase allowances in anticipation of downturn economic conditions, but this requires some level of forecasting accuracy. Based on the experience from the two most recent crisis episodes, we can expect economic forecast uncertainty to increase significantly during periods of stress and CECL projections to be significantly impacted. The effect of economic forecasting errors on allowances is unlikely to be homogeneous. It will vary across portfolios and across risk segments of a portfolio. It will also vary across model specifications.

B. Model Misspecification Error

Financial institutions applied sizeable judgment-based adjustments to their provisioning models during COVID-19 in an attempt to mitigate the effects of model misspecification error paired with a highly unusual — and out of historical range — macroeconomic environment. The differences between allowances and charge-offs in Figures 1, 2, and 3 during the pandemic affects both the CECL and the incurred loss frameworks. However, as Figure 1 shows, the effect is particularly dramatic for CECL adopters. This suggests that the impact of model misspecification error on future expected losses, in addition to incurred losses, was significant under CECL.

In the next paragraphs, we analyze conceptually the potential effects of model misspecification error on CECL projections. We draw on the literature on model specification and forecasting to better understand the challenges that can impact CECL. We begin with a sample statistical representation of the problem of generating forward-looking estimates k periods into the future of a certain quantity of interest y,

$$y_k = \psi_k(s, m_k) + \in_k$$

where y_k represents the value of y k periods into the future, which is a function of portfolio characteristics s, reasonable and supportable forecasts of economic conditions up to k periods into the future, denoted m_k, and a residual stochastic unpredictable component $\in_k$ which accounts for additional unexplained variability in outcomes. In practice, a forecast $\hat{y}_k$ requires forecasts $\widehat{m}_k$ of macroeconomic conditions and unbiased estimates $\widehat{\psi}_k$ of fundamental relationships. In some cases, it may also require estimates of certain aspects of the distribution of $\in_k$. The projection can then be computed as, $\hat{y}_k = \widehat{\psi}_k(s, \widehat{m}_k)$.

In order to better understand modeling challenges, we adopt the terminology of Hendry and Mizon (2014).[22] These authors classify the problem of unpredictability in econometric modeling and forecasting into three distinctive categories with different implications. Informally, these three categories can be described as: (1) anticipated stochastic variation in forecasts, (2) unexpected instances of outliers, more commonly known as “black swans” and (3) unanticipated shifts in the relevant relationships postulated by the model, also known as “regime shifts” in certain contexts. More formally, the authors define these categories as: (1) intrinsic unpredictability, (2) instance unpredictability, and (3) extrinsic unpredictability, respectively. This categorization offers a useful tool to better understand modeling challenges across different economic environments and for formulating strategies to minimize their impact.

Intrinsic unpredictability is the result of innate uncertainty in forecasts; thus, it is inherently unavoidable. The second and third categories are conceptually different but may be difficult to distinguish in practice. The case of *instance unpredictability* can be described by a probabilistic process subject to a nonnegligible probability of a nonpersistent unexpected “black swan” event. This case can be explained within the framework of the postulated probabilistic process, perhaps as a result of fat tails in the distribution of the model residual. By contrast, the case of *extrinsic unpredictability* refers to a persistent distributional shift that cannot be reasonably explained within the framework of the postulated probabilistic process. After a

[22] Zhang, Singh, Ghassemi and Joshi (2023) analyze the problem of model performance from the perspective of the machine learning literature. Breeden (2018) presents an early study of the impact of model specification assumptions on the cyclicality of CECL projections before CECL implementation and before the pandemic.

distributional shift, outliers may become a common occurrence. A persistent change in economic relationships for an extended period of time would fall into the category of *extrinsic unpredictability*. Intuitively, this may be the primary differentiating feature between categories two and three.

The two most recent crises, the Great Recession and the pandemic, are arguably examples of *extrinsic unpredictability*. In the case of the Great Recession, mortgage defaults increased considerably while home prices experienced unprecedented drops.[23] Lenders' recoveries from defaulted mortgages also decreased markedly as a combination of lower home prices and increased time to foreclosure and sale. This level of stress in the mortgage market persisted for several years and was significantly different from prior experience.[24] In the case of the pandemic, life as we knew it changed suddenly and dramatically, as did important economic variables, like unemployment. The impact of the pandemic and the resulting government policies had a long-lasting impact on borrower behaviors.

The pandemic triggered unprecedented levels of government intervention, which included direct assistance to households, extensions of unemployment benefits, as well as programs directly targeted at consumer lending, primarily mortgage and student loan forbearance programs. Credit cards and auto loans were also impacted by forbearance efforts, although to significantly lesser degrees. It is not surprising that significant government interventions, unaccounted for in models during the pandemic, could lead to significant bias in model projections. The unprecedented level of government assistance impacted the future credit performance of banks' loan portfolios and contributed to a breakdown in the traditional relationships between economic variables and measures of credit risk, and portfolio loss, for a prolonged period. Incorporating the effect of government assistance in our theoretical equation results in the following expression,

$$y_k = \Phi_k(s, m_k, g_k) + \epsilon_k,$$

[23] fred.stlouisfed.org/series/csushpinsa
[24] fred.stlouisfed.org/series/DRSFRMACBS

where g_k represents diverse government assistance programs that were introduced at different points during the pandemic and were omitted from pre-pandemic models, as they were absent from the historical data.

While an equation as the one postulated above cannot generally be directly estimated, given the lack of historical data along with other identification challenges, it can still inform us about sources of model misspecification, along with potential strategies to mitigate bias in projections. Model misspecification error can lead to biased projections, even in the case of accurate economic forecasts. In our case, the relationship Φ_k may differ substantially from the estimated relationship ψ_k before the pandemic. Thus, the typical sources of model misspecification, functional form misspecification and omitted variables, are represented in the above equation. Predictions $\hat{y}_k$ relying on precrisis estimates of $\hat{\psi}_k$ will likely lead to systematic forecast bias, consistent with the case of extrinsic unpredictability, unless model misspecification bias is acknowledged and properly addressed. Models trained with historical data over the period of the Great Recession were poorly equipped to forecast the impact of the pandemic as well as the effects of fiscal and monetary policy actions. The level of government support significantly minimized the severity of economic outcomes.[25] Therefore, it is perhaps not surprising to observe a disconnect between allowances and charge-offs, as depicted in Figure 3.[26]

C. Mitigating the Impact of Forecasting Error and Model Misspecification

Error in macroeconomic forecasts and a more fundamental problem of model misspecification are potential sources of CECL bias, as discussed above. Macroeconomic forecasts are inherently uncertain, and the level of uncertainty generally increases in challenging economic environments, like the early stages of a financial crisis or a pandemic. Thus, lessons learned from prior crises suggest that reasonable and supportable forecast horizons are likely to be shorter in periods of high uncertainty. It may also be helpful to translate uncertainty in forecasts into CECL projections, for example, by considering multiple scenarios with the importance assigned to

[25] International accounting standard setters have emphasized that banks should consider the impact of government policies in their analysis of allowances (De Araujo, Cohen, and Pogliani, 2021). The results in Degryse and Huylebroek (2022) are consistent with a positive impact of government fiscal policy on banks' credit risk and profitability.

[26] The experience of the Great Recession also generated significant debate about model performance during crisis periods (see, for example, Gerardi, Lehnert, Sherlund, and Willen (2008) and Frame, Gerardi, Lehnert, and Willen (2015)).

different scenarios commensurate with the level of confidence. During periods of elevated economic uncertainty, it may also be helpful to look for novel sources of information and external benchmarks, as well as to consider more frequent development and validation of forecasts.

How can we mitigate CECL sensitivity to model error under extrinsic unpredictability conditions? Extrinsic unpredictability conditions can lead to long-lasting changes in model-postulated relationships. Thus, in these instances, it may be necessary to adapt and modify models to the realities of a novel environment in order to be able to overcome ingrained misspecification bias. The models and strategies to be considered in periods of uncharted economic conditions can be informed by insights from econometric theory, by an analysis of primary and auxiliary data after the shock, as well as by expert judgment.[27]

Hendry and Mizon (2014) point out that it may be possible to address the effects of extrinsic unpredictability ex post. Novel evidence available after a shock can inform model selection and re-estimation, and sources of misspecification and forecast failure can be potentially addressed.[28] Econometric theory suggests that model factors that have the largest correlations with relevant unaccounted factors, or omitted variables, will have the largest impact on misspecification bias. Thus, simple economic reasoning and expert judgment can help us address model shortcomings and identify model specifications that are more suitable to the novel environment. Simple model specifications that use robust sources of information and downplay potentially biased information may prove useful after a shock. It may also be helpful to analyze potential divergences between early indicators of stress and model predictions of loss, this can enhance the information set after the shock and serve as an early warning of model performance bias. It may also be possible to leverage the information of early indicators to ascertain the performance of standard measures of portfolio risk and to discriminate across model candidates. Overreliance on a single model is probably not an optimal strategy in times of stress. In fact, while models conditional on macroeconomic factors generally performed poorly, not all relationships

[27] Model misspecification during a crisis is only one possible source of forecasting bias. For example, measurement error in input variables broadly defined could be considered as another candidate for further analysis.

[28] A recent speech by Fed Governor Waller offers advice for tackling challenges, beyond forecasting errors, that often arise during periods of economic stress arising from unprecedented circumstances. Waller advises that "when the shock is unique, adapt fast." This requires careful analysis of the novel shocks and may also require modifying and adapting models to the novel environment.

"broke down" during COVID-19, as we argue in the empirical section of the paper. Thus, it may be useful to regularly evaluate the strength and weaknesses of different model specifications.

V. An Application for Consumer Finance Portfolios

In the previous section, we highlighted the advantages of a flexible and adaptable modeling framework that can be quickly adapted to the challenges of a novel crisis. In this section, we advance our views by presenting an econometric framework that is nimble and adaptable and consistent with the typical modeling framework implemented by the most sophisticated CECL adopters. We also leverage this framework to analyze the usefulness of some of the strategies previously discussed to palliate the impact of model misspecification error.

The modeling framework considered can be estimated and deployed rapidly, irrespective of the size of the portfolio considered. For this reason, the approach is particularly valuable in consumer finance, in which the typical loan portfolio comprises many millions of loans, like personal loans, mortgages, auto loans, or credit card loans. Mortgages and credit cards have received significant attention in the literature, especially regarding their performance during the Great Recession. We consider an application for auto loans, which have not previously received the same level of attention.

A. The Data

We employ data from the FRBNY Consumer Credit Panel/Equifax (CCP), and specifically its associated Auto Tradeline panel data. The CCP is a panel data set comprising information from anonymized individual credit bureau reports starting with the first quarter of 1999. The panel comprises a nationally representative 5 percent random sample of individuals with a credit history.[29] The Auto Tradeline panel associated with the CCP was constructed to provide additional loan-specific information on associated auto loans. The CCP auto tradeline includes snapshots of the auto tradelines in the credit bureau data in six-month intervals. It includes loan-specific origination information such as origination date and loan balance, and monthly performance information that is updated periodically. Tradeline information can be

[29] Lee & Van der Klaauw (2010) describes the data in more detail.

complemented with additional borrower-specific credit bureau information available in the main CCP panel, like borrower Risk Score.

While the tradeline data provide valuable information about the performance of auto loans, they also have some limitations for the analysis of allowances. Specifically, they do not include information on recovery values in the case of default — information that is readily available to lenders. For this reason, our empirical framework will focus on the analysis of default rather than the analysis of loss. We also restrict our sample to loans issued by banks and credit unions in order to focus our analysis on depository financial institutions. We complement the Auto Tradeline data with additional information on key macroeconomic variables, primarily state unemployment.

Banks and credit unions have generally a higher concentration of safer loans when compared with the overall market, with nonbank lenders having a larger concentration of subprime borrowers. Figure 5 depicts changes over time in early delinquency for our representative portfolio. The figure highlights the significant increase in default risk over the period 2008–11 around the time of the Great Recession. In contrast, delinquency generally decreased during the pandemic, particularly severe delinquency.

Intuitively, a model's forecasting ability is in good part determined by the information embedded in the historical training data. With this in mind, in figure 6, we parse out the variation in unemployment rates across states, which is the primary source of macroeconomic variation informing our models. The figure provides information that will help us understand the performance of models with different sets of training data. Most of the variation in the unemployment rate from 2001 to 2007 is concentrated in unemployment rates between 3 percent and 7 percent. This contrasts with the 2009–11 period, during which unemployment increased significantly across the board, with unemployment rates concentrated between 6 percent and 12 percent. The experience in 2020 was even more remarkable. Suffice it to say that the year started with an aggregated unemployment rate of 3.5 percent that jumped to 14.7 percent in April of that year, at the onset of the pandemic. Unemployment across states in the first half of 2020 was concentrated within the range of 2.2 percent to 28.5 percent, with the largest value achieved in April in Nevada, a state that was severely impacted by lockdown mandates.

B. The Empirical Framework

Consider a loan portfolio that can be divided into S segments of loans with broadly homogeneous risk characteristics. Each segment is composed of loans with the same, a priori, independent probability of default p. It follows then that the aggregated default distribution for a segment of N loans will follow a binomial distribution $B(N, p)$. Furthermore, for an N large enough, the Poisson distribution *Poisson(λ)*, with λ=Np, represents an excellent approximation to the $B(N, p)$ distribution. Thus, our empirical strategy considers the estimation of segment-level Poisson models for the number of defaults n_{sl} in each period ahead of the postulated life of the loan $l = 1, \dots, L$, for each segment S of N_s loans for each vintage in our estimation data set. Specifically, we postulate that the number of defaults n_{sl} associated with segment S in period l can be represented by the Poisson distribution,

$$n_{sl} \sim P(\lambda_{sl}, N_l) \text{ for } s = 1, \dots, S \text{ and } l = 1, \dots, L$$

In our empirical specification, we consider a `standard parametrization $\lambda_{sl} = \lambda_l(X_s, m)$, with X_s representing segment specific characteristics and m representing region-period specific macroeconomic drivers. We also consider a more flexible, segment-specific parametrization, which is ultimately our specification of choice.

The impact of economic conditions on the risk profile of a portfolio of consumer loans is typically identified by the historical variation in economic variables over time and across geographic regions, most often across states. With loan level data representing T snapshots, or cohorts, of a loan portfolio and credit performance up to L periods ahead, we can leverage the heterogeneity in macroeconomic conditions and performance across regions and over time. Loan-level data can be aggregated at the segment-geography level as,

$$\{(N_{sgt}, n_{sgtl}, m_{sgtl}): s = 1, \dots S;\ g = 1, \dots, G;\ t = 1, \dots, T; l = 1, \dots, L\}$$

with N_{sgt} representing the number of loans in a specific segment-geography for a particular snapshot t, n_{sgtl} representing the number of associated defaults in performance period l, and m_{sgtl} representing macroeconomic conditions in geographic unit g at period l.

With modern statistical software, after data manipulation, the approach can be implemented with a single line of code, for example, using the General Structural Equation Modeling (GSEM) available in Stata.[30] In our preferred model specification, we consider a segment-specific parametrization of λ_{sl} with unemployment and the six-month unemployment change as macroeconomic risk drivers. The approach can be applied to the estimation of unconditional (our preferred method) or conditional probabilities. The estimation of unconditional probabilities is often more robust because the conditional probability framework can be impacted by potential propagation of model forecasting error. Using the estimated Poisson framework, along with macroeconomic projections $\widehat{m}_{sgtl}$, we can derive forecasts of segment defaults $\hat{n}_{sgtl}(\widehat{m}_{sgtl})$ or segment default rates $\hat{n}_{sgtl}(\widehat{m}_{sgtl})/N_{sgt}$.

Importantly, note that the size of the resulting data set, after the segmentation scheme has been determined, is a function of S, the number of segments, rather than a function of the portfolio loan sample size. Thus, the sample size of the original portfolio becomes muted. This is particularly important for consumer loan portfolios of mortgages, autos, and cards, with potentially tens of millions of loans, or even hundreds of millions of loans in the case of credit cards. As a result, we can conduct the empirical analysis on a portfolio of any size without increasing the computational burden. Specifically, in our empirical example, we employ the whole sample of auto loans in the consumer credit panel from 2001 to 2022, consisting of 5 percent of all loans originated in the United States and reported to the credit bureau. Our models can be estimated and deployed in minutes.

C. Selection of Segmentation Scheme

[30] See Canals-Cerda (2022) for a description of the GSEM framework and an illustration of this powerful framework.

One potential problem with the approach described in the previous subsection is that it becomes impractical when the number of segments is large enough. The approach can incorporate continuous variables as long as they are constant within segments, but not necessarily constant over time. This is typically the case for macro variables, which are constant at a certain aggregated geographic level, state in our case. However, other continuous variables will have to be incorporated into a segmentation scheme in order to be included in the empirical framework. In order to select an optimal segmentation scheme, we employ a ML classification algorithm on a 20 percent random sample of the data across all vintages during 2001–20, with the target variable defined as the two-year forward default, which takes the value one if the loan defaults within two years and zero otherwise, and with features including credit score and loan size at origination.[31] Other options could be an expert judgement segmentation based on business needs, or a segmentation inspired by regulatory requirements. For example, the Federal Reserve FR Y-14Q Auto submission requires banks to report portfolio information at the segment level by product type, age, original LTV, credit score, delinquency, and geography, resulting in a segmentation scheme with a few thousand segments.

Figure 7 reports receiver operating characteristic (ROC) metrics from our ML segmentation approach applied across vintages. Figure 7.a graphically depicts the ROC performance as the maximum depth of the tree increases from one to 23. While the ROC performance continues to increase with the maximum depth with the training data, using the test data instead, we observe that the ROC does not significantly increase after a maximum debt of three. Therefore, for our empirical application, we select an optimal segmentation scheme based on a maximum depth of three, resulting in a segmentation scheme with eight segments.

In order to analyze the stability of the segmentation scheme over time, we consider the ROC performance of the segmentation scheme across year cohorts, from 2001 to 2020. This information is reported in Figure 7.b. As the figure indicates, the ROC of the segmentation scheme remains stable over time, taking values that range from 0.81 to 0.84. Perhaps not

[31] We experimented with several ML classification techniques and eventually selected a decision tree classifier based on the entropy criterion (ref. sklearn.tree.DecisionTreeClassifier — scikit-learn 1.2.1 documentation).

surprisingly, the ROC metric recorded its lowest values in 2006–09 and 2020, i.e., around times of significant economic uncertainty.

VI. Discussion of Empirical Findings

Our focus in this empirical exercise will be on the problem of model misspecification error, which has received less attention from practitioners than the problem of macroeconomic forecasting error. Model misspecification errors have usually been addressed by practitioners with model overlays and overrides relying primarily on expert judgment and other auxiliary information, without directly tackling the roots of the problem. Here, we leverage our simple empirical framework to examine the effects of model misspecification error in times of high economic uncertainty and analyze strategies to mitigate its impact. We take advantage of our rich historical data, which encompass two periods of significant economic uncertainty. We analyze model shortcomings specific to each crisis period separately. We observe that each crisis episode offers its own lessons that may offer some useful guidance for future crises.

A. Model Performance During the Great Recession

In order to focus on the impact of model misspecification bias, we assume perfect macroeconomic foresight, and a nine-quarter reasonable and supportable forecast period. Figure 8a presents realized and forecast nine-quarter default rates for cohorts of newly originated auto loans from the 2001 to 2014 cohorts. The figure illustrates the impact of different training data sets on the out-of-sample performance of model projections. The solid line depicts the realized nine-quarter forward default rate, while all other lines represent model projections using our preferred model specification estimated with different training data sets, including the 2001–05, 2001–07, 2001–08, and 2001–09 cohorts. The model estimated with data from the 2001–05 cohorts performs well in times of benign economic conditions, before and after the period of the Great Recession, but it performs poorly during the period of the Great Recession, characterized by significantly higher defaults. In order to understand this performance, note that the 2001–05 cohorts experienced mostly benign economic conditions, characterized by relatively low levels of unemployment during the first nine quarters after origination, as illustrated in figure 6. A model estimated using the 2001–07 cohorts performs much better during the period of the Great Recession. Considering data from the 2001–08 cohorts further improves model fit during the period of the Great Recession, adding additional cohorts does not improve performance significantly.

