Federal Reserve Bank of Philadelphia

Working Papers
Research Department

WP 20-22
June 2020
https://doi.org/10.21799/frbp.wp.2020.22

Forecasting Consumption Spending Using Credit Bureau Data

Dean Croushore
University of Richmond and
Federal Reserve Bank of Philadelphia Consumer Finance Institute Visiting Scholar

Stephanie Wilshusen
Federal Reserve Bank of Philadelphia Consumer Finance Institute

Consumer Finance Institute

ISSN: 1962-5361

Forecasting Consumption Spending Using Credit Bureau Data

Dean Croushore and Stephanie M. Wilshusen

June 2020

Abstract

This paper considers whether the inclusion of information contained in consumer credit reports might improve the predictive accuracy of forecasting models for consumption spending. To investigate the usefulness of aggregate consumer credit information in forecasting consumption spending, this paper sets up a baseline forecasting model. Based on this model, a simulated real-time, out-of-sample exercise is conducted to forecast one-quarter ahead consumption spending. The exercise is run again after the addition of credit bureau variables to the model. Finally, a comparison is made to test whether the model using credit bureau data produces lower or higher root-mean-squared-forecast errors than the baseline model. Key features of the analysis include the use of real-time data, out-of-sample forecast tests, a strong parsimonious benchmark model, and data that span more than two business cycles. Our analysis reveals evidence that some credit bureau variables may be useful in improving forecasts of consumption spending in certain subperiods and for some categories of consumption spending, especially for services. Also, the use of credit bureau variables sometimes makes the forecasts significantly worse by adding noise into the forecasting models.

Keywords: consumer credit information, real-time data, forecasting, consumption spending

JEL Classification Codes: D12, D14, C53, C55, E27

The authors thank participants at the System Committee on Macroeconomics Conference, the Credit Scoring and Credit Control Conference, and seminar participants at the Federal Reserve Bank of Philadelphia. Please send comments to Stephanie Wilshusen at stephanie.wilshusen@phil.frb.org or to Dean Croushore at dcrousho@richmond.edu.

1. Introduction

This paper considers whether, as a source of real-time data, the inclusion of information contained in consumer credit reports might improve the out-of-sample predictive accuracy of rigorous macroeconomic forecasting models. Specifically, this paper will assess whether aggregate consumer credit information can be useful in forecasting consumption spending.

We emphasize that, when properly conducted, this is a very high hurdle to overcome. Assuming one has developed a strong forecasting model — built using real-time data — including additional variables typically worsens the out-of-sample precision of the model as measured by root-mean-squared-forecast error (RMSFE). We are not surprised to find that, in many instances, adding consumer credit variables does not improve the accuracy of consumption forecasts and, in a few instances, can make them worse. However, we do find instances in which the addition of a variable derived from credit bureau data can improve out-of-sample consumption forecasts.

We examine the existing literature in the next section, but it is important here to highlight the key methodological considerations that distinguish this paper from the related research and the forecasting literature more generally. First, as much as possible, forecasts and the evaluation of forecasts should be based on real-time data — the data that would have been available to the researcher at the time the forecast was developed. A great deal of macroeconomic data, especially national income accounting data, are subject to revision after the fact. A researcher developing or evaluating forecasts years later should not base the analysis on data that were not available at the time. Doing so confounds the information problem the forecaster actually faces. The analysis presented in this paper relies on real-time data.

Second, the gold standard for evaluating macroeconomic forecasts are out-of-sample forecast tests. Frequently, models are developed that perform well in terms of predicting outcomes within the sample, but they perform badly in forecasting outcomes out of sample. In other words, models can be developed that do a remarkable job of predicting past outcomes, but they do not do very well in predicting the future. The analysis presented in this paper examines the out-of-sample RMSFE of the models.

Third, with the preceding point in mind, the best models for predicting consumption out of sample are extremely parsimonious. We demonstrate this point in Section 4, but it is one that is widely accepted in the literature on the basis of both theory and empirics. It is fairly easy to show that adding more information to a weak model will improve its performance. It is difficult to find additional variables that will improve upon a well-conceived benchmark model in the sense of out-of-sample precision.

Last, this final point is related to the previous one. The best out-of-sample models for predicting consumption benefit more from a longer time series than they do from the addition of variables, especially when this reduces the time series over which the forecasting model is estimated. One advantage of the data used in this paper is that they span more than two business cycles in the U.S. economy.

It is therefore not surprising that many credit bureau variables that seem sensibly related to consumption do not improve well-designed out-of-sample consumption forecasts. Nevertheless, it is worth documenting this fact, given the conflicting results found in the literature. At the same time, we do find instances in which certain real-time credit bureau variables can moderately improve consumption forecasts, and this is also worth documenting as it provides a basis for future research.

We summarize some key findings here. For overall consumption spending over the entire sample period, the addition of credit bureau variables to the baseline model produces no statistically significant improvements. Similarly, for components of consumption spending over the entire sample period, we find that credit bureau variables are not helpful in forecasting spending on durables, nondurables, or services.

We also considered whether the Great Recession had a distortionary effect on outcomes and looked at sample periods that end prior to the recession. We found that credit bureau variables do have some statistically significant forecasting power for consumption of services.

As a sensitivity analysis, we varied the pre-forecasting and forecasting periods. A longer pre-forecasting period provides better coefficient estimates, but the forecast period is shorter and less likely to show significant effects. Conversely, a shorter pre-forecasting period provides worse coefficient estimates, but the longer forecast period might be more likely to show significant effects.

Using a shorter forecast period produced few significant results, except for one variable that made the forecasts of services consumption significantly worse. Using a longer forecasting period, pre-recession, generated the most significant results. For consumption of durables and nondurables, the use of credit bureau variables makes the forecasts significantly worse. However, three of the credit bureau variables produce a statistically significant forecast improvement for consumption of services.

Using a longer forecasting period including the recession yields similar results: mixed, not significant results for overall consumption, generally worse forecasts of durables and nondurables consumption, and generally improved forecasts of services consumption when including credit bureau variables, in one case a statistically significant improvement.

In a final exercise, we illustrate the effect of the choice of an appropriate baseline model on whether the inclusion of a credit bureau variable helps to improve forecasts. The results are sensitive to the number of lags of consumption spending used in the baseline model. Had we chosen a weaker baseline model, the credit bureau variables could have been shown to significantly improve the forecasts of consumption spending.

Of course, we are not the first to examine the relationship between consumer credit measures and macroeconomic conditions, specifically consumption. Haberler (1942) studied the impact of consumer installment credit on consumption and economic stability, while Hendricks and Youmans (1973) focused on spending on durable goods financed by consumer installment credit. More recently, Maki (2002) looked at whether credit market variables can improve forecasts of consumer spending and found that "credit quality" variables such as debt service burden, delinquency, and bankruptcy do not predict spending.

The rest of the paper is organized as follows. Section 2 provides an overview of recent literature on macroeconomic forecasting using financial transaction data. Section 3 describes the data used in the paper. Section 4 details the baseline forecasting model that uses real-time data, the out-of-sample evaluation of the baseline model, and the addition of certain credit bureau variables to the baseline model. In Section 5, we detail the results of the main estimation, for subcategories of consumption spending, for variations in forecast periods taking into account the Great Recession, and an illustration of the effects of choosing a weak versus strong baseline model. Section 6 concludes.

2. Related Literature

Studies from several countries have investigated the use of high-frequency electronic financial transaction data for macroeconomic forecasting and nowcasting. Galbraith and Tkacz (2007) investigate the usefulness of debit card transaction data to predict forecast errors and data revisions for Canadian GDP growth and personal consumption expenditures. While data limitations did not allow out-of-sample forecasting over the 2001 to 2005 forecast period, the authors report suggestive results that the inclusion of suitably transformed debit card transaction data improved initial estimates of GDP growth and personal expenditures, and predicted revisions of these national accounts data. In Galbraith and Tkacz (2015), the authors expand their electronic payments data information set to examine the contribution of debit, credit, and checking transactions to nowcasts of Canadian GDP growth. The value of the payments data — particularly debit card data — to reduce nowcast errors appears in a quarter's earliest nowcast because of its immediacy. The authors note that their results are suggestive rather than conclusive because of the limited observations afforded by quarterly GDP data over the 2000–2009 main sample.

Noting delays in the release of quarterly GDP data in Australia, Gill, Perera, and Sunner (2012) assess the usefulness of electronic purchase credit and debit transaction data as timely indicators of growth in economic activity. Their central question is whether these payment data can improve the fit of baseline autoregressive models of spending. While the transaction data are available for a short period of time (beginning in late 2002), Gill et al. find modest improvements in the fit of autoregressive models of spending, as well as in the models' out-of-sample predictive ability as measured by reductions in mean absolute error. However, the inclusion of the payments data did not reduce the forecast errors for any of the models' economic variables.

Ermişoğlu, Akçelik, and Oduncu (2013) use credit data to nowcast Turkish GDP growth and explain its variability. In this exercise, the authors use two credit flow measures (new borrowing and the change in new borrowing) for business loans and retail loans (the sum of consumer loans and credit cards). Using real-time data and an out-of-sample analysis, the authors find that — compared with a benchmark model based on the Turkish Purchasing Managers' Index — the models that include the credit flow measures show significantly reduced mean absolute errors and RMSFEs, which signifies improvement in their ability to nowcast quarterly GDP growth over the sample period 2006–2012.

Duarte, Rodrigues, and Rua (2017) use high-frequency electronic data from automated teller machine (ATM) and point-of-sale (POS) transactions from 2000 to 2014 to forecast private consumption. The authors use mixed data sampling (MIDAS) regressions to assess the ability of these high-frequency data to monitor quarterly macroeconomic activity in Portugal. The authors use the ATM/POS data, the retail sales index, and a consumer confidence indicator to nowcast and forecast one-quarter ahead nominal quarterly private consumption (excluding durables). They assess the performance of their forecasting models using an out-of-sample forecast evaluation that covers half of the sample period (the beginning of 2008 through the end of 2014). Based on an assessment of the relative root-mean-squared-forecast error for each variant of the MIDAS model, and for each predictor compared with the univariate benchmark model, Duarte et al. conclude that the use of ATM/POS data statistically significantly improves forecasting performance.

Gil, Pérez, Sánchez, and Urtasun (2018) nowcast and forecast quarterly Spanish private consumption using real-time monthly indicators, including ATM withdrawals and payment card transactions at the POS. They estimate mixed-frequency models over the period 2001–2016 including out-of-sample forecasting exercises and find that adding aggregate POS and ATM transaction value data leads to significant nowcast and forecast improvement.

Verbaan, Bolt, and van der Cruijsen (2017) examine the value of monthly electronic POS debit card payments data in lowering errors in nowcast and one-quarter ahead forecasts of Dutch private consumption. Their analysis does not use real-time data over the two sample periods 2008–2015 and 2012–2015 and one period out-of-sample forecasts. Including debit card payments data in a macroeconomic policy model reduced the RMSFE of household consumption nowcasts.

3. Credit Bureau Data

In this paper, we examine whether using aggregated consumer credit variables that originate in big data help us to forecast consumption spending. The information contained in consumer credit reports is an example of big financial data. The data contained in consumer credit reports are provided to credit reporting agencies by approximately 10,000 to 15,000 data furnishers, including creditors, government agencies, and collection agencies.[1] Consumer credit report information is maintained by each of the three leading national consumer credit reporting agencies (CRAs): Equifax, Experian, and TransUnion. Our data set comes from TransUnion. Examples of the information contained in credit reports include applications for credit, new credit obtained by type of credit, balances on many different categories of credit, and utilization rates, which are the ratio of balances to credit lines on revolving credit, delinquencies, and bankruptcies.

The consumer credit report data used in this paper are from TransUnion and are based on the data archived at the time; they have not been revised. Some other research data sets based on credit report data also have this property, but it is important to emphasize that the credit report data provided to lenders for making credit decisions are subject to revision. Under the Fair Credit Reporting Act, a consumer can dispute inaccurate information in his or her credit report and erroneous information will be eliminated from the history available to lenders.

The data are frequently updated, typically on a monthly basis; in this paper, we use data with a quarterly frequency that quickly reflects trends and changes in credit use, which enables researchers to examine the dynamic nature of credit use by consumers. They can observe consumers' decisions about debt payments (payment hierarchy) and to the extent those decisions are based on expectations of future macroeconomic conditions, such repayment behavior may be useful in improving the accuracy of forecasts. In addition, changes in levels and composition of consumer debt and applications for credit and new credit granted potentially contain information about consumer and lender expectations about future economic conditions.

The data set in this paper was developed to provide quick insights about how consumers are acquiring and using consumer credit. It consists of representative snapshots of consumer credit activity that were archived at the time the data were reported. The data in this paper are aggregates and were available from the first quarter of 1992 to the third quarter of 2015 for most

[1] See Federal Trade Commission (2019), Accuracy in Consumer Reporting: An FTC/CFPB Workshop, transcript, *Accuracy in Consumer Reporting, Part 2-1*, p. 44.

credit characteristics, spanning 2.5 business cycles.[2] The types of variables we test include applications for credit (inquiries), balances, utilization of credit lines, delinquencies, and changes in the distribution of a credit score.

Figure 1 plots a few credit bureau variables (overall delinquencies and mortgage delinquencies) as well as real consumption growth over the course of the last few business cycles.[3] An obvious inverse relationship between delinquencies and consumption is apparent. This is one of many examples we could show.[4] The following sections explore whether incorporating consumer credit variables such as delinquencies improves the accuracy of out-of-sample consumption forecasts.

4. The Forecasting Model

Our goal is to investigate whether credit bureau variables are helpful in forecasting consumption spending. To do so, we first develop a baseline forecasting model for consumption spending based on the existing literature. Then, we add a variety of different credit bureau variables to the model to see if the resulting forecast errors are less than those of the baseline model.

The literature on forecasting consumption spending is vast, but few of the past studies performs simulated real-time out-of-sample tests. Instead, most studies just run regressions based on final, revised data, and suggest that a significant coefficient in that in-sample test is sufficient to prove a variable's forecasting ability. Mostly, this has been due to a lack of real-time data and thus the inability of researchers to run a more desirable simulated real-time out-of-sample test. However, the development of the Real-Time Data Set for Macroeconomists made the latter much easier to perform. For example, Croushore (2005) performs such an exercise for testing whether consumer confidence variables are helpful in forecasting consumer spending. Croushore's model is based on Bram-Ludvigson (1998), which set up a useful forecasting structure but without real-time data.

---

[2] Wiermanski and Wilshusen (2015) describe the credit bureau data and discuss its properties.

[3] *Delinquencies* are defined here and throughout the paper as loans that are 60 days or more past due.

[4] Additional examples are found in Wiermanski and Wilshusen (2015).

4.1 Baseline Model

The baseline forecasting model in Bram-Ludvigson (1998) and Croushore (2005) is a regression equation in which consumption spending growth in quarter *t* is a function of four quarterly lags of consumption growth, income growth, growth in real stock prices, and the change in the short-term interest rate. Although all four variables are significant in-sample using final, revised data, Croushore (2005) found that the forecasts were improved in a real-time out-of-sample simulation by dropping the income variable and the real interest rate, and just including four quarterly lags of consumption growth and real stock prices.

To represent consumer spending, we use real personal consumption expenditures, which is a variable reported by the Bureau of Economic Analysis. We also wish to simulate how a forecaster could use the data in real time, which is complicated by the fact that data are revised. So, we will use the Real-Time Data Set for Macroeconomists produced by the Federal Reserve Bank of Philadelphia.[5]

The quarterly data on consumer spending are released initially one month after the end of the quarter based on an incomplete sample. The data are revised in each of the following two months as additional sample data become available. Then the data are revised in July (usually) in each of the following three years, along with the annual revision to all National Income and Product Account data. Finally, about every five years or so, there are benchmark revisions, which incorporate definitional changes to the accounts.

To provide a sense of the magnitude of the revisions, Figure 2 shows two versions of the consumer spending data. The data plotted are annualized growth rates from one quarter earlier. The figure shows the initial release of the data and the data as they occur in today's data set, which have been revised many times. It further reveals that the data are revised substantially between their initial release and the final, revised data available in today's data set. Revisions are large, averaging 1.0 percentage point in absolute value, compared with an average growth rate of 2.9 percent. So, the average revision size is over one-third of the magnitude of spending growth.

To represent income, we use two different variables: personal income and disposable (after-tax) personal income. Both variables are in nominal terms, so we adjust both for inflation

---

[5] See Croushore and Stark (2001) for details on the data set.

using the price index for personal consumption expenditures (referred to as the PCE price index). For stock prices, we use the Wilshire 5000 stock price index, adjusting for inflation using the PCE price index. For interest rates, we use the interest rate on three-month Treasury bills.

The Wilshire 5000 stock price index and the interest rate on three-month Treasury bills are not revised, so we take their values from the Federal Reserve Bank of St. Louis FRED database.[6] The other variables are all subject to revisions, so we must use real-time data for them. From the Real-Time Data Set for Macroeconomists, we obtain vintages of data for each variable from 1965Q4 to 2016Q4. Each vintage shows the complete time-series history of that variable as it would have appeared to a forecaster in the middle of the quarterly vintage. For example, the 1965Q4 vintage shows what a forecaster would have observed in a macroeconomic database on November 15, 1965. In running our simulated forecasting exercise, we are careful to never look ahead at future data but implicitly assume that the forecaster just used the data available at the time. Croushore (2011) discusses forecasting with real-time data and the procedures necessary to perform true out-of-sample forecasting experiments.

A first baseline model is thus:

$$\Delta c_t = \alpha_0 + \sum_{i=1}^{4} \alpha_1^i \Delta c_{t-i} + \sum_{i=1}^{4} \alpha_2^i \Delta y_{t-i} + \sum_{i=1}^{4} \alpha_3^i \Delta r_{t-i} + \sum_{i=1}^{4} \alpha_4^i \Delta s_{t-i} + \varepsilon_t, \quad (1)$$

where Δc_t is the quarterly growth rate of real consumption spending, Δy_t is the quarterly growth rate of real income (either real disposable personal income or real personal income), Δr_t is the quarterly change in the nominal interest rate on 3-month Treasury bills, and Δs_t is the quarterly growth rate of real stock prices.

4.2 Out-of-Sample Evaluation of the Baseline Model

To find the best benchmark model, we perform a simulated out-of-sample forecasting exercise, following these steps. First, consider a forecaster's analysis at a particular date, for example, 2002Q1. Using the vintage data from 2002Q1, which would include data on all the variables from 1992Q1 to 2001Q4, run the regression in equation (1). Use the estimated coefficients to make a forecast of the quarterly growth rate of consumption spending for 2002Q1. Then, roll forward one period to consider a forecaster standing in 2002Q2. Using the vintage data from

[6] The Federal Reserve Bank of St. Louis FRED database is available at *fred.stlouisfed.org*.

2002Q2, which would include data on all the variables from 1992Q1 to 2002Q1, run the regression in equation (1). Use the estimated coefficients to make a forecast of the quarterly growth rate of consumption spending for 2002Q2. Continue in this fashion until forecasts have been created for all dates between 2002Q1 and 2015Q3. Now we have an entire series of forecasts that we can use to compare with other forecasts.

To see which baseline model is best, consider dropping one or all of the variables in equation (1) and repeat the exercise. Do this for all combinations of the variables, thus generating numerous sets of forecasts, as in Croushore (2005). Also, once the ideal subset of variables has been determined, we can test for the optimal number of lags of each variable, which might differ from the four lags illustrated in equation (1).

Because the data on consumer spending are revised, we must make a decision about how to calculate the forecast errors; in particular, how do we determine which release of the data to consider the true value? This is a crucial issue, given redefinitions of variables in benchmark National Income and Product Account revisions. No method of forecasting can anticipate redefinitions of variables, so the best forecasting model may not be the one that turns out to be best at forecasting the final, revised value of the variable. Instead, we will consider several alternatives: the initial release of the variable, the first annual release of the variable, and the pre-benchmark release of the variable (that is, the last release of a variable right before it gets redefined). Croushore (2011) discusses these alternatives in greater detail. For comparison purposes, we will also show the forecast errors based on the final, revised data, although we think these are of less interest than the other measures.

The statistic used to determine which model is best is the Root-Mean-Squared-Forecast Error (RMSFE).[7] We analyze many variations of the baseline model, showing the RMSFE for each model using data beginning in 1992Q1, using the first 10 years of data to estimate the model, then running rolling forecasts for the forecast period from 2002Q1 to 2015Q3. The model with the lowest RMSFE is the best model. No matter which variable is used as "actual" (Initial, Annual, Pre-benchmark, or Final), the lowest RMSFE occurs when just three lags on consumption spending growth are used in the model. While it may seem surprising that using

[7] The use of RMSFE implicitly means that the loss function for evaluating the forecasts is quadratic.

any of the other variables makes forecasts worse, the result is actually consistent with economic theory.[8] In addition, forecasting theory suggests that many variables that seem significant in-sample just add noise in a forecasting context and should be removed from a forecasting model.[9] So, our baseline forecasting model is now:

$$\Delta c_t = \alpha_0 + \sum_{i=1}^{3} \alpha_1^i \Delta c_{t-i}. \tag{2}$$

4.3 Augmenting the Baseline Model Using Credit Bureau Variables

Next, we must choose some reasonable credit bureau variables that might be helpful in forecasting consumer spending and add them to the model. The model that includes a credit bureau variable is:

$$\Delta c_t = \alpha_0 + \sum_{i=1}^{3} \alpha_1^i \Delta c_{t-i} + \sum_{i=1}^{n} \beta^i \, b_{t-i}, \tag{3}$$

where *b* represents a credit bureau variable of some type, and *n* is the number of lags (to be determined). Experimentation finds the best results when $n = 1$ sometimes with the credit bureau variable in levels, $n = 1$ other times when the credit bureau variable should be in changes, and $n = 2$ with the level and change in the credit bureau variable both included.

Which credit bureau variables might be most helpful in forecasting consumer spending? The variables that seem most likely, a priori, to help forecast consumption growth are variables that measure inquiries, delinquencies, balances, utilization of credit, and credit score.[10] To experiment, we will try one or two variables of each of those types and see if any or all of them are helpful in forecasting consumer spending.

For inquiries, we chose the percent of consumers with inquiries in the past 90 days (variable name INQUIRE). For delinquencies, we chose as variables the percent of borrowers

[8] See Hall (1978), who notes that "no variable apart from current consumption should be of any value in predicting future consumption" (p. 971). Croushore (2005) found that lagged real stock prices were also somewhat helpful in forecasting consumer spending, but his sample period was very different (1968 to 2002).

[9] See Diebold (2007).

[10] Wiermanski and Wilshusen (2015) discuss simple lead-lag relationships between credit bureau variables and macroeconomic variables. For example, using data for 2000–2011, TransUnion found that lags of captive auto finance limits Granger cause changes in retail sales. Lags of a number of credit bureau variables (e.g., applications for new auto and mortgage loans, credit limits on bankcards, and balances on revolving trades) were positively correlated with retail sales or measures of housing activity. In addition, we did some experimentation with other variables including these indices for specific types of credit but found no variables with strong forecasting power.

who are delinquent 60 days or more (DELINQ), and the percent of mortgage borrowers who are delinquent 60 days or more (MDELINQ). For balances, we used total debt per borrower (AVGDEBT). For utilization, we used the average utilization of all accounts (UTILIZE). For credit score, we used a variable derived from an account management model developed by TransUnion that measured the probability of an account going 90 or more days past due within the 24 months following the score calculation (SCORE).[11] The financial effects of recessions, and especially in the Great Recession, are both direct and indirect. The direct effect is more keenly felt by consumers with credit scores on the lower end of the credit score distribution; that is, those people who are facing constraints and who have less latitude to make adjustments in their spending and repayment behavior. The indirect effect falls on consumers who are able to adjust their borrowing and repayment behaviors in response to changes in expectations that occur during recessions — these are consumers with relatively high credit scores. This motivates us to examine changes in the tails of the credit score distribution. In this paper, we focus on the thickness of the right (upper tail) of the distribution as compared with a reference period.[12]

5. Results

In the tables that follow, we report the relative root-mean-squared-forecast error (RRMSFE), which is the ratio of the root-mean-squared-forecast error when we include a credit bureau variable in the forecasting equation divided by the RMSFE of the baseline model. That is:

$$RRMSFE = \frac{RMSFE(with\ credit\ bureau\ data)}{RMSFE(baseline)}, \tag{4}$$

where *RRMSFE* > 1 means the inclusion of the credit bureau variable makes the forecasts worse, and *RRMSFE* < 1 means that the inclusion of the credit bureau variable helps improve the forecasts. The tables also report the *p*-value of the test statistic of the null hypothesis of no

[11] Specifically, the variable is the proportion of scores that fall into the 90–95 percentiles of the probability score, where the cutoff values are defined for a reference period. This measure compares the thickness of the tails of the credit risk distribution relative to a point in time.

[12] After some experimentation, we focus on the proportion of scores that fall into the 90–95 percentiles of the probability score, where the cutoff values are defined for a reference period. We chose this for several reasons. First, the right tail is considerably fatter than the left tail — there are more consumers there. Second, it represents the set of consumers with the most access to credit, which means they account for a very large share of consumer borrowing and repayment behavior that could affect aggregate consumption.

significant difference between the forecasts; so, a p-value < 0.05 means that the model with the lower RMSFE is statistically significantly better than the model with the higher RMSFE.[13]

5.1 Exercise 1 — Main Results

Our first exercise uses data on all the variables from 1992Q1 to 2001Q4, then starts the forecast period at 2002Q1, rolling through one quarter at a time, finishing in 2015Q3.[14] We then analyze the statistical significance of the difference in RMSFEs. Table 1 shows the results of this exercise.

In Table 1, about half of the time, the use of a credit bureau variable improves the forecasts (RRMSFE < 1), and the other half of the time, the use of a credit bureau variable makes the forecasts worse. However, with p-values well above 0.10 in all cases, none of the RRMSFEs is statistically significantly different from 1.

5.2 Exercise 2 — Predicting Subcategories of Consumption Growth

With no significant impact of credit bureau variables on overall consumption spending, we next ask if subcategories of consumption spending might be forecastable with credit bureau data. A priori, we might think, for example, that people use credit to buy durable goods, and thus credit bureau variables might do a good job at forecasting consumption spending on durable goods.

To investigate this possibility, we run a second exercise, breaking down consumption spending into its three components: durables, nondurables, and services. Just as we did with overall consumption spending, we need to establish a baseline model. Testing finds that for consumption on durables and services, the optimal baseline model is one with three lags of overall consumption spending; for consumption of nondurables, the optimal baseline model is one with five lags of overall consumption spending. Next, we run the models with and without the credit bureau variables and report the RRMSFEs in Table 2.

---

[13] The test is that of Harvey, Leybourne, and Newbold (1997).

[14] The exceptions are for the inquiries measure and the credit score measure. Data on inquiries begin in 2000, and data on credit score begin in 1998; so, for both of these variables, we do not start the forecast period until 2007Q1, to allow a sufficient amount of data to estimate the model.

In Table 2, we see results similar to those in Table 1, but now more than half of the RRMSFEs are larger than 1, meaning that the forecasts become worse when using a credit bureau variable. In just two cases in Table 2, there is a *p*-value less than 0.10, so most of the RRMSFEs are not statistically significantly different from 1. Nothing in Table 2 suggests that the credit bureau variables are helpful in forecasting the subcategories of consumption spