Chaire en macroéconomie et prévisions

Forecasting Bank Failures in a Data-Rich Environment

Jean-Armand Gnagne[1] et Kevin Moran[2]

Cahier de recherche
Working paper
2020-13

Juin / June 2020

ESG UQÀM
École des sciences de la gestion
Université du Québec à Montréal

[1] Université Laval
[2] Université Laval

Chaire en macroéconomie et prévisions

La Chaire en macroéconomie et prévisions est fière de s'appuyer sur un partenariat avec les organisations suivantes:

Les opinions et analyses contenues dans les cahiers de recherche de la Chaire ne peuvent en aucun cas être attribuées aux partenaires ni à la Chaire elle-même et elles n'engagent que leurs auteurs.

Opinions and analyses contained in the Chair's working papers cannot be attributed to the Chair or its partners and are the sole responsibility of the authors.

Forecasting Bank Failures in a Data-Rich Environment*

Jean Armand Gnagne
Université Laval

Kevin Moran
Université Laval

May 31, 2020

Abstract

This paper develops a monitoring and forecasting model for the aggregate monthly number of commercial bank failures in the U.S. We extract key sectoral predictors from the large set of macroeconomic variables proposed by McCracken and Ng (2016) and incorporate them in a hurdle negative binomial (HNB) model to predict the number of monthly commercial bank failures. Our in-sample analysis uncovers a significant and robust relationship between the predictor related to the housing sector and the occurrence of bank failures, suggesting the importance of the link between developments in that sector and banking vulnerabilities. Out-of-sample exercices, conducted by sequentially re-estimating our HNB model at every step using the real-time vintages of the McCracken and Ng (2016) data, confirm the value of our forecasting approach, which outperforms other alternatives.

JEL Classification: E60, F37, F38, G01.

Keywords: Financial Regulation, Financial Crises, Factors Models, Diffusion Index Models.

*We thank Benoît Carmichael, Gilles Koumou, Jonathan Morneau-Couture and Christian Friedrich for useful comments and discussions. Any errors and omissions are own.

1 Introduction

Banking crises and other episodes of financial market distress have important macroeconomic consequences: they disrupt the flow of credit, raise the risks of corporate and personal failures, lead to output losses relative to trend and to sharp declines in tax revenues and the fiscal health of governments. The 2007-2009 crisis has reaffirmed this fact and caused significant worldwide economic damage.[1]

Considering the costs they generate, an important body of work has sought to analyze banking crises and identify "early warning" variables –key factors associated with heightened crisis probabilities– forecasting future crisis occurences. This body of work, originating in contributions such as Demirgüç-Kunt and Detragiache (1998), Kaminsky and Reinhart (1999) or Borio and Lowe (2002), has been energized by the 2007-2009 events and has since grown considerably.[2]

This paper provides an original and complementary contribution to the literature studying banking crises. We develop a count-data framework to model the monthly aggregate number of bank failures in the United States and assess its out-of-sample forecasting ability. Specifically, we employ a hurdle-negative binomial (HNB) model, an extension of the standard Poisson count-data process designed to accommodate high frequencies of zero counts (the absence of bank failures in a given month for our data) and the high dispersion in the number of such failures, when they do occur. Our explanatory variables are factors extracted from the several dozen time-series in the McCracken and Ng (2016) database, in accordance with the literature showing how a few predictors summarizing information contained in a large number of variables can outperform other forecasting approches (Stock and Watson, 2002a,b, 2006; Bai and Ng, 2008, 2009).

Our in-sample results demontrate that the HNB model outperforms alternatives like the standard Poisson counterpart. It also shows that the factor related to the housing industry block in the McCracken and Ng (2016) database contains the most robust, statistically and economically meaningful information about the future occurrence of bank failures. This echoes findings in related research showing links between the housing sector and banking industry disruptions (Barrell et al., 2010; Bernanke, 2013; Ghosh, 2015). It

[1] Reinhart and Rogoff (2013) and Laeven and Valencia (2013) present assessments of the fiscal consequences of banking crises. In addition, Laeven and Valencia (2013) document the extent to which economies suffer output and bank equity losses following such crises. See also Hutchison and McDill (1999) for an exploration of the consequences of Japanese banking crises.

[2] A non-exhaustive review of recent contributions includes Bussiere and Fratzscher (2006), Davis and Karim (2008), Borio and Lowe (2009), Barrell et al. (2010), Barrell et al. (2010), Duca and Peltonen (2013), Betz et al. (2014), Gogas et al. (2018) or Antunes et al. (2018).

also accords well with intuitive links between the boom-and-bust cycles in housing and the health of the banking sector, with the boom featuring rising home values, loosening of lending and refinancing terms and rising bank exposure to loan defaults and the bust, by contrast, being characterized by house price decreases, households, increases in mortgage delinquencies or defaults, and hightened bank vulnerabilities.

Next, we assess the out-of-sample forecasting ability of our approach as follows. We extract factors and estimate our model on the sample $1975M1 - 1999M12$ using the real-time vintages of the McCracken and Ng (2016) database and then forecast the number of bank failures for up to twelve months ahead. We repeat this procedure using expanding windows of data (one new observation each month), recomputing the factors at each step using the appropriate data vintage. We then repeat this analysis for a variety of alternative models. We document that our HNB framework has the best forecasting ability at all horizons and that most of these superior performances are statistically signifiant according to the Diebold and Mariano (1995) test.

Our approach is a contribution to the general research on banking crises and as such offers three potential advantages to more conventional strategies in that literature. First, using the aggregate number of bank failures as the proxy for crises provides an interesting alternative to measures used elsewhere.[3] Second, the *monthly* frequency of our framework can provide regulatory authorities early insights about developing financial vulnerabilities, relative to other work using annual data on bank failures (Davutyan, 1989). Finally, a framework to monitor and forecast the aggregate occurrence of bank failures in the United States is important in its own right, particularly for institutions such as the Federal Deposit Insurance Corporation (FDIC) mandated with such monitoring responsibilities.

The rest of this paper is organized as follows. Section 2 briefly reviews the literature on the determinants of bank failures and other forms of banking-sector distress. Section 3 describes the data used in our contribution. Section 4 presents the econometric framework and Section 5 the results. Section 6 concludes.

[3] How to actually measure banking crises is a recurring challenge. The database of Laeven and Valencia (2013) notes banking crises with dummy variables indicating "significant signs" of financial distress in banking systems (bank runs, losses and bank liquidations) or "significant banking policy interventions". This is related to Reinhart and Rogoff (2013)'s measure where a bank run or a government assistance to banks (closure, merging and other large-scale regulatory actions) defines crises. Other measures add additional distress signals –such as nonperforming banking assets– to define crises (Demirgüç-Kunt and Detragiache, 1998, 2005).

2 Monitoring and forecasting bank failures

Monitoring financial systems is one key task of regulatory authorities and has typically focused on bank-specific, industry-specific and macroeconomic determinants of bank failure. We hereafter briefly review these areas of analysis.

Poor management is seen as playing the major role among bank-specific factors leading to bank failures (Berger and DeYoung, 1997; Salas and Saurina, 2002; Podpiera and Weill, 2008). Profit-seeking incentives may sometimes encourage bank managers to take innovative actions that result in poor credit scoring, spurious collateral appraisal, inadequate borrowers monitoring and subpar overall loan quality. A lack of diversification in such activities may also exacerbate these problems, with diversification usually proxied by the proportion of non-interest income as a share of total income and expected to to be negatively related to non-performing loans. Finally, insufficient loan loss provisions may reflect the overall disinterest of banks towards risks control as increases in such provisions could be perceived by investors and shareholders as signals of trouble and bad management.

Researchers have also identified important industry-specific factors driving bank failures, related to monetary policy or to banking regulation (Keeton, 1999; Bernanke, 2013). An over-accommodating monetary policy stance characterized by low interest rates and growing money supply may be associated with rapid expansions of credit and subsequent deterioration in credit-allocation standards. In addition, weak banking regulation, such as low capital requirements in a competitive industry as well as generous deposit insurance, may encourage banks managers to take on too much risk. A lively ongoing debate about the impacts of deposit insurance and the role of central banks as lenders of last resort during times of financial system instability is exemplified by contributions in Boyd and Gertler (1994), Stern and Feldman (2004), Ennis and Malek (2005) or Bernanke (2013). Insufficient banking regulation may be exacerbated by the inability of regulators to adequately monitor banking activities. Development of sophisticated financial instruments also add difficulties to the supervision of the banking industry by the regulatory authorities.

Finally, aggregate macro-financial factors also play a key role in financial system stability (Demirgüç-Kunt and Detragiache, 1998; Kaminsky and Reinhart, 1999; Louzis et al., 2012). Sustained output growth and well-anchored inflation are generally positively associated with banking system stability. Low unemployment rate and dynamic housing industry foster booms in banking activities. Breuer (2006) suggests that other national factors such as corruption may also be important.

3 Data

As stated above, this paper's goal is to provide a workable forecasting tool for the aggregate monthly number of commercial bank failures in the United States. To this end, we analyze monthly-frequency data on bank failures and relate them to the information contained in the McCracken and Ng (2016) dataset, which comprises a large set of macro-financial explanatory variables and is updated on a timely basis.[4] In addition, our out-of-sample experiment employs the real-time vintages of the McCracken and Ng (2016) dataset, which reflect the information historically available at each given period of time.

3.1 The monthly number of bank failures

Our variable of interest is the monthly number of bank failures and assistances reported by the Federal Deposit Insurance Corporation (FDIC).[5] A bank failure is defined as the closing of a financial institution by its chartering authority, while an assistance pertains to a situation where a failing institution is acquired by another (healthy) institution, possibly with financial assistance from the FDIC.

Figure 1 illustrates the evolution of the US banking industry between 1975 and 2017. As depicted in Panel (a) of the figure, more than 14,000 commercial banks were operating in the United States in the mid 1970s, largely as a result of strict regulations on branching. In the 1980s, progressive easements in branching regulation induced waves of mergers and the number of banks with no branch steadily decreased whereas the number of banks with branches increased till the late 1980s (but has slowly declined since). These two effects combined to create a a significant downward trend in the total number of commercial banks in the United States.

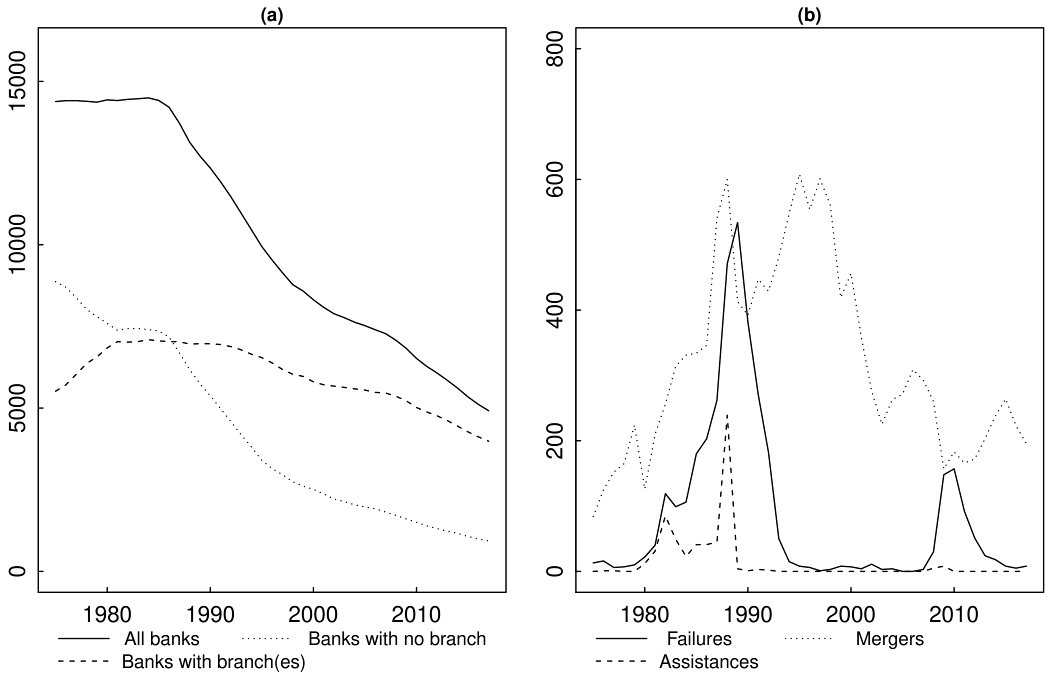
Next, panel (b) of Figure 1 provides the data on failures, assistances and mergers. The evolution of failures and assistances clearly depict the two major disruptive episodes experienced by the U.S. banking system over the last 40 years, namely the *Savings and Loans crisis* (late 1980's) and the *subprime crisis* (2007-2009).

Figure 2 scrutinizes further the monthly number of bank failures and assistances, by depicting the level –Panel (a)– and the ratio of bank failures and assistances to the beginning-of-the year number of banks, Panel (b). The magnitude of the 2007-2009 subprime crisis

[4]See De Nicolo and Lucceta (2016), Smeekes and Wijler (2018) or Forni et al. (2018), among others, for recent uses of the McCracken and Ng (2016) dataset in forecasting.

[5]As the primary deposit insurance provider for US banks, the FDIC supervises both federally-chartered banks as well as most of their state-chartered counterparts. Each insured bank must report to the FDIC and the agency is involved in the majority of proceedings arising from bank failures or assistances.

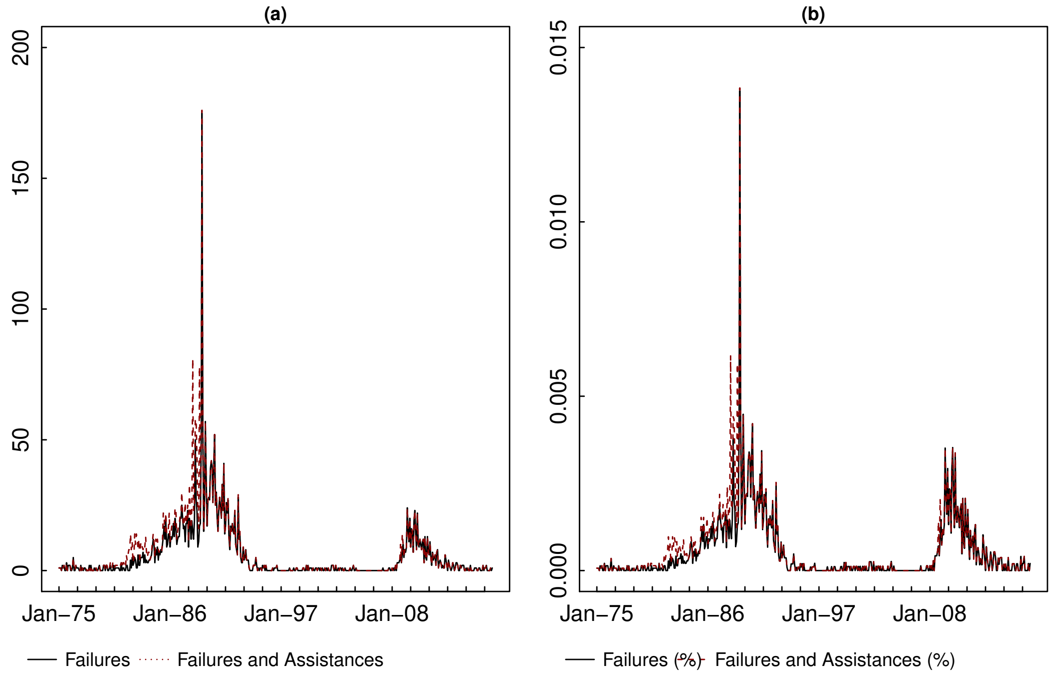
Figure 1: Evolution of the U.S. banking industry: 1975 - 2018



Notes: Data on the U.S. banking industry are expressed in levels and sourced from the Federal Deposit Insurance Corporation. Panel (a) depicts the progressive concentration of the U.S. banking industry. Panel (b) reports mergers as well as failures and assistances of U.S. banks.

naturally appears slightly amplified when bank failures are reported as a proportion of the total number of banks, but the evolution of the two measures appears very similar. Our work below emphasizes the number of bank failures but could be extended to the proportion of bank failures. The within-sample analysis of the same data contained in Gnagne and Moran (2018) suggests however that results would be similar.

Figure 2: U.S. bank failures and assistances (in levels and in proportion of total)



Source: FDIC

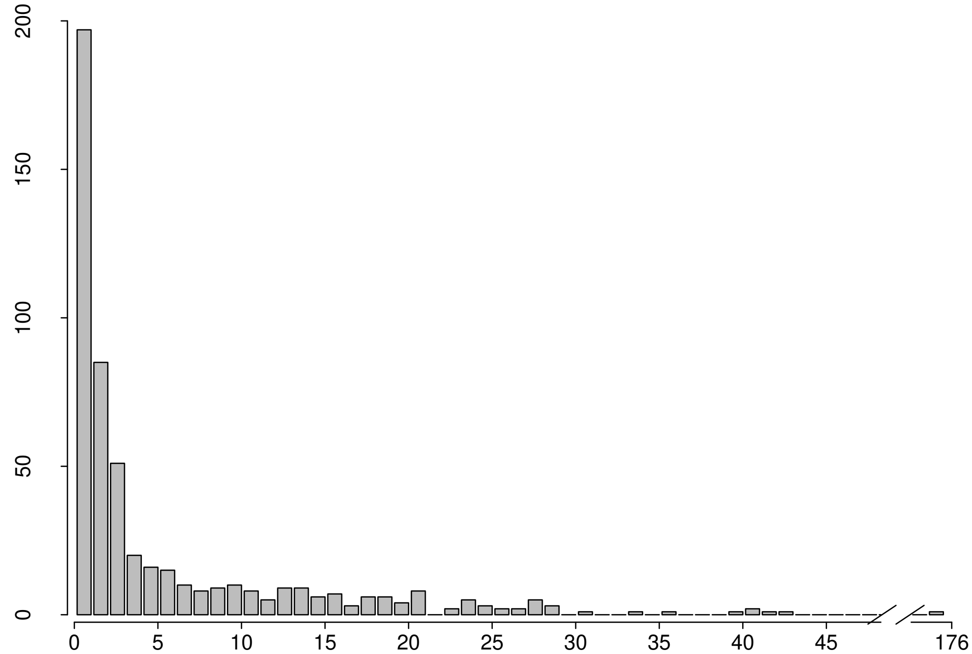
Table 1: U.S. bank failures and assistances: descriptive statistics

Period	Nb. of Failures	Monthly Mean	Std. Dev.
1975 - 1984	438	3.65	4.12
1985 - 1994	2550	21.25	20.92
1995 - 2004	55	0.46	0.66
2005 - 2014	524	4.37	5.61
2015 - 2018	21	0.44	0.68
1975- 2018	3588	6.80	13.19

Source: FDIC

Table 1 provides additional information about the process of bank failures. From 1975 to 2018, the U.S. banking system experienced an average of almost seven bank failures each month, but that average masks significant fluctuations. Indeed the two distress episodes (the *Savings and Loans* and *subprime* crises) are clearly perceivable and the table shows that the period 1985-1994 saw an average of more than 21 banks failures each month whereas the corresponding figure for the period 2005-2013 was only about 5 failures a month. Signs of overdispersion, wherein the variance of count data is higher than its mean are visible, particularly during the distress episodes.

Figure 3: Histogram of the monthly number of U.S. bank failures and assistances



Note: Total monhtly number of bank failures and assistances in the U.S: the x-axis reports the number of bank failures and the y-axis the number of months in the 1975-2018 sample during which a corresponding number of bank failures occurred. Data are from FDIC.

Figure 3 depicts an alternative representation of the data, via a histogram showing the number of failures on the x-axis and the number of months during which the corresponding number of bank failures occurred. The figure shows that bank failures remain a relatively rare event: nearly 200 months in our sample experienced no bank failure. Conversely, the distress episodes imply that a relatively fat tail is present, with some months experiencing high numbers of failures: in March 1989, for instance, 175 banks went into bankruptcy. Our dependent variable is thus characterized by a large proportion of zeroes and overdispersion, features that our econometric strategy will take into account.

3.2 Explanatory Variables

McCracken and Ng (2016) propose a comprehensive, easily accessible and regularly updated database containing several dozen macroeconomic time series for the United States, organized by sectors. They aim to provide a convenient starting point for research on the forecasting ability of large datasets. Additionally, they provide real-time vintages of these data, allowing a better picture of the information available at specific moments of time. Our out-of-sample experiments below make use of the vintages when assessing the ability of our framework to forecast bank failures.

The database we use thus contains 131 different variables observed at a monthly frequency over the sample 1975M1 - 2018M12, where in accordance with McCracken and Ng (2016) all series have been transformed to induce a weakly stationary behaviour. Most $I(1)$ series are therefore used in first difference of logarithms. The monthly frequency of these data has the potential to help identify occurrence of banking difficulties in a timely manner.

The large number of variables in the database ensures one can take advantage of the largest amount of available relevant information. Considering the large number of variables however, a procedure by which the dimension of the estimation is reduced becomes necessary and our analysis via principal components is designed to achieve this, through the use of one predictor extraced from each sector of the larger database.

4 Econometric Framework

4.1 Predictors

A large literature has shown that the efficient use of the information contained in several dozen time series can make significant contributions to forecasting. In this context, extracting a few latent factors summarizing the common sources of variability in these large datasets, and then using these factors as predictors has been shown to provide significant forecasting ability in a wide variety of settings and is now a standard part of the forecaster's toolkit. These methods have, however, been used more sparingly in the banking literature and one importante innovation of the present paper is to use this approach to identify predictors for the number of bank failures.[6]

Specifically, we extract the principal components of each group (or sector) of variables in the McCracken and Ng (2016) dataset and then use the first principal component from each sector as the predictor summarizing its information contribution for forecasting bank

[6] Factor modeling for forecasting was popularized by Stock and Watson (2002a,b) and essential contributions in this field include Forni et al. (2005), Boivin and Ng (2006), Bai and Ng (2008) and Bai and Ng (2009). Stock and Watson (2006) reviews the literature.

failures.[7] Concretely, denote $\mathbf{X}^j$ as the data matrix for the N_j time series in sector j (one of the 8 present in our dataset); the principal component decomposition of $\mathbf{X}^j$ will uncover $\mathbf{F}_i^j, i = 1, \ldots, N_j$ components with each $\mathbf{F}_i^j$ a linear combination of the underlying data, such that

$$\mathbf{F}_i^j = \mathbf{X}^j \mathbf{c_i}, \tag{1}$$

with $\mathbf{c_i}$ the ith eigenvector associated to the variance-covariance or correlation matrix of $\mathbf{X}^j$. We then keep $\mathbf{F}_1^j$, the first principal component of sector i, as summarizing the information contained in that sector.

The strategy whereby we extract one predictor for each sector of the McCracken and Ng (2016) dataset is meant to provide some economic intuition for the likely source of a predictor's value-added to forecasting. In our analysis, this means that the grouping of variables under the sector "Housing" from McCracken and Ng (2016), consisting of various data on housing starts or building permits in the United States, is best synthetized as one variable related to the general health of the housing market in the United States.[8] An alternative strategy would be to search for the factor within that sector that has the most promising forecasting ability instead of selecting the first principal component (Bai and Ng, 2009). However, a comparison between these two strategies, conducted from the viewpoint of the within-sample exercises in Gnagne and Moran (2018), suggests that they would offer comparable performance.

4.2 Model

We now present the econometric strategy we follow to analyze the monthly occurrence of aggregate commercial bank failures in the United States. We first discuss the standard Poisson model often used as a starting point in the count data literature, before introducing refinements to this model aimed at accommodating data features such as overdispersion and excess zero counts.

[7] Principal components analysis is the tool most often used to extract factors from a large dataset (Stock and Watson, 2006). It is a multivariate statistical procedure that transforms a set of N correlated variables into a new set of N *principal components*, linear combinations of the original variables that are orthogonal to each other and form a basis on which the observations are projected. Principal components are ordered in that the first principal component explains the largest fraction of the overall covariance of the N original variables. For extensive discussion, see Jolliffe (1986) or Abdi and Williams (2010).

[8] The strategy ressembles that of Boivin and Giannoni (2006), where the authors consider that a variable in a highly-aggregated quantitative model, say consumption, is best represented as a factor drawn from several timeseries related to the concept of consumption.

4.2.1 Standard Poisson Model

The Poisson distribution generally represents the starting point in modeling count data. Its probability mass function (p.m.f) is given by:

$$f_{Y_t}(y_t) = \frac{e^{-\lambda_t}\lambda_t^{y_t}}{y_t!}, \tag{2}$$

where y_t represents the realization of a count variable of interest Y_t (the number of bank failures during month t in our case) and λ_t is the corresponding mean and variance, as both coincide in the standard model:

$$E[Y_t] = V[Y_t] = \lambda_t. \tag{3}$$

The standard Poisson regression model uses (3) to relate predictors to the conditional mean of y_t via the following:

$$E[Y_t|X_t] = \lambda_t = \exp(X_t'\beta), \tag{4}$$

with X_t the vector of predictors and β the vector of associated parameters.

This framework has been used to analyze the determinants of health services demand, insurance and accident claims and several other types of count data; see Cameron and Trivedi (2013) for a survey. It has, however, seldom been applied to the study of bank failures, with the notable exception of Davutyan (1989). Davutyan's analysis, however, studies the *annual* count of bank failures; by contrast, our analysis pertains to the monthly count of bank failures, an arguably more relevant objective for policy purposes. However, the standard Poisson regression model cannot be applied successfully to all count data analysis. Notably, features such as overdispersion (where the variance exceeds the mean) and excess zero-counts are at odds with the implications of the standard model. We now discuss extensions of the standard model designed to accommodate these features.

4.2.2 The Negative Binomial Model

Equidispersion, a feature of the standard Poisson model, refers to the equality of the mean and the variance of a count data variable of interest. By constrast, overdispersion (underdispersion) occurs when this property is violated and the variance exceeds (is less than) the mean. One class of count data model that can account for dispersion is the negative binomial (NB) model. Negative binomial models relax the strict assumption of equality of mean and variance and instead work with models admitting the following relationship between the conditional mean and the conditional variance of the variable of interest:

$$V[Y_t] = \lambda_t + \frac{\lambda_t^p}{\alpha}, \quad p \in \mathbb{R}, \tag{5}$$

where the two most common parameterizations specify $p = 1$ or $p = 2$. In the latter case, the expression thus becomes

$$V[Y_t] = \lambda_t + \frac{\lambda_t^2}{\alpha}, \tag{6}$$

and α is an overdispersion parameter to be estimated. This specification is the *NegBin2* model discussed in Cameron and Trivedi (2013) and the one we use below.[9]

4.2.3 The Hurdle Negative Binomial Model

Hurdle models were introduced by Mullahy (1986) and are designed to handle count data featuring excess zeros and overdispersion. These *two-part* models specify a process for the zero counts (the absence of bank failures in our case) that is different from the process for the positive counts (the number of bank failures when occurring). An economic interpretation of this structure could therefore be that two regimes can affect banking activities, namely *normal times*, for which $k = 0$, and *abnormal times* with increasing severity according to which $k = 1, 2, ...$

More specifically, let $f_1(0)$ denote the probability that y_t takes a zero value and $f_2(k)$, a truncated p.m.f. governing the intensity for values greater than zero ($k = 1, 2, ...$). Note that the two distributions functions underlying these probabilities are not constrained to be the same processes and/or to depend on the same predictors. The distribution of such a "hurdle-at-zero" model is given by:

$$f_{Y_t}(y_t = k) = \begin{cases} f_1(0), & k = 0, \\ \\ \dfrac{(1 - f_1(0))f_2(k)}{1 - f_2(0)}, & k = 1, 2, ... \end{cases} \tag{7}$$

where $f_1(\cdot)$ and $f_2(\cdot)$ then depend on the various predictors examined; $f_2(\cdot)$ is typically defined as a Poisson or negative binomial model, while $f_1(\cdot)$ can be a binomial or a

[9] Note that (6) is obtained by introducing an idiosyncratic, unobserved and multiplicative disturbance ϵ in the standard model, so that the p.d.f. now reads

$$f_{Y_t}(y_t) = \frac{e^{-\lambda_t \epsilon_t} (\lambda_t \epsilon_t)^{y_t}}{y_t!},$$

and then assuming a Gamma distribution for ϵ and solving for the unconditional first moments for y, which implies the relationship between $V[Y_t]$ and $E[Y_t]$ expressed in (6). See Cameron and Trivedi (2013) for details.

geometric model. The expected value arising from (7) is

$$E(Y_t) = \frac{(1 - f_1(0))}{1 - f_2(0)} \sum_{k=1}^{\infty} k f_2(k), \tag{8}$$

while the variance obeys

$$Var(Y_t) = \frac{(1 - f_1(0))}{1 - f_2(0)} \sum_{k=1}^{\infty} k^2 f_2(k) - \left[\frac{(1 - f_1(0))}{1 - f_2(0)} \sum_{k=1}^{\infty} k f_2(k) \right]^2. \tag{9}$$

Parameters of hurdle models are estimated with maximum likelihood and the log-likelihood function (L) of a hurdle-at-zero model is expressed as follows:

$$L = \sum_{t=1}^{T} \mathbb{I}_{\{y_t=0\}} \log f_1(0; \theta_{1,t}) + \mathbb{I}_{\{y_t>0\}} \log(1 - f_1(0; \theta_{1,t})) + \sum_{t=1}^{n} \mathbb{I}_{\{y_t>0\}} \log \frac{f_2(y_t; \theta_{2,t})}{1 - f_2(0; \theta_{2,t})} \tag{10}$$

with $\theta_{1,t} = \{X_t, \beta_1\}$, $\theta_{2,t} = \{X_t, \beta_2\}$, T the number of observations and β_1 and β_2 the parameter vectors associated to f_1 and f_2, respectively.

The specific assumptions this paper employs are as follows. Our benchmark model forecasts the future number of bank failures using the hurdle-with-negative-binomial (HNB) model (7), where the binomial function (f_1) governs the process generating the zeros and the negative binomial distributions explains the positive counts (f_2): the "hurdle-at-zero" feature is designed to capture the high occurrence of zeros noticed in Figure (3), while the negative binomial aspect seeks to address the high dispersion of positive counts. The vector of explanatory variables $\mathbf{X_t}$ consists of the principal components (or factors) extracted for the entire McCracken and Ng (2016) dataset.[10]

Recall that the distribution function of a binomial distribution is

$$f_1(s; n, p_s) = \frac{n!}{s!(n-s)!} p_s^r (1 - p_s)^{n-s}, \tag{11}$$

with n the number of trials, p_s the success probability for each trial and s the number of success. We posit a logit function for the binomial regression so that the probability p_s of success for each trial (the presence of non-zero bank failures for that month) is related to our predictors in the following manner:

$$log(\frac{p_s}{1 - p_s}) = \mathbf{X_t'}\beta. \tag{12}$$

[10] A related strategy to address high counts of zeros is known as the zero-inflated model (Cameron and Trivedi, 2013). It considers that zeros can arise either from the occurrence of Regime 1, which always results in a zero-count, or from Regime 2, a standard count model which includes the possibility of zeros. The in-sample analysis provided in Gnagne and Moran (2018) suggests that the forecasting potential of the two frameworks are similar.

5 Results

This section presents our main results. It first reports on in-sample experiments –estimations obtained using the full sample available– conducted to showcase the potential of our Hurdle-with-negative-binomial (HNB) model. These results also suggest that the factor associated with the housing-industry block of variables in McCracken and Ng (2016) is associated with that favorable performance. Second we document how the HNB model outperforms alternatives in out-of-sample forecsting exercises, using the real-time vintages of the data from McCracken and Ng (2016).

5.1 In-sample Estimation

Table 2 reports estimation results for the standard Poisson model and the HNB variant. These results are obtained by using the largest workable sample (1975M1-2018M12). For each model, the variable to be explained is the future number of bank failures (we report results for the one-month, three-month and twelve-month-aheads horizons) while the predictors consist of one principal component for each sector. Recall that two probability functions are specified in the HNB model, $f_1(\cdot)$ and $f_2(\cdot)$, where the first controls zero-counts –the absence of bank failures– while the other governs positive counts ie. the intensity of bank failures when present. As such, two sets of parameter estimates are reported.

The first standout result in Table 2 is the robust significance of the predictor associated with the housing sector. Interestingly, it is for the *extensive* margin of the HNB model –the presence or not of bank failures– that this variable is the most robustly significant. As such, the health of the housing industry, as represented by that variable, might be key to forecast the future presence of an episode of bank distress. Other sector-specific predictors do not exhibit an equivalent robustness: the predictor associated with the 'Orders' group of variables is notably significant for the Poisson process but not the HNB extensions. A few other sectoral variables appear statistically significant through the lens of the standard Poisson model, only to lose that significance when assessed using the HNB model.[11]

Second, Table 2 suggests that the HNB model has the better overall performance. All three classical measures of model performance reported, the log likelihood, the Akaike Information criterion (AIC) and the Bayesian Information criterion (BIC) clearly favour the HNB model. Our out-of-sample experiments below confirm this suggestion.

[11] Recall however that factors are identifiable only up to a square matrix and as such interpretation of their sign may be misleading.

Table 2: In-sample Estimation Results

Predicting at	$t+1$ month-ahead			$t+3$ months-ahead			$t+12$ months-ahead		
	Poisson	HNB		Poisson	HNB		Poisson	HNB	
		Zeros	NB2		Zeros	NB2		Zeros	NB2
Output/Income	-0.02	-0.25	0.06	-0.15^{***}	-0.08	-0.28	0.05	0.16	-0.06
Labor Market	-0.06	0.01	0.12	0.28^{***}	0.06	0.48^{*}	0.03	0.01	0.36
Housing Industry	0.19^{***}	0.84^{***}	-0.06	0.15^{***}	0.89^{***}	-0.16	0.04^{**}	0.88^{***}	-0.32^{**}
Orders	0.20^{***}	-0.06	0.12	-0.20^{***}	-0.19	-0.11	0.00	-0.03	-0.02
Money & Credit	0.28^{***}	0.14	0.30	0.00	0.02	0.13	-0.05	-0.01	-0.08
Int & Exch. Rate	0.07^{**}	0.20	0.01	0.01	0.23	-0.06	0.09^{***}	0.18	0.08
Prices	0.08^{***}	0.09	0.09	0.01	-0.05	0.01	0.01	-0.08	0.01
Stock Market	-0.12^{***}	-0.04	-0.07	0.04	0.04	-0.04	-0.04^{*}	0.09	-0.05
−Log Likelihood	3796	1314		3805	1308		3821	1297	
AIC	7610	2667		7629	2655		7660	2633	
BIC	7649	2748		7667	2736		7698	2713	

Symbols *,** and *** indicate statistical significance at 10%, 5% and 1% level.

Next, Table 3 repeats this analysis within a model where six lagged values of the dependent variable have been added as potential predictors. The facts documented above in Section 3, notably the tendency of bank failures to occur during a limited number of multi-months bank distress episodes, suggest that the explanatory power of the lagged number of bank failures ought to be assessed.[12] Indeed the table shows that these lagged values have statistically-significant explanatory power in many cases. However, this occurs mostly for the standard Poisson model, whereas the HNB model appears to have less need for these additional predictors. As was the case in Table 2 above, the housing industry is a key source of explanatory power for the model, notably for the extensive component (the presence or not of bank failures) of the HNB model. Further, all three criteria continue to strongly favour the HNB model.

[12] Gnagne and Moran (2018) present further analysis of the explanatory power arising from the lagged values of the dependent variable, in addition to several other explorations of the in-sample performance of our framework.

Table 3: In-sample Estimation Results: *Model with Lagged Bank Failures*

Predicting at	$t+1$ month-ahead			$t+3$ months-ahead			$t+12$ months-ahead		
	Poisson	HNB		Poisson	HNB		Poisson	HNB	
		Zeros	NB2		Zeros	NB2		Zeros	NB2
Explanatory Variables									
Output/Income	-0.03	-0.22	0.10	-0.17^{***}	0.05	-0.04	0.11^{***}	0.29	0.18
Labor Market	-0.19^{***}	0.35	-0.02	0.20^{***}	0.52	0.14	0.04	0.37	0.29
Housing Industry	0.19^{***}	0.36^{*}	0.08	0.13^{***}	0.54^{***}	0.04	-0.01	0.73^{***}	-0.14^{*}
Orders	0.12^{***}	-0.03	-0.08	-0.26^{***}	-0.44	-0.04	-0.09^{**}	-0.02	-0.15
Money & Credit	0.46^{***}	-0.02	0.14	0.06	-0.22	0.02	-0.02	-0.15	-0.09
Int & Exch. Rate	0.21^{***}	-0.20	0.21^{*}	0.07	-0.19	0.10	0.15^{***}	-0.20	0.11
Prices	0.13^{***}	0.11	0.04	0.03	-0.05	0.04	0.01	-0.08	0.04
Stock Market	-0.07^{***}	-0.06	0.02	0.02	0.03	-0.01	0.06^{**}	0.20	0.08
Lagged Response Variable									
Bank Failures (t)	0.02^{***}	0.23^{**}	0.04^{***}	0.01^{***}	0.19^{*}	0.02^{**}	0.01^{***}	0.34^{***}	0.04^{***}
Bank Failures ($t-1$)	0.01^{***}	0.12	0.00	0.01^{***}	0.23^{*}	0.01	0.01^{***}	0.10	0.01
Bank Failures ($t-2$)	0.01^{***}	0.12	0.01^{*}	0.01^{***}	0.12	0.01	0.01^{***}	-0.03	0.01
Bank Failures ($t-3$)	0.01^{***}	0.26^{**}	0.01	0.01^{***}	0.12	0.02	0.01^{***}	0.02	0.02
Bank Failures ($t-4$)	0.01^{***}	0.13	0.01	0.01^{***}	-0.10	0.02^{*}	0.00^{***}	-0.06	0.00
Bank Failures ($t-5$)	0.01^{***}	0.12	0.02^{**}	0.01^{***}	0.09	0.03^{***}	0.01^{***}	0.01	0.01
Bank Failures ($t-6$)	0.01	-0.04	0.01^{*}	0.01^{***}	0.19^{*}	0.02^{**}	0.01^{***}	0.04	0.01
−Log Likelihood	2039	1057		2194	1068		2436	1108	
AIC	4110	2180		4420	2202		4904	2282	
BIC	4178	2320		4488	2342		4971	2421	

Symbols *,** and *** indicate statistical significance at 10%, 5% and 1% level.

5.2 Out-of-Sample Experiments

5.2.1 The Experiment

The forecasting experiment we conduct is as follows. We consider the hurdle-with-negative-binomial (HNB) model with one pricnipal component/factor per group of variables (using the *real-time* version of the McCracken and Ng (2016) data for that date) as our benchmark; the set of explanatory variables $\mathbf{X_t}$ thus consists of 8 variables. We start with the sample $1975M1 - 1999M12$, estimate the model's coefficients and use them to provide forecasts for the future number of bank failures up to twelve months ahead, from $2000M1$ to $2000M12$. The procedure is repeated with an expanding window of data (one new observation each month) and recomputing the principal components at each step using the updated real-time McCracken and Ng (2016) data. At the end of the procedure, we thus have twelve time series (one- to twelve-months-ahead forecasts) covering the period $2000M1$ to $2018M12$. Table 4 summarizes the experiment's structure.

We then repeat this exercise firstly for the standard Poisson model and then versions of the HNB and standard Poisson model with lagged values of the dependent variable as additionnal predictors. We also assess different (competing) models, versions of the Poisson and the HNB model with two factors (principal component) per sector and, *squared factors* (one per sector) to accommodate possible non-linearities (Bai and Ng, 2008).

Table 4: The Forecasting Experiment (2000 - 2018)

Estimation Sample	Forecast h months ahead				
	$h=1$	$h=2$	$h=3$	$\cdots$	$h=12$
$1975M1 \longrightarrow 1999M12$	$2000M1$	$2000M2$	$2000M3$	$\cdots$	$2000M12$
$1975M1 \longrightarrow 2000M1$	$2000M2$	$2000M3$	$2000M4$	$\cdots$	$2001M1$
$1975M1 \longrightarrow 2000M2$	$2000M3$	$2000M4$	$2000M5$	$\cdots$	$2001M2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$1975M1 \longrightarrow 20017M11$	$2017M12$	$2018M1$	$2018M2$	$\cdots$	$2018M11$
$1975M1 \longrightarrow 20017M12$	$2018M1$	$2018M2$	$2018M3$	$\cdots$	$2018M12$

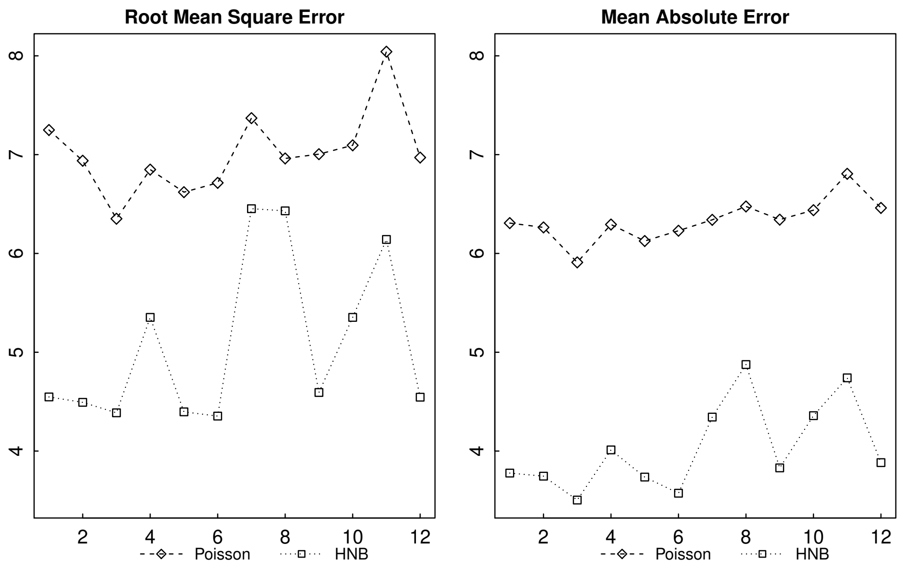
Note: Structure of the forecasting experiment: for each period, the underlying principal components predictors are computed using the real-time data available at the time according to the McCracken and Ng (2016) database.

5.2.2 Results

A first indicator of the superior forecasting ability of the HNB model is represented in Figure 4. The figure reports the root-mean-squared and mean-absolute errors across the 12 forecasting horizons considered in the exercise, for the standard Poisson model (diamonds) and our HNB framework (squares), in the case where no lagged values of the dependent variable are used. The figure's main takeaway is that the HNB model appears to be a better forecasting framework for all horizons considered, an impression validated by the formal tests conducted below.

Figures 5 and 6 depict these results through an alternative angle. In each of these two figures, the predicted and (eventually) realized number of bank failures are charted side-

Figure 4: Forecasting performance: 2000- 2018



Note: *RMSEs* (left-panel) and *MAEs* (right-panel) resulting from the 2000-2018 real-time forecasting experiment, standard Poisson and HNB models. X-axis represents the forecasting horizons and Y-axis the forecasting performance measure in units.

by-side for different forecasting horizons, with the standard Poisson on the left side of the graphs and the HNB framework on the right. The inability of the standard Poisson model to account for the frequent occurence of zero failures is readily apparent in the figures. In addition, the HNB model appears to better track the magnitude of the banking distress episodes of the late 2000s. Finally, although predicting the onset of a banking distress episodes is difficult for both frameworks, the HNB model appears to better identify their end.

We then repeat this analysis for the model versions where lagged values of the dependent variables are used as additionnal predictors; Figure 7, 8 and 9 report the results thus obtained. Overall, the quality of the forecasts appear to improve markedly with this addition: compare for example the scales of the Y-axis in Figure 7 relative to that in Figure 4. As such our formal tests comparing the forecasting performance of different models will be for the cases where the lagged values of bank failures are used (see below). Interestingly however, Figures 7, 8 and 9 also reveal that the relative comparisons between the standard Poisson and our HNB benchmark are largely unchanged from above: the HNB model has consistently-lower RMSEs and MAEs, has the better ability to correctly forecast the

Figure 5: Realized and Predicted Bank failures - 1 and 3-months-ahead horizons

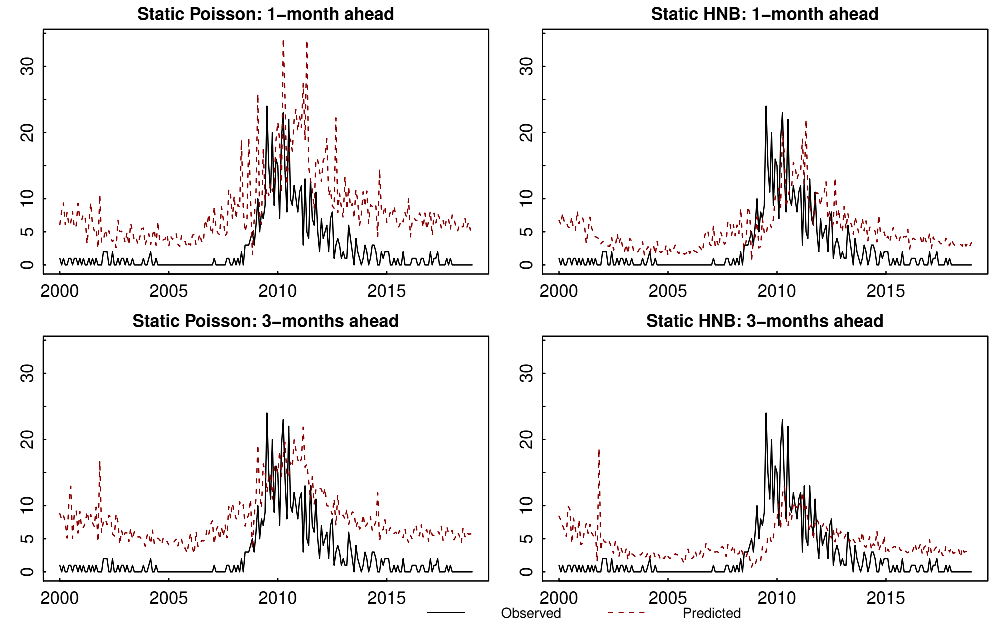


Figure 6: Realized and Predicted Bank failures - 6 and 12-months-ahead horizons

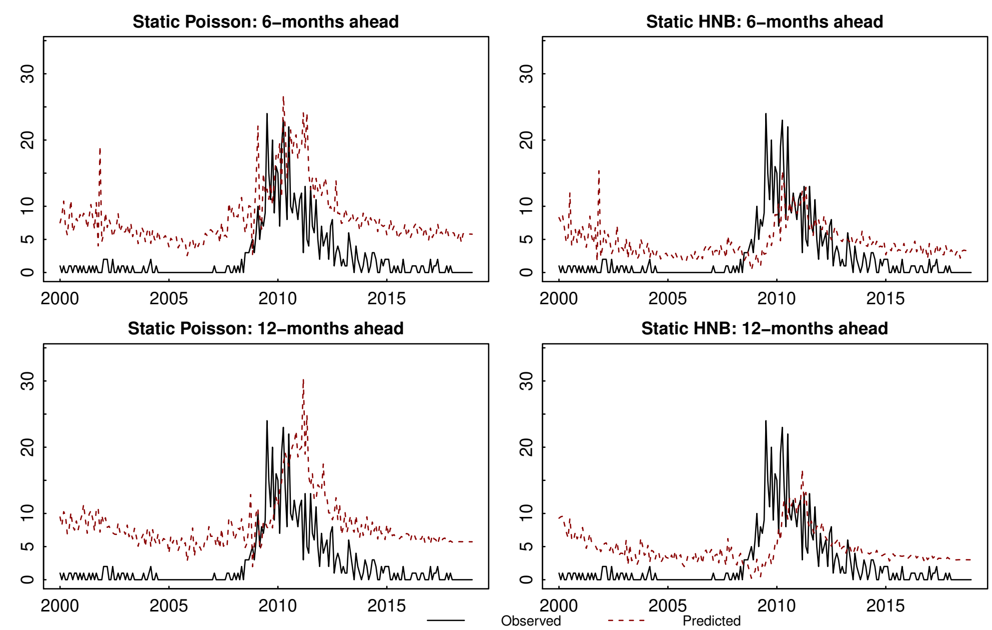
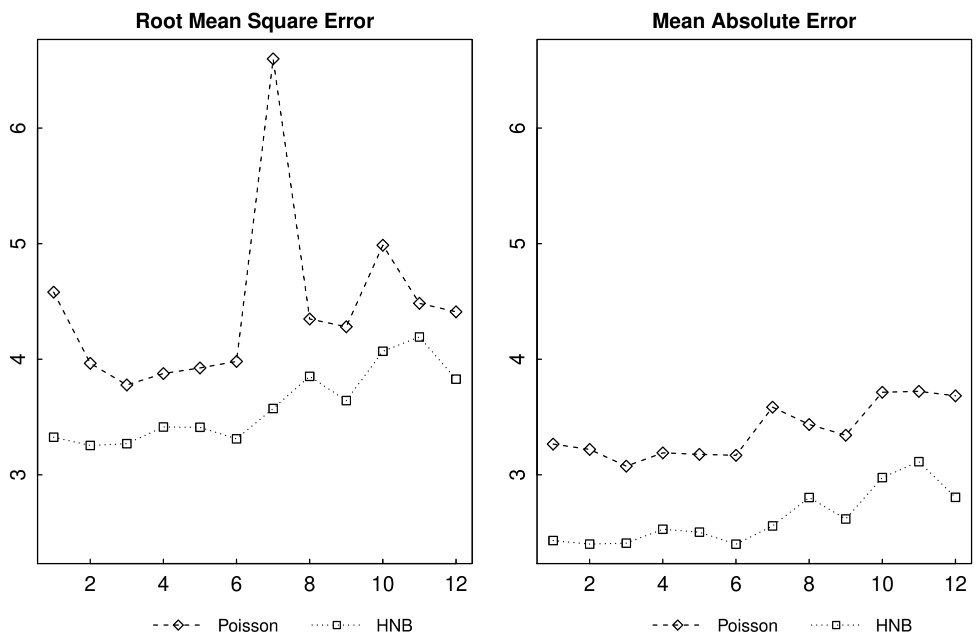


Figure 7: Forecasting performance with lagged values of bank failures



Note: *RMSEs* (left-panel) and *MAEs* (right-panel) resulting from the 2000-2018 real-time forecasting experiment, standard Poisson and HNB models with lagged values of bank failures as additional predictors. X-axis represents the forecasting horizons and Y-axis the forecasting performance measure in units.

no-banking-distress episodes and appears to better identify the end of such episodes when they occur. The fact that the HNB model retains its edge over the standard Poisson even when lagged values of bank failures are allowed attests to the robustness of the approach.

Finally, Table 5 documents the relative predictive accuracy of our benchmark, HNB-one-factor model against several alternatives. Each entry in the table reports the Root Mean Square Error (RMSE) for the alternative model and horizon considered *relative* to that of the benchmark.[13] For example, the first row and column of Table 5 reports that the RMSE of a one-factor standard Poisson model over the one-month-ahead horizon is 1.38 that of the RMSE of our benchmark over the same horizon. As such entries above 1 in the table suggest that our benchmark has the better forecasting ability for the future number of bank failures. In addition, symbols *,** and *** that these differences in predictive ability are statistically different at the 10%, 5% and 1% levels, respectively, according to

[13]As indicated above, all models considered for Table 5 are the versions with the lagged values of bank failures included as additional explanatory variables.

Figure 8: Forecasting performance with lagged values of bank failures:
1 and 3-months-ahead horizon

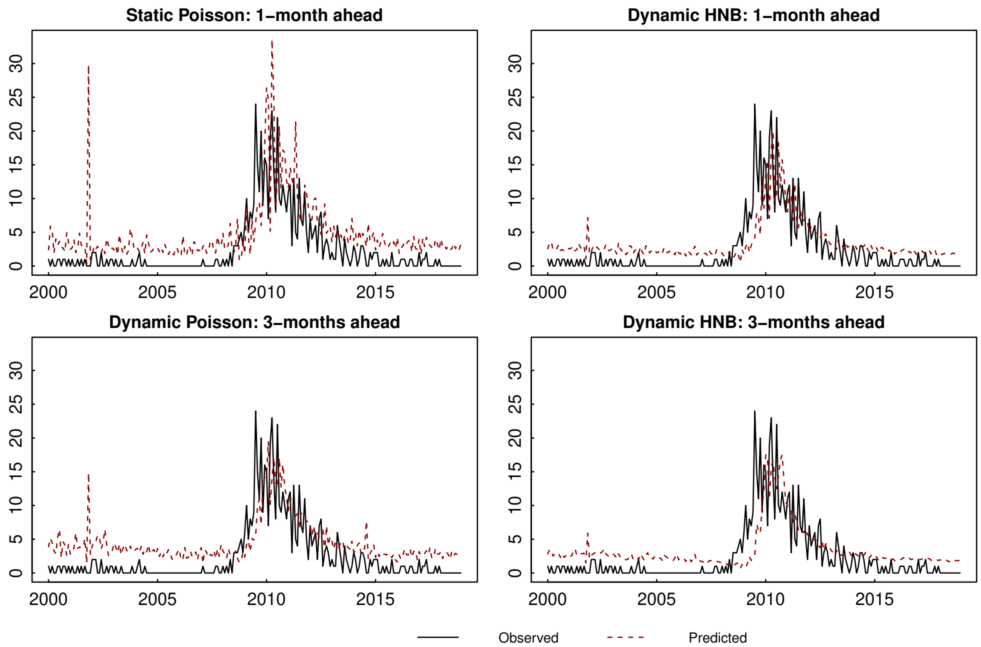
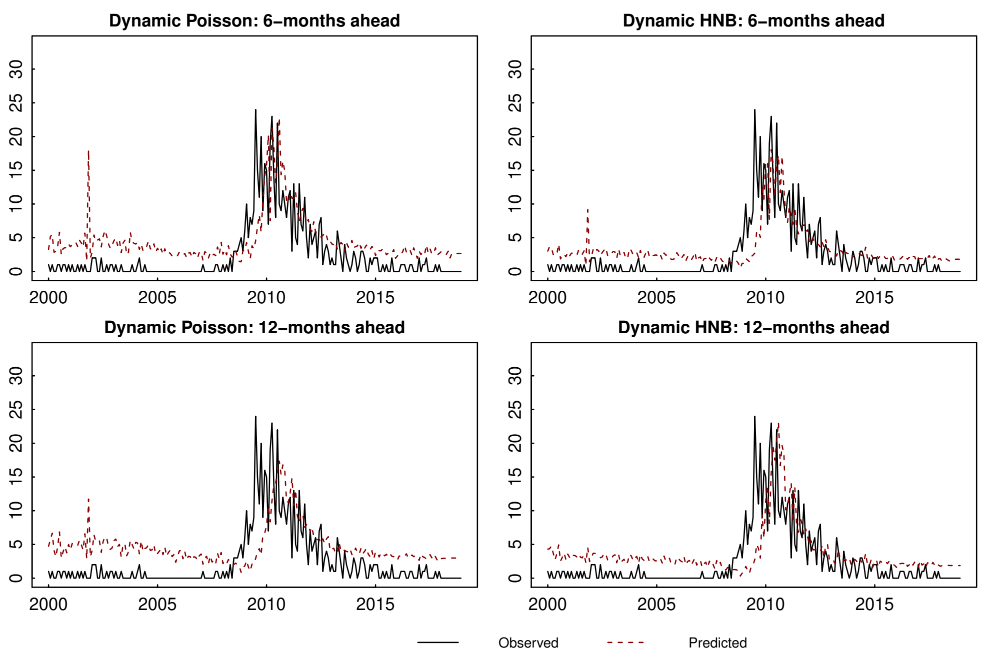


Figure 9: Forecasting performance with lagged values of bank failures:
6 and 12-months-ahead horizon



the Diebold and Mariano (1995) test.

Table 5: Testing for equal forecasting accuracy, 2000M1 - 2018M12

	Models				
	One Component	Two Components		Non-Linear Component	
Forecasting Horizon	Poisson	Poisson	HNB	Poisson	HNB
t+1 month	1.38***	1.66***	1.18	1.18**	1.09
t+2 months	1.22***	5.80	1.30	3.81	5.10
t+3 months	1.16	1.18**	1.10*	1.57**	1.30*
t+4 months	1.36***	2.85	1.50	1.32*	1.21
t+5 months	1.15***	6.00	0.99	1.05	1.18
t+6 months	1.20***	1.20*	1.08	1.09	2.54
t+7 months	1.85	5.45	1.16*	1.25**	1.05
t+8 months	1.13*	2.11	1.06*	1.17*	1.24
t+9 months	1.18***	1.70*	1.05	1.62	1.02
t+10 months	1.23	5.80	1.14*	1.11	1.39
t+11 months	1.07*	1.08	1.05	1.54	2.73
t+12 months	1.15**	1.06	1.76	1.13***	3.00

Note: Ratios of the Root Mean Square Errors (RMSE) of alternative models when forecasting the number of bank failures, to the RMSEs of our HNB benchmark, for the cases where lagged bank failures appear as additional predictors. Entries over 1 suggest superior forecasting performance of our benchmark and symbols *,** and *** indicate statistically significant differences, at 10%, 5% and 1% levels respectively (Diebold and Mariano, 1995).

Overall, the results reported in Table 5 clearly suggest that our HNB benchmark is the better forecasting framework. Notably, the first column of the table shows that it dominates the standard Poisson model for all forecasting horizons: all the relative RMSEs are substantially over 1, most of them in a statistically-significant manner. Interestingly, the two middle columns of the table also show that selecting only one predictor per group of variables (or sector), as our benchmark model does, also leads a stronger forecasting ability that allowing for two predictors per sector, although the latter strategy might fit the model better in-sample. As such, this represents another examples of how parsimony often produces the best out-of-sample forecasts. Finally, note from the last two columns of the table that using squared factors does not improve the ability of the framework significantly; overall then the key message from Table 5 is that our HNB benchmark with lagged values and one predictor per sector represents the best forecasting framework assessed.

6 Conclusion

This paper develops a monitoring and forecasting framework for the monthly aggregate occurrence of bank failures in the United States. To this end, we extract key sectoral predictors from the large set of variables in the McCracken and Ng (2016) database and incorporate them in a hurdle negative-binomial model for bank failures counts. Our result uncover a strong and consistent relationship between housing industry variables and banking failures and our out-of-sample forecasting exercise, using the real-time vintages of the McCracken and Ng (2016) data, documents the promising ability of the forecasting framework.

References

Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

A. Antunes, D. Bonfim, N. Monteiro, and P. MM Rodrigues. Forecasting banking crises with dynamic panel probit models. *International Journal of Forecasting*, 34(2):249–275, 2018.

Jushan Bai and Serena Ng. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317, 2008.

Jushan Bai and Serena Ng. Boosting diffusion indices. *Journal of Applied Econometrics*, 24(4):607–629, 2009.

R. Barrell, E. P. Davis, D. Karim, and I. Liadze. Bank regulation, property prices and early warning systems for banking crises in OECD countries. *Journal of Banking & Finance*, 34(9):2255–2264, 2010.

A. N. Berger and R. DeYoung. Problem loans and cost efficiency in commercial banks. *Journal of Banking & Finance*, 21(6):849–870, 1997.

Ben S Bernanke. *The Federal Reserve and the financial crisis*. Princeton University Press, 2013.

F. Betz, S. Oprică, T. Peltonen, and P. Sarlin. Predicting distress in European banks. *Journal of Banking & Finance*, 45:225–241, 2014.

J. Boivin and M. Giannoni. DSGE models in a data-rich environment. NBER Working Paper No. 12772, December 2006.

J. Boivin and S. Ng. Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):169–194, 2006.

C. Borio and P. Lowe. Assessing the risk of banking crises. BIS Quarterly Review, December 2002.

C. Borio and P. Lowe. Assessing the risk of banking crises–revisited. BIS Quarterly Review, March 2009.

John H Boyd and Mark Gertler. The role of large banks in the recent US banking crisis. *Federal Reserve Bank of Minneapolis. Quarterly Review-Federal Reserve Bank of Minneapolis*, 18(1):2, 1994.

Janice Boucher Breuer. Problem bank loans, conflicts of interest, and institutions. *Journal of Financial Stability*, 2(3):266–285, 2006.

M. Bussiere and M. Fratzscher. Towards a new early warning system of financial crises. *Journal of International Money and Finance*, 25:953–973, 2006.

A. C. Cameron and P. K. Trivedi. *Regression analysis of count data*. Econometric Society Monograph 53, Cambridge University Press, 2013.

E. P. Davis and D. Karim. Comparing early warning systems for banking crises. *Journal of Financial Stability*, 4:89–120, 2008.

Nurhan Davutyan. Bank failures as Poisson variates. *Economics Letters*, 29(4):333–338, 1989.

G. De Nicolo and M Lucceta. Forecasting tail risks. *Journal of Applied Econometrics*, 32 (1):159–170, 2016.

A. Demirgüç-Kunt and E. Detragiache. The determinants of banking crises in developing and developed countries. *IMF Staff Papers*, 45:81–109, 1998.

A. Demirgüç-Kunt and E. Detragiache. Cross-country empirical studies of systemic bank distress: A survey. IMF Working Paper, May 2005.

F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–263, 1995.

M. L. Duca and T. A. Peltonen. Assessing systemic risks and predicting systemic events. *Journal of Banking & Finance*, 37(7):2183–2195, 2013.

Huberto M Ennis and H Sam Malek. Bank risk of failure and the too-big-to-fail policy. *Federal Reserve Bank of Richmond Economic Quarterly Volume*, 91/2, 2005.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association*, 100:830–840, 2005.

M. Forni, A. Giovannelli, M. Lippi, and S. Soccorsi. Dynamic factor model with infinite-dimensional factor space: Forecasting. *Journal of Applied Econometrics*, 33(5):625–642, 2018.

Amit Ghosh. Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states. *Journal of Financial Stability*, 20:93–104, 2015.

J.-A. Gnagne and K. Moran. Monitoring bank failures in a data-rich environment. Working Paper No. 2018-15, CRREP, August 2018.

P. Gogas, T. Papadimitriou, and A. Agrapetidou. Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting*, 34(3): 440–455, 2018.

M. Hutchison and K. McDill. Are all banking crises alike? The Japanese experience in international comparison. *Journal of the Japanese and International Economies*, 13(3): 155–180, 1999.

I. T. Jolliffe. Principal component analysis and factor analysis. In *Principal component analysis*, pages 115–128. Springer, 1986.

G. Kaminsky and C. M. Reinhart. The twin crises: The causes of banking and balance-of-payments problems. *The American Economic Review*, 89:473–500, 1999.

William R Keeton. Does faster loan growth lead to higher loan losses? *Economic Review-Federal Reserve Bank Of Kansas City*, 84(2):57, 1999.

L. Laeven and F. Valencia. Systemic banking crises database. *IMF Economic Review*, 61 (2):225–270, June 2013. IMF Working Paper no.12/163.

D. P. Louzis, A. T. Vouldis, and V. L. Metaxas. Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios. *Journal of Banking & Finance*, 36(4):1012–1027, 2012.

M. W. McCracken and Serena Ng. Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589, 2016.

John Mullahy. Specification and testing of some modified count data models. *Journal of econometrics*, 33(3):341–365, 1986.

J. Podpiera and L. Weill. Bad luck or bad management? emerging banking market experience. *Journal of Financial Stability*, 4(2):135–148, 2008.

C. R. Reinhart and K. S. Rogoff. Banking crises: An equal opportunity menace. *Journal of Banking and Finance*, 37:4557–4573, 2013.

V. Salas and J. Saurina. Credit risk in two institutional regimes: Spanish commercial and savings banks. *Journal of Financial Services Research*, 22(3):203–224, 2002.

S. Smeekes and E. Wijler. Macroeconomic forecasting using penalized regression methods. *International Journal of Forecasting*, 34(3):408–430, 2018.

G. H. Stern and R. J. Feldman. *Too big to fail: The hazards of bank bailouts*. Brookings Institution Press, 2004.

J. H. Stock and M. W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002a.

J. H. Stock and M. W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162, 2002b.

James H Stock and Mark W Watson. Forecasting with many predictors. *Handbook of Economic Forecasting*, 1:515–554, 2006.