It may also be helpful to analyze model performance across segments. Looking at figure 9, we observe that the performance of models across risk segments follows a similar pattern as the performance at the aggregate level. However, for the riskiest segments, the default rate seems to deteriorate more rapidly in the early stages of the Great Recession. This suggests that there is value in tracking the performance of specific segments, as some segments can be early indicators of risk for the rest of the portfolio.

Finally, while the focus until now has been on model performance over a nine-quarter period, we may be able to draw additional insights from looking at the lifetime allowance performance prescribed in CECL. For the purpose of our analysis, we assume a life of a loan of five years, with a nine-quarter period of reasonable and supportable forecast. Figure 10 reports realized lifetime defaults across cohorts and modeled CECL lifetime estimates of default. The estimated lifetime default rates combine a nine-quarter estimate of default under perfect foresight of economic conditions, with a remaining-life-of-the-loan estimate of default beyond nine quarters that represents an over-the-cycle estimate over a mix of economic conditions. We observe that the model-projected lifetime default rate generally lies above the realized default rate in periods of good economic conditions, while it lies below the realized default rate during the period of the great recession. What explains the performance of CECL projections? On the one hand, defaults during the Great Recession remained elevated beyond the assumed nine quarters of perfect macroeconomic foresight; this explains the CECL underprediction during the Great Recession. On the other hand, the long-run average default rate estimated with the 2001–07 cohorts includes the period of the Great Recession, resulting in estimates that are overly conservative during periods of benign economic conditions. Thus, lifetime CECL projections average out good and bad economic environments beyond the reasonable and supportable timeframe, and this explains the observed differences between realized and projected lifetime default rates.

The experience from the Great Recession indicates that fitted models usually underperform when presented with uncharted economic environments characterized by out-of-sample macroeconomic conditions. However, when models are re-estimated with additional data that include some exposure to the new macroeconomic environment, performance can improve significantly. It is also useful to analyze model performance across segments, as model underperformance in certain segments can act as an early warning for more widespread underperformance. Furthermore, CECL long-run projections by design average out economic cycles to a certain extent, although short-term economic conditions are a key determinant of CECL allowances. Thus, the experience during the Great Recession offers useful insights that may assist modelers in building model infrastructures that are resilient and adaptable in future crises.

B. Model Performance During the Pandemic

The COVID-19 pandemic generated significant stress among retail borrowers. It also triggered unprecedented levels of government assistance, including forbearance programs directly targeted at consumer lending. We have argued in this paper that this unprecedented level of government assistance contributed to a breakdown in the traditional relationship between economic variables and consumer credit risk, which prompted the divergence between historical charge-offs and allowances reported in figure 3.

Figure 11 looks at the evidence of model performance during the pandemic in our empirical application for auto loans. The figure compares nine-quarter realized default rates across cohorts, with projected default rates across different model specifications under perfect economic foresight. The solid line represents the realized nine-quarter default rate across cohorts; the dotted line represents projected nine-quarter default rate for a model estimated using our preferred model specification and data from the 2001–17 cohorts. The model provides a reasonable fit of the data up to the 2018 cohort. In contrast, model projections deviate significantly from realized outcomes for cohorts with a nine-quarter projection period overlapping with the period of the COVID-19 pandemic. Thus, our preferred model specification, based on data before the pandemic, forecasts dramatic increases in defaults due to the macroeconomic experience in the early days of the pandemic. As we well know, the dramatic increase in defaults projected by the model never materialized.

Figure 11 also depicts projected defaults from a model estimated using our preferred specification and data from the 2001–20 cohorts (long-dashed line), which includes the period of COVID-19 pandemic. We also report projections from a model estimated with data from the 2001–17 cohorts but for a model specification that does not include macroeconomic drivers (dashed line). Thus, changes in projected default rates for this last model are driven only by cohort-specific risk characteristics. Observe that the default rate projections from these two models are almost the same. This is consistent with our intuition that model forecast error during the pandemic resulted in good part from the misspecification of the impact of macroeconomic variables during that period. This misspecification is the result of generous government policies directed to mitigate the effects of lockdown policies. Informing the models with data from the pandemic significantly improves the fit of model projections. Excluding macroeconomic factors in models estimated with data from before the pandemic improves accuracy as well, in this case because these excluded factors are the most likely source of model misspecification, as we postulated in a prior section.

Figure 12 expands on figure 11 by depicting model performance across cohorts and risk segments. Interestingly, we observe a significant divergence in model performance between high-risk and low-risk segments. For the highest-risk segments, we observe that realized default rates decreased most significantly with respect to the pre-pandemic trend. This suggests that government policies had the largest impact on these segments of consumers. In contrast, we observe the largest impact of model misspecification in the lowest-risk segments, with the model estimated using the economic experience before the pandemic (dotted line) experiencing the largest divergence from observed outcomes in these lower-risk segments, proportionally. Thus, also in this case, there is significant value in tracking the performance of specific risk segments. Specifically, observed performance across the highest-risk segments may serve as an early indicator of the impact of government policies, while model projections across the lowest-risk segments may serve as an early indicator of the impact of model misspecification error.

The experience from the pandemic offers insights consistent with the experience during the Great Recession, as well as novel insights. Consistent with the experience during the Great Recession, we observe that out-of-sample macroeconomic conditions can lead to model underperformance. Also, re-estimated models that incorporate some data from the current crisis generate much improved forecasts. One important facet that was unique to the pandemic was the unprecedented level of government assistance targeted specifically at retail borrowers. This level of assistance precipitated a breakdown in the traditional relationship between economic variables and consumer credit risk. Leveraging econometric theory insights, we explore the performance of models without macroeconomic drivers, which are a significant source of model misspecification, and verify that the projections from these models are much more in line with the observed performance of auto loan portfolios during the pandemic. Thus, it can be fruitful to consider a variety of model specifications, particularly in times of crisis, and to leverage econometric theory insights of the potential impacts across model specifications of specific aspects of a crisis.

While we argue in favor of a flexible model infrastructure, we should also acknowledge the challenges of this strategy, especially for heavily regulated financial institutions. A recent report (Kumar, Laurent, Rougeaux, and Tejada, 2022) indicates that validation of Tier 1 models in the U.S. requires 12 weeks on average, while Tier 2 and 3 models require six and four weeks, respectively. Undoubtedly, validation resources get strained during periods of crisis. Model simplicity can assist with the validation process. It is important to plan ahead and to consider every aspect of the model life cycle as part of the model development process. A model infrastructure that is nimble and adaptable and that can leverage the lessons of a variety of model specifications can ease validation constrains in periods of crisis.

VII. Conclusions

CECL represents a significant change in the way financial institutions compute their allowance for credit losses. The new framework focuses on lifetime expected losses rather than incurred losses, and it is expected to add transparency to financial statements. We analyze the performance of CECL across financial institutions around the time of its implementation, which coincided with the start of the COVID-19 pandemic. The experience supports claims from studies before implementation. Specifically, claims regarding CECL cyclicality during unanticipated crisis episodes, as well as higher peak allowances than the prior allowance framework. The recent crisis also increased our awareness on the sensitivity of CECL allowances to model and macroeconomic forecast errors.

The focus of our empirical application is on the implementation of CECL to consumer finance portfolios, perhaps the most challenging area of CECL implementation, given their size and complexity. Specifically, we examine auto loans portfolios, which have received less attention in the consumer finance literature. Our empirical implementation combines machine learning techniques with standard statistical principles. The approach considered is simple without compromising performance, it can easily accommodate multiple models, and it allows for quick and simple model redevelopment, redesign, and deployment, irrespective of the size of the loan portfolio. The simplicity of the framework can also streamline the model validation process.

We analyze problems of forecasting macroeconomic bias and model misspecification bias that are likely to impact CECL implementation during crisis periods. We look back at more than 20 years of data and evaluate model performance during the Great Recession as well as the COVID-19 pandemic. Both events share some similarities, but the COVID-19 episode differs substantially on the level of government assistance directly to retail borrowers. We observe that models usually underperform when presented with uncharted economic environments characterized by out-of-sample macroeconomic conditions. However, when models are re-estimated with additional data that includes some exposure to the new macroeconomic environment performance can improve significantly. Econometric theory can offer insight into the sources of model underperformance. Specifically, macroeconomic risk drivers were a significant source of model misspecification during the pandemic. As a result, models without macroeconomic risk drivers generated default projections more in line with the observed performance during the pandemic. We also observed that certain portfolio risk segments can act as early warning of stress. Furthermore, CECL long-run projections by design average out economic cycles to a certain extent, although short-term economic conditions are a key determinant of CECL allowances.

Insights from our empirical exercise include avoiding overreliance on single models, focusing on the resiliency and adaptability of models and model infrastructure in times of crisis, and considering flexible forecasts and forecast horizons. Simple models, whenever possible, can hold some advantages over more complex models. Simple models may be more robust and easier to diagnose than more complex models. They can also be useful as benchmarks, can provide guidance when overrides or overlays are applied to primary models, and can also help identify areas of weakness in more complex models. There is value in leveraging multiple models and understanding their strengths and weaknesses. There is also value in considering redevelopment or redesign of models in environments that challenge stablished economic relationships. Thus, when building models and model infrastructures, it is important to consider resiliency and adaptability to new shocks. While we argue in favor of a flexible model infrastructure, we also acknowledge the challenges that regulated institutions face, especially taking into account expectations about model validation standards. Thus, it is important to plan ahead and to consider every aspect of the model life cycle as part of the model development process.

VIII. References

Bank for International Settlements (2021). "The Procyclicality of Loan Loss Provisions: A Literature Review." Basel Committee on Banking Supervision Working Paper 39.

Bank for International Settlements (2022). "Newsletter on COVID-19 Related Credit Risk Issues." www.bis.org/publ/bcbs_nl26.htm.

Beck, A. K., and P. J. Beck (2022). "Effects of ASU 2016-13 and COVID-19 on Banks' Loan Loss Allowances." *Advances in Accounting* 56, 100581.

Breeden, J. (2018). "CECL Procyclicality: It Depends on the Model." Working paper, www.prescientmodels.com/articles/CECL-Procyclicality.pdf.

Chae, Sarah, R. Sarama, C. Vojtech, and J. Wang (2018). "The Impact of the Current Expected Credit Loss Standard (CECL) on the Timing and Comparability of Reserves." Finance and Economics Discussion Series Working Paper 2018-020, https://www.federalreserve.gov/econres/feds/files/2018020pap.pdf.

Canals-Cerdá, J. (2020). "From Incurred Loss to Current Expected Credit Loss (CECL): A Forensic Analysis of the Allowance for Loan Losses for Credit Cards Portfolios." *Journal of Credit Risk* 16(4), December.

Canals-Cerdá, J. (2022). "Can We Take the "Stress" Out of Stress Testing? Applications of Generalized Structural Equation Modeling to Consumer Finance." *Journal of Risk Model Validation, Volume 16, N.2.*

Canals-Cerdá, J. and B.J. Lee (2023). "Who Provides Credit in Times of Crisis? Evidence from COVID-19 and the Auto Loan Market." The Fifth Biennial Conference on Auto Lending, Federal Reserve Bank of Philadelphia, July 12–13, 2022.

Chen, J., Y. Dou, S. G. Ryan, and Y. Zou (2022). "Does the Current Expected Credit Loss Approach Decrease the Procyclicality of Banks' Lending?" Working paper, dx.doi.org/10.2139/ssrn.4110109.

Covas, Francisco, and W. Nelson (2018). "Current Expected Credit Loss: Lessons from 2007-2009." Banking Policy Institute Working Paper, bpi.com/wp-content/uploads/2018/07/CECL-Lessons-2007-2009-WP-July-12-2018.pdf.

De Araujo, D. K. G., B. H. Cohen, and P. Pogliani (2021). "Bank Loan Loss Provisioning During the COVID Crisis," www.bis.org/publ/qtrpdf/r_qt2103w.htm.

Degryse, H., and C. Huylebroek (2023). "Preventing a Banking Crisis: Fiscal Support and Loan Loss Provisions During the COVID-19 Pandemic." *Journal of Financial Stability 67, August 2023.*

DeRitis, C., and M. Zandi (2018) "Gauging CECL Cyclicality," Moody's Analytics Analysis Paper, www.economy.com/mark-zandi/documents/2018-12-03-Gauging-CECL-Cyclicality.pdf.

Financial Stability Forum (2009). *Report of the Financial Stability Forum on Addressing Procyclicality in the Financial System*. Basel, Switzerland: Financial Stability Board, www.fsb.org/wp-content/uploads/r_0904a.pdf.

Frame, S., K. Gerardi, and P. S. Willen (2015). "The Failure of Supervisory Stress Testing: Fannie Mae, Freddie Mac, and OFHEO," Federal Reserve Bank of Boston Working Paper 15-4.

Gerardi, K., A. Lehnert, S. Sherlund, and P. Willen (2008). "Making Sense of the Subprime Crisis." Brookings Papers of Economic Activity (Fall 2008).

Haoran Zhang, Harvineet Singh, Marzyeh Ghassemi, Shalmali Joshi (2023). "Why Did the Model Fail?": Attributing Model Performance Changes to Distribution Shifts." Working paper. arxiv.org/pdf/2210.10769.pdf.

Hendry, D. F., Mizon, G. E., 2014. "Unpredictability in economic analysis, econometric modeling and forecasting." *Journal of Econometrics* 182 (2014) 186–95.

Kumar, Pankaj, Marie-Paule Laurent, Christophe Rougeaux, and Maribel Tejada (2022). "Model Risk Management 2.0 Evolves to Address Continued Uncertainty of Risk-Related Events." *McKinsey on Risk* 12 (April).

Loudis, Bert, and B. Ranish (2019). "CECL and the Credit Cycle." Finance and Economics Discussion Series 2019-061. Washington, D.C.: Board of Governors of the Federal Reserve System, doi.org/10.17016/FEDS.2019.061.

Loudis, B., S. Pechenik, B. Ranish, C. M. Vojtech, and H. Xu (2021). "New Accounting Framework Faces Its First Test: CECL During the Pandemic." *FEDS Notes* (December 3), www.federalreserve.gov/econres/notes/feds-notes/new-accounting-framework-faces-its-first-test-cecl-during-the-pandemic-20211203.htm.

Pinello, A., and E. L. Puschaver (2020). "CECL Encounters a 'Perfect Storm.'" *CPA Journal* 90(7/8), 20–8.

Pinello, A. S., and L. Puschaver (2022). "Implementing CECL During the Pandemic: Unexpected Consequences Revealed." *CPA Journal* 92(1/2), 20–7.

Rosenblum, S. and P. Lai (2020). "CECL Adoption and Q1 Results Amid COVID-19," available online at: www.moodysanalytics.com/articles/2020/cecl-adoption-and-q1-results-amid-covid-19.

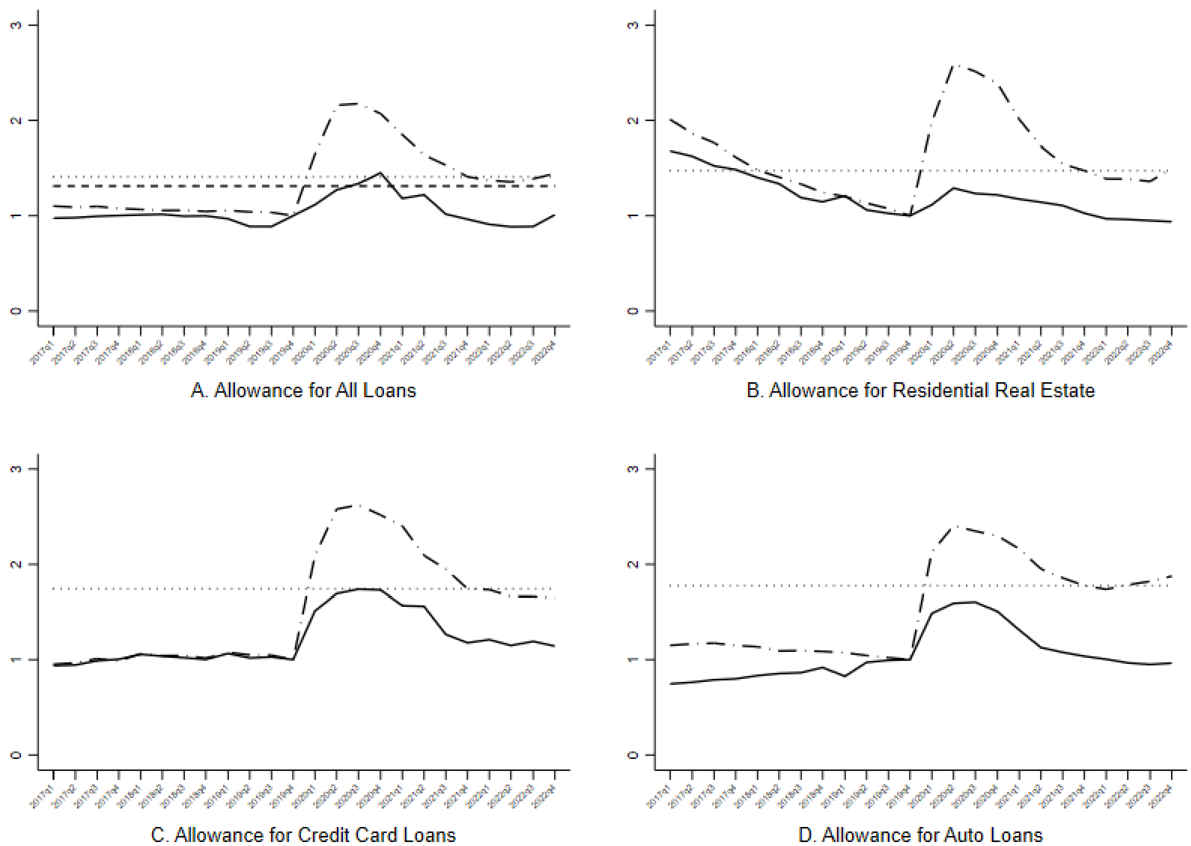
Wall, L. D., (2020). "Banks' Accumulated Credit Losses in the First Quarter." *Notes from the Vault*, www.atlantafed.org/cenfis/publications/notesfromthevault/05-banks-accumulated-credit-losses-in-the-first-quarter-2020-05-29.

Waller C. J. (2021). "A Hopeless and Imperative Endeavor: Lessons from the Pandemic for Economic Forecasters," speech at the Forecasters Club of New York (December 17), www.federalreserve.gov/newsevents/speech/waller20211217a.htm.

IX. TABLES AND FIGURES

Figure 1: ALLL During the Pandemic, CECL Adopters and Nonadopters

Allowances reported as a percentage of the allowances in the fourth quarter of 2019 for CECL adopters (dash/dot-dash line) and CECL nonadopters (solid line). We include the day one impact in graph A (dash horizontal line). Also, as an additional reference, we include the CECL allowances in the fourth quarter of 2021 (dotted horizontal line), a quarter of mild economic conditions.



Data source: Y9C public submissions.

Table 1: Disaggregated Quarterly Changes in ALLL with Respect to 2019Q1

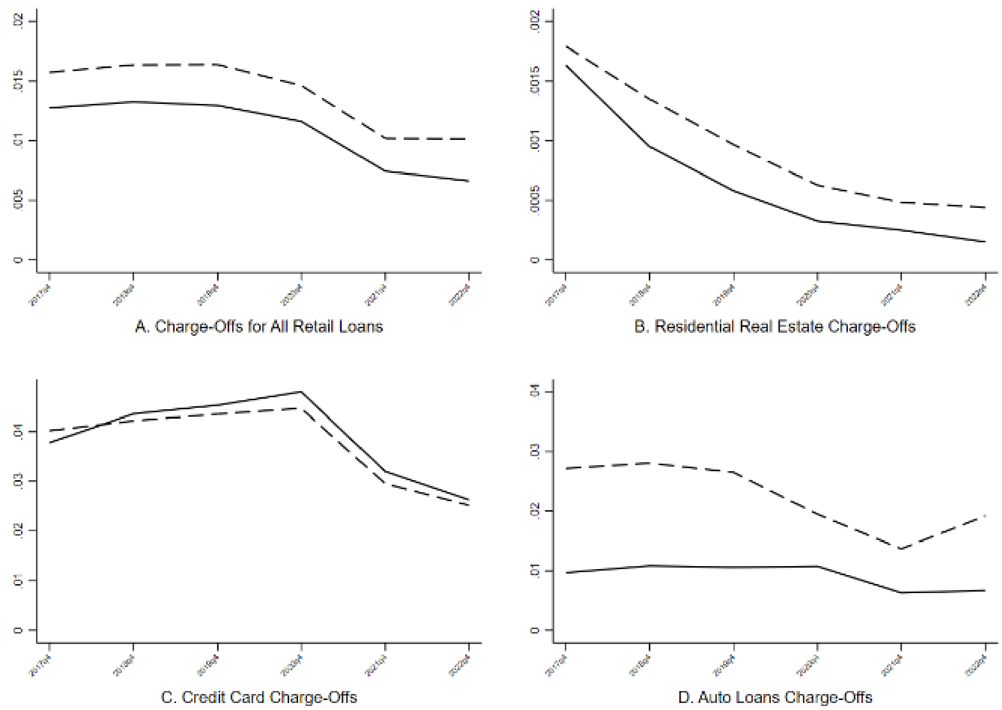
The top panel reports the relative percentual increase in allowances for CECL nonadopters, while the bottom panel represents the percentual increase in allowances for CECL adopters, with 1 representing ALLL in 2019Q1. Values in parentheses represent the incremental value of ALLL for CECL adopters vs. nonadopters.

VARIABLES	Residential Loans	Credit Cards	Auto Loans
	Non adopters		
2020Q1	1.13	1.49	1.46
2020Q2	1.29	1.68	1.57
2020Q3	1.25	1.71	1.58
2020Q4	1.21	1.69	1.49
2021Q1	1.15	1.53	1.31
2021Q2	1.13	1.54	1.12
2021Q3	1.09	1.22	1.06
2021Q4	1.02	1.13	1.02
2022Q1-Q4	0.99	1.13	0.96
	CECL adopters		
2020Q1*CECL	2.01 (0.88)	2.16 (0.67)	2.22 (0.76)
2020Q2*CECL	2.63 (1.34)	2.68 (1.00)	2.66 (1.09)
2020Q3*CECL	2.53 (1.28)	2.72 (1.01)	2.54 (0.96)
2020Q4*CECL	2.45 (1.24)	2.60 (0.91)	2.49 (1.00)
2021Q1*CECL	2.03 (0.88)	2.46 (0.93)	2.27 (0.96)
2021Q2*CECL	1.78 (0.65)	2.13 (0.59)	2.03 (0.91)
2021Q3*CECL	1.6 (0.51)	1.98 (0.76)	1.93 (0.87)
2021Q4*CECL	1.53 (0.51)	1.77 (0.64)	1.84 (0.82)
2022Q1-Q4*CECL	1.45 (0.46)	1.70 (0.57)	1.88 (0.92)
R-squared	0.78	0.96	0.97

Data source: Y9C public submissions.

Figure 2: Charge-Off Rates During the Pandemic

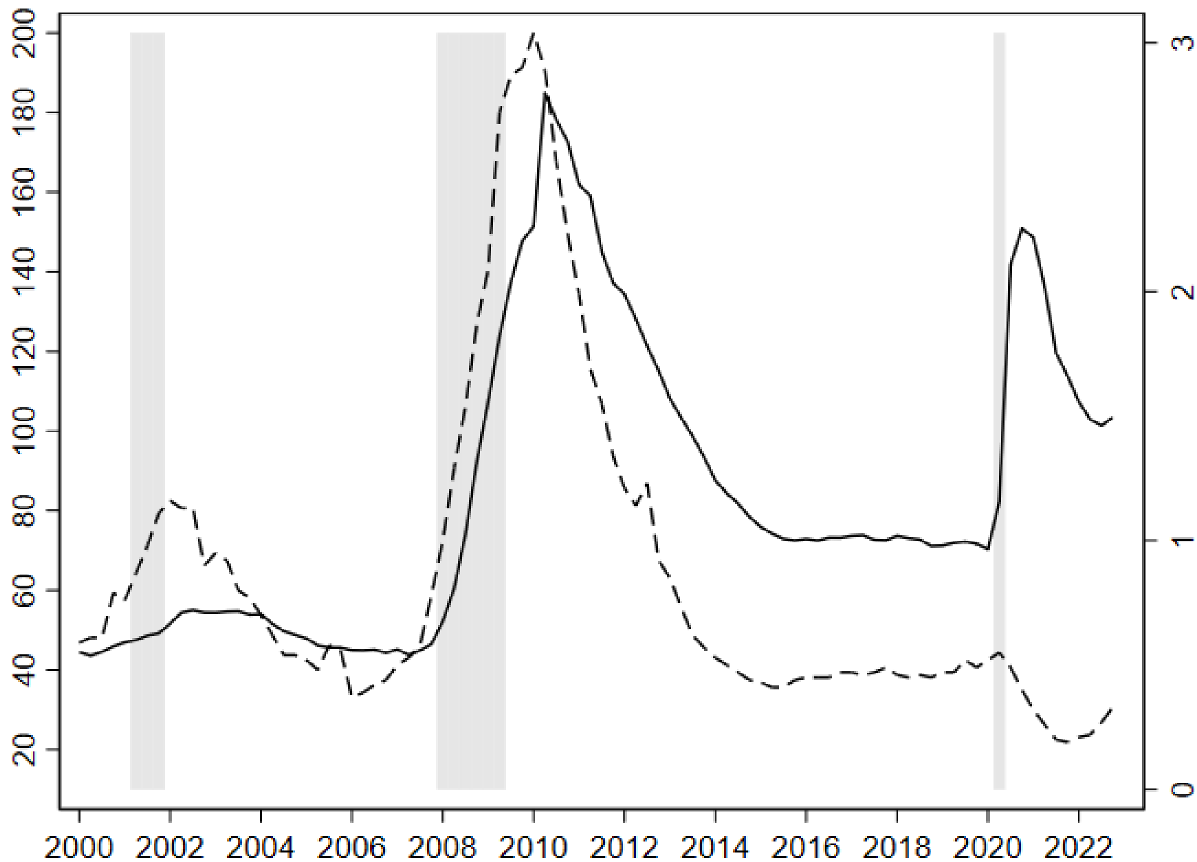
Charge-off rates for retail portfolios for CECL adopters (dash line) and nonadopters (solid line).



Data source: Bank Y9C public submissions.

Figure 3: Historical Allowances and Charge-Off Rates

The figure depicts the aggregated charge-off rate on all loans at all commercial banks (dashed line, right axis) and allowances for loan and lease losses, large domestically chartered commercial banks (solid line, left axis).[32]

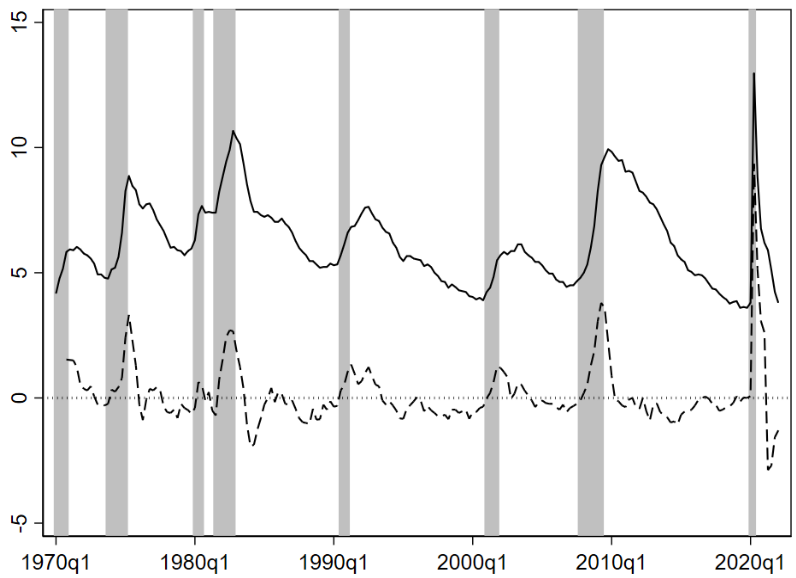


Data source: fred.stlouisfed.org/.

[32] The allowance for all commercial banks follows a similar pattern to the allowance for large commercial banks, we report the allowance for large commercial banks here because of the availability of historical data in Fred prior to the Great Recession.

Figure 4: Professional Forecasters' Error

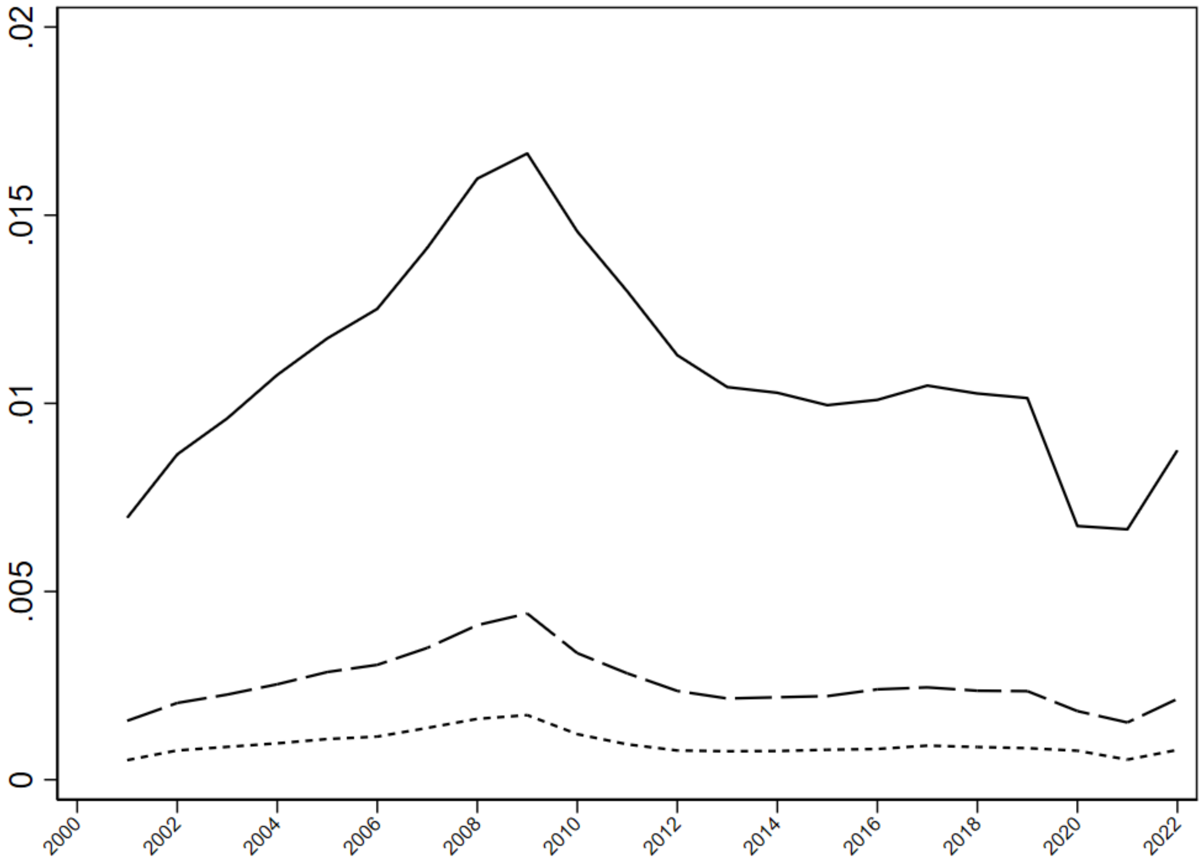
The figure depicts the realized unemployment rate, the four-quarter ahead unemployment rate forecast, and the forecast error. Forecasts are from the Philadelphia Fed's *Survey of Professional Forecasters*.[33] The solid line represents the unemployment rate; the dashed line represents the one-year-ahead unemployment forecast error. The forecast error was 4 percent during the Great Recession and up to 9 percent during the COVID-19 lockdown.



[33] Figure from "From Incurred Loss to Current Expected Credit Loss (CECL): A Forensic Analysis of the Allowance for Loan Losses for Credit Cards Portfolios." *Journal of Credit Risk 16:4, December 2020.*

Figure 5: 30+, 60+ and 90+ Delinquency Rates

The figure depicts changes over time in 30+, 60+ and 90+ delinquencies in auto loans.



Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

Figure 6: State Unemployment Rate Over Time

The charts depict kernel density functions that illustrate the variation in the unemployment rate over time and across states for different time periods. The top figures depict the distribution of state unemployment rate over the periods 2001–19 and 2001–10, respectively. The top-right figure depicts the distribution of state unemployment for the periods 2001–07 (solid line), 2008 (long dashed line), 2009 (dotted line) and 2010 (dashed line). The bottom-left figure depicts 2011 (solid line), 2012 (long dashed line), 2014 (dotted line) and 2014 to the first two months of 2020 (dashed line). Finally, the bottom-right figure depicts March 2020 to May 2020 (solid line), June 2020 to December 2020 (long dashed line), full-year 2021 (dotted line), and full-year 2022 (dashed line).

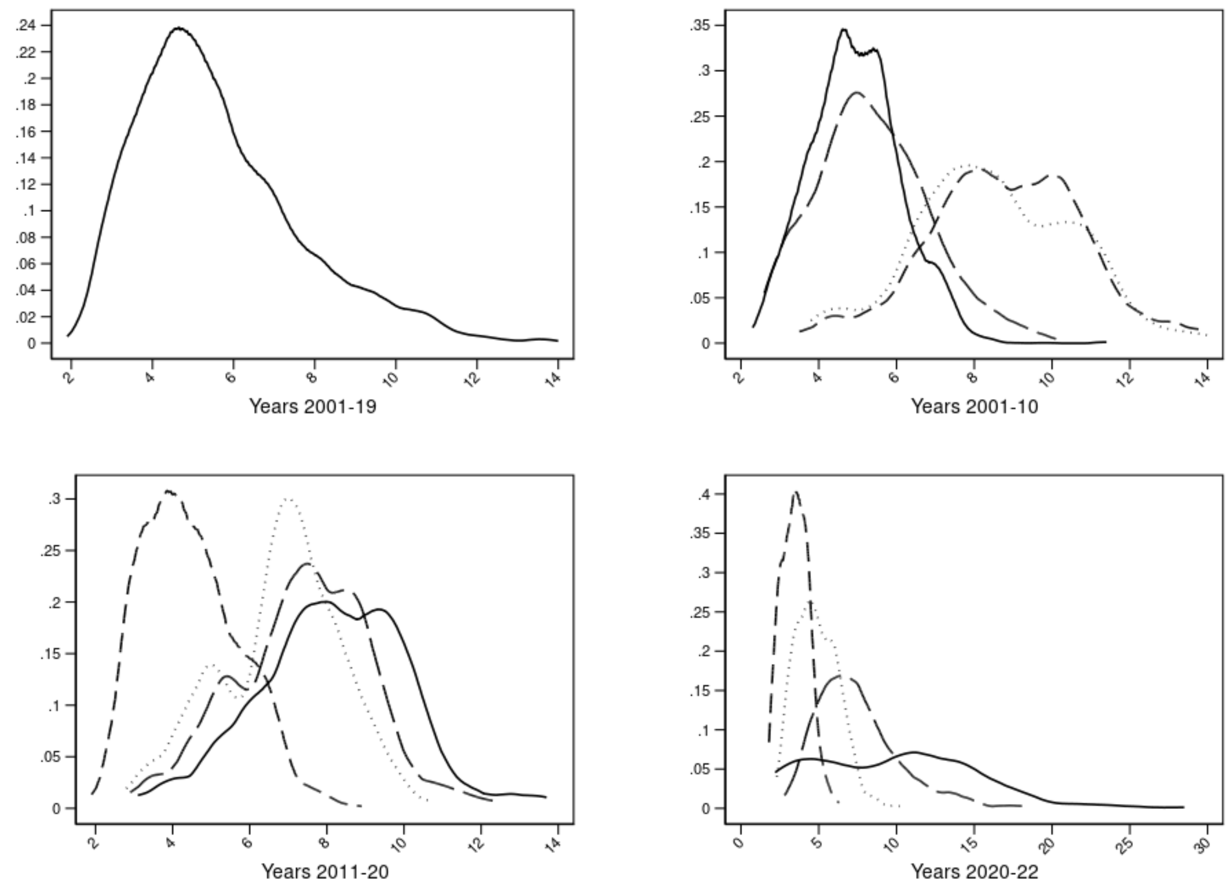
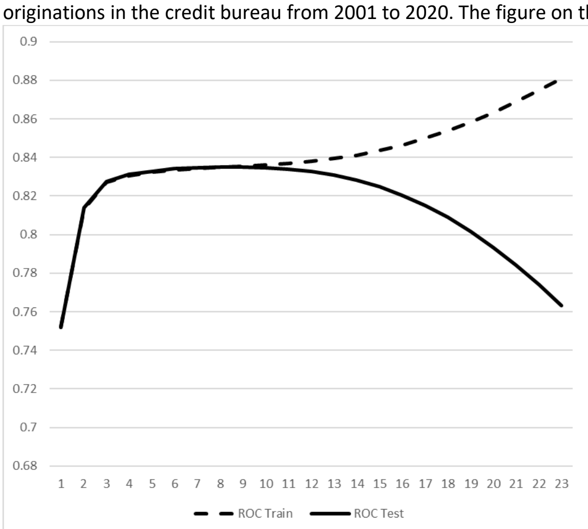
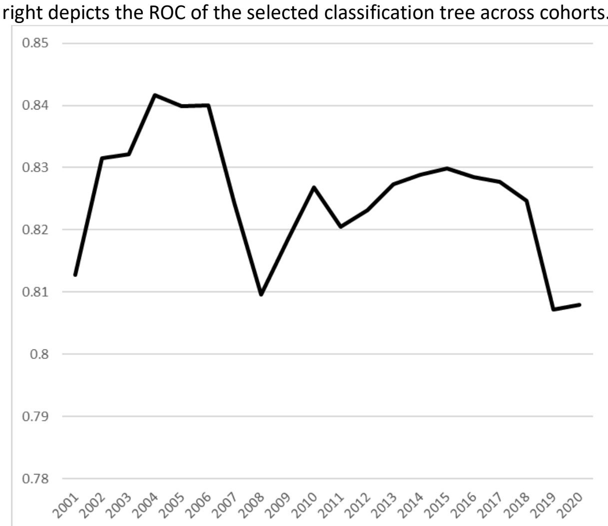


Figure 7: ROC Performance Across Models and Over Time

The figure on the left depicts receiver operating characteristic (ROC) metrics for a decision tree classifier of the two-year forward-looking default, as the maximum depth of the tree increases from one to 23, for test and training data sets from the overall population of auto loan originations in the credit bureau from 2001 to 2020. The figure on the right depicts the ROC of the selected classification tree across cohorts.



a.- ROC performance across models.

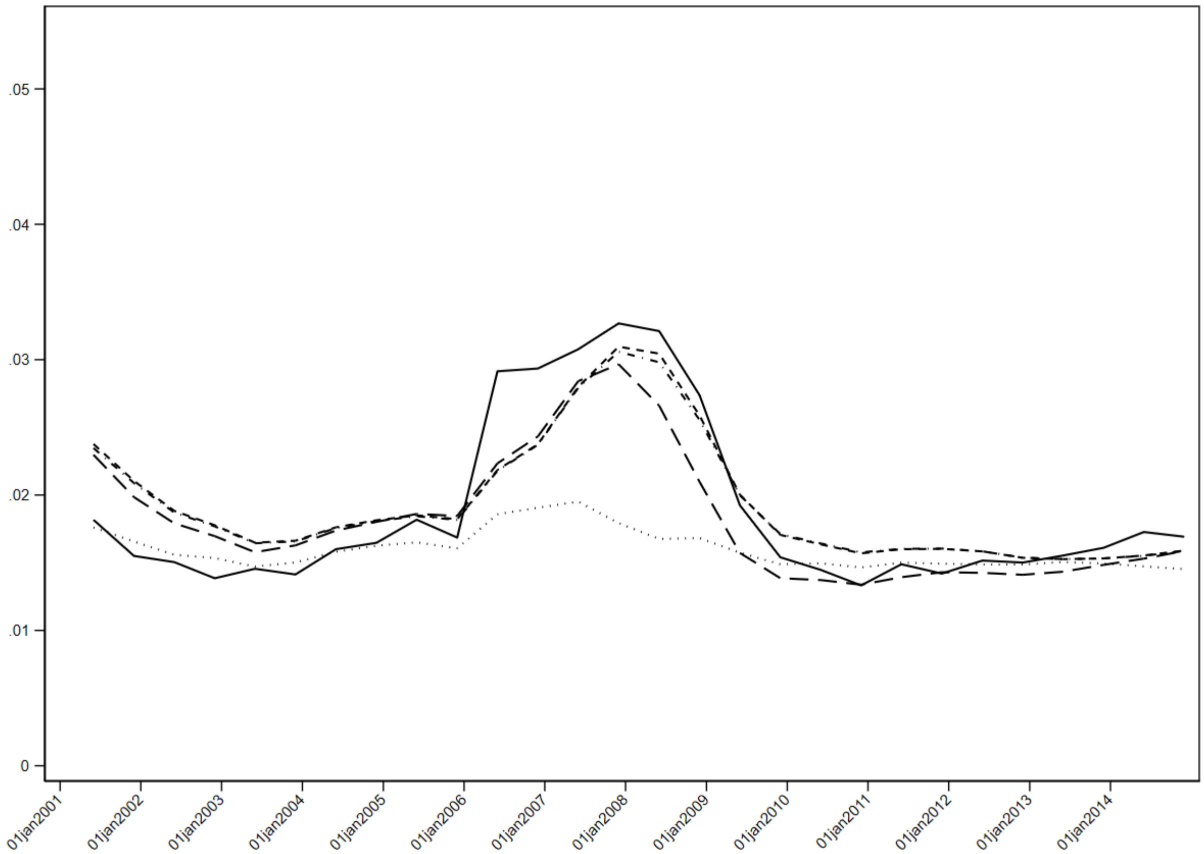


b.- Two-year ROC performance over time.

Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

Figure 8: Cumulative Default Rates Across Cohorts

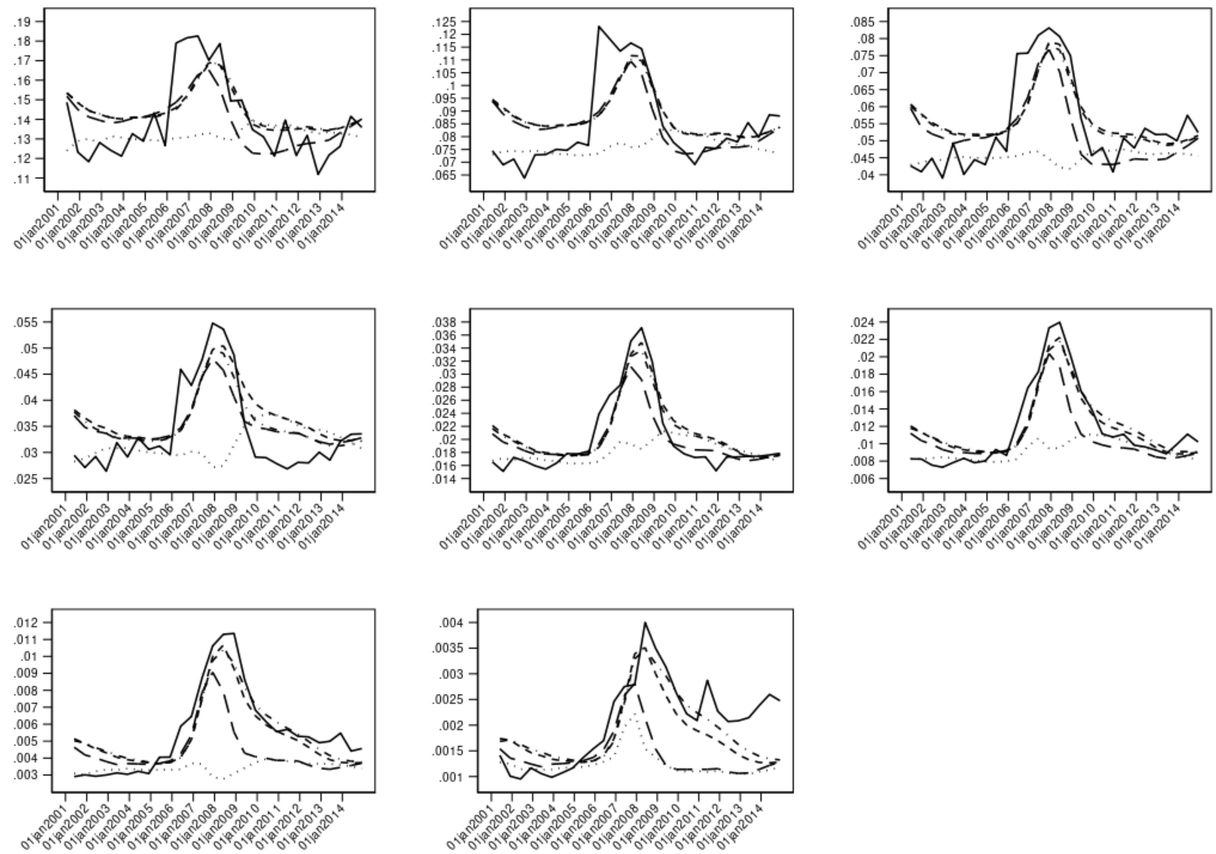
The figure depicts realized nine quarters cumulative default rates across cohorts (solid line), as well as forecasted values for models estimated with data including nine quarters of performance from the 2001–05 cohorts (dotted line), 2001–07 cohorts (long dash line), 2001–08 cohorts (dash line) and 2001–09 cohorts (dot-dash line).



Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

Figure 9: Cumulative Default Rates Across Cohorts and Risk Segments

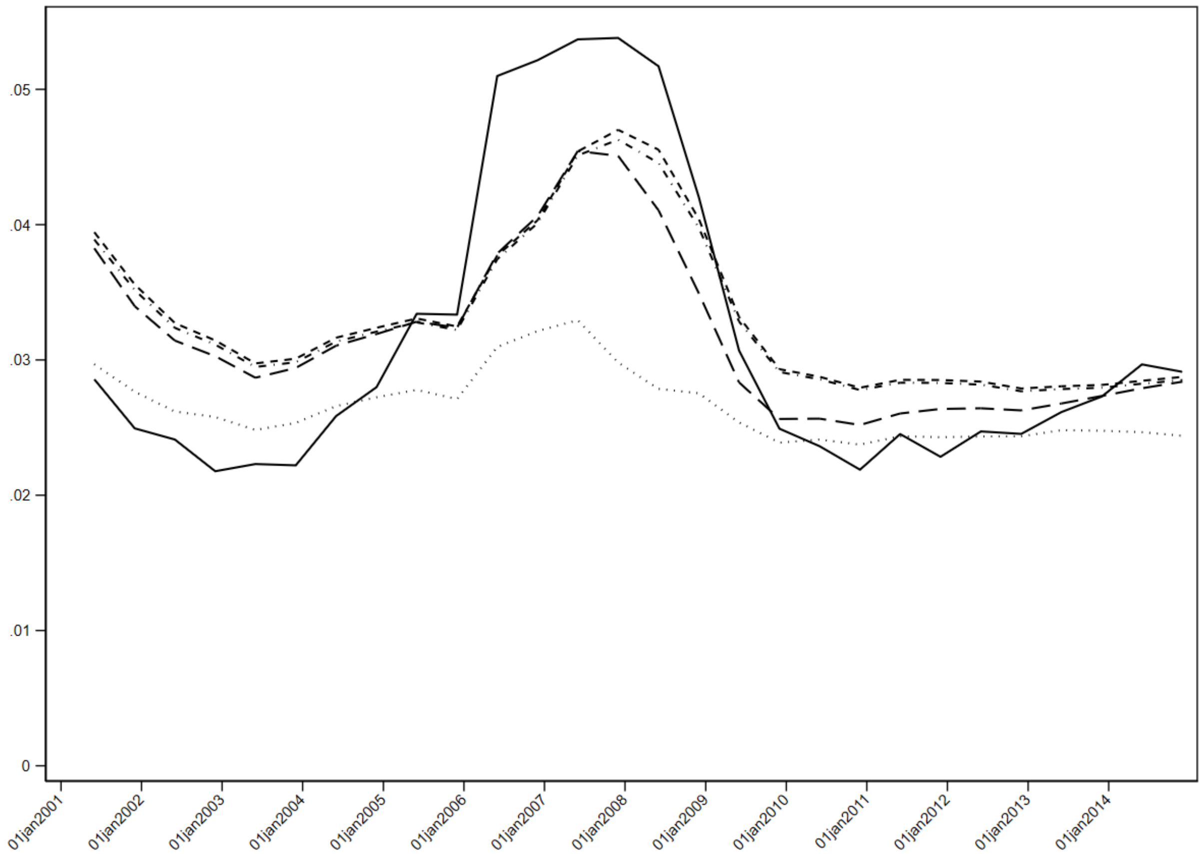
This figure expands on figure 8.a by depicting cumulative default rates across risk segments, for segments with decreasing risk from left to right and from top to bottom. Each individual chart depicts the realized nine-quarter cumulative default rates across cohorts (solid line), as well as forecast values for models estimated with data including nine quarters of performance from the 2001–05 cohorts (dotted line), 2001–07 cohorts (long dashed line), 2001–08 cohorts (dashed line), and 2001–09 cohorts (dot-dash line).



Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

Figure 10: Lifetime Default Rates Across Cohorts

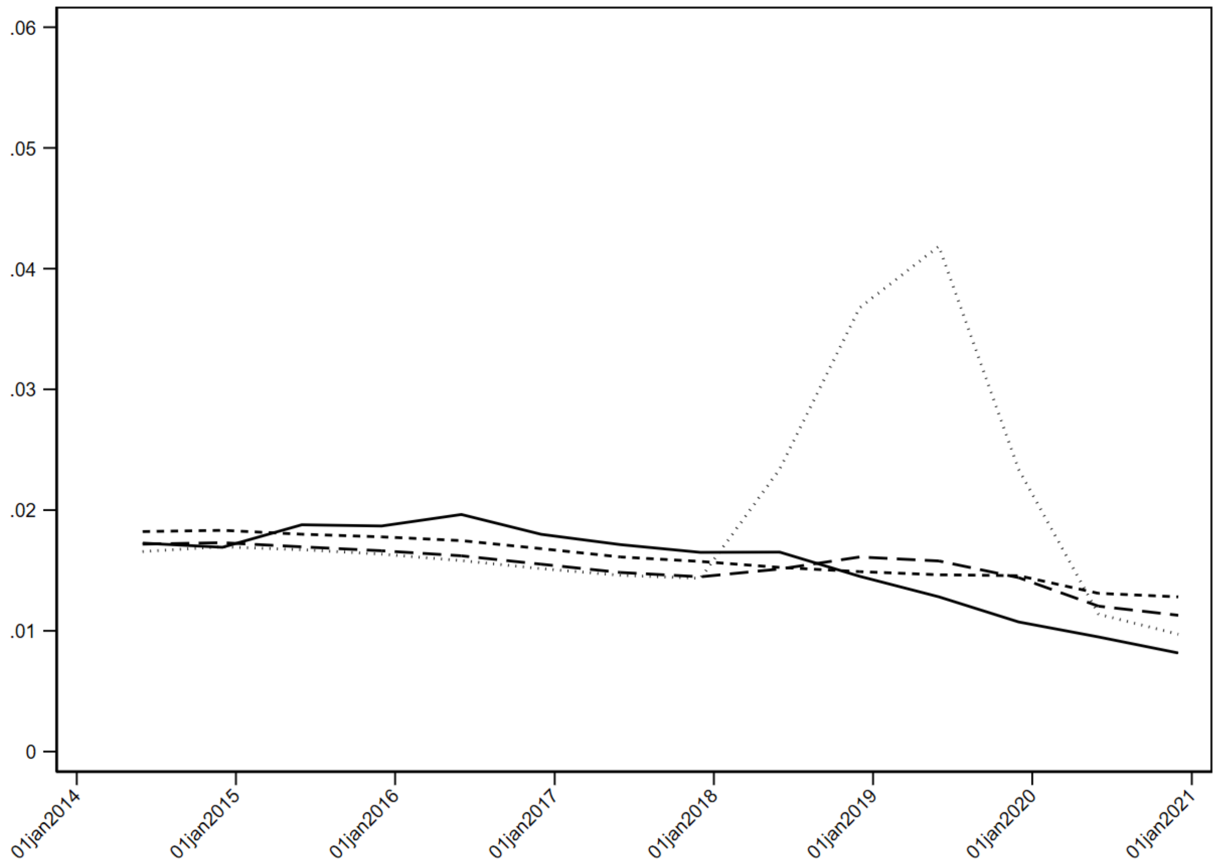
The figure depicts the realized lifetime cumulative default rates across cohorts (solid line), as well as forecast values for models estimated with data including nine quarters of performance from the 2001–05 cohorts (dotted line), 2001–07 cohorts (long dashed line), 2001–08 cohorts (dash line), and 2001–09 cohorts (dot-dash line). In the analysis, we define the life of the loan as the smaller of five years or time to payment.



Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

Figure 11: Cumulative Default Rates Across Cohorts, Including the COVID-19 Period

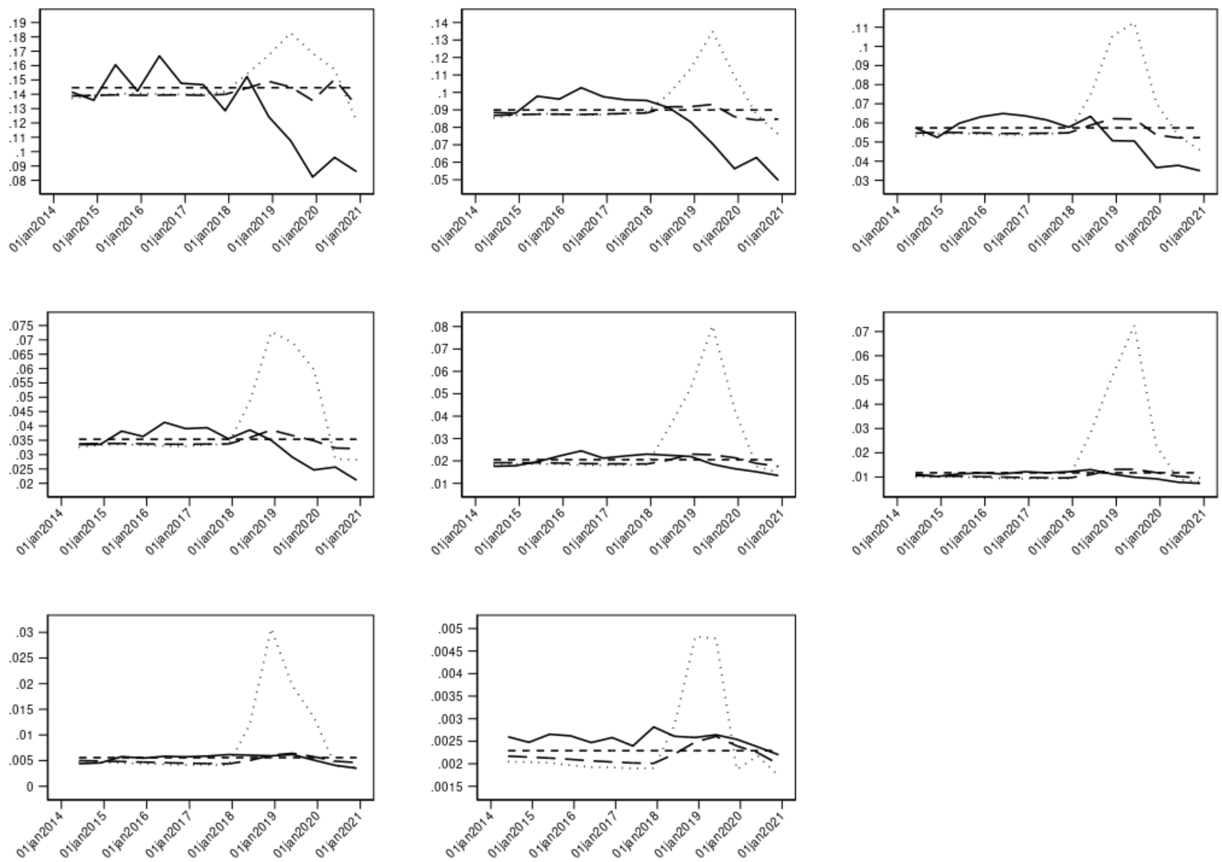
The figure depicts the realized nine-quarter cumulative default rates across cohorts (solid line), as well as forecast values for models estimated with data including nine quarters of performance from the 2001–17 cohorts (dotted line), 2001–20 cohorts (long dashed line), and 2001–17 cohorts without macro variables (dashed line).



Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

Figure 12: Cumulative Default Rates Across Cohorts and Risk Segments, Including the COVID-19 Period

This figure expands on figure 11 by depicting cumulative default rates across risk segments, for segments with decreasing risk from left to right and from top to bottom. The figure depicts the realized nine-quarter cumulative default rates across cohorts (solid line), as well as forecast values for models estimated with data including nine quarters of performance from the 2001–17 cohorts (dotted line), 2001–20 cohorts (long dashed line), and 2001–10 cohorts without macro variables (dashed line).



Data source: Auto tradeline data from the FRBNY Consumer Credit Panel/Equifax.

A. APPENDIX: Regulatory Guidance on CECL Implementation.

As FASB staff has indicated in multiple instances, the CECL standard allows for flexibility in determining the best approach for computing the allowance. CECL is by design nonprescriptive about the methodology that should be employed when computing the allowance, as well as the economic projections that should be considered when determining the reasonable and supportable forecast. This level of flexibility is intended to facilitate CECL implementation across financial institutions with different levels of complexity.

For the less sophisticated financial institutions, banking regulators have contributed examples of acceptable methodologies, like the snapshot/open pool approach, the vintage approach, and the remaining life/weighted average remaining maturity (WARM) approach.[34] The methods differ primarily on the way the lifetime historical charge-off rate is calculated. For example, the snapshot approach computes the lifetime historical charge-off rate as the ratio of total lifetime charge-offs associated with the snapshot loan portfolio to loan portfolio balance. Specific adjustments to current conditions, and reasonable and supportable forecasts should be considered when computing the CECL allowances using these simpler methods.[35] A FASB staff Q&A transcript clarifies that it is acceptable to adjust historical loss information for current and future forecast economic conditions through a qualitative approach properly documented.[36]

The Federal Reserve has developed a simple Excel-based tool to assist smaller community banks with total assets of less than $1 billion in calculating their allowances under CECL. This method, known as SCALE (scaled CECL allowance for losses estimator), uses publicly available Call Report data to derive expected lifetime credit loss rates. The Federal Reserve has also developed an Excel-based expected loss estimator (ELE) tool for the WARM method, primarily intended for community financial institutions.[37] The method allows for the use of a financial institution's own

[34] https://www.supervisionoutreach.org/cecl/methodologies-and-examples
[35] Additional details can be found in the following interagency slide presentation: www.supervisionoutreach.org/-/media/files/supervisionoutreach/cecl/22718-ask-the-regulator-presentation.pdf?sc_lang=en&hash=95EEAD092807060791975C482B16B553
[36] https://www.fasb.org/page/PageContent?pageId=/standards/Transition/credit-losses-transition/fasb-staff-qatopic-326-no-1whether-the-weightedaverage.html
[37] https://www.supervisionoutreach.org/cecl/ele

loan data. WARM has been reviewed by FASB staff and deemed one of many methods that could be used to estimate allowances for less complex financial asset pools.[38]

Complex financial institutions may consider more sophisticated model frameworks including discounted cash flow approaches, roll rate approaches, and methodologies that decompose losses in terms of the probability of default (PD), loss given default (LGD), and exposure at default (EAD). On the one hand, sophisticated modeling frameworks can better accommodate changes in portfolio characteristics and macroeconomic scenarios. On the other hand, they may be more sensitive to model and forecasting error that may be difficult to diagnose and troubleshoot, in part because of the intricacy of the modeling framework. Challenges to CECL models may be particularly severe in times of crisis.

[38] https://www.fasb.org/page/PageContent?pageId=/standards/Transition/credit-losses-transition/fasb-staff-qatopic-326-no-1whether-the-weightedaverage.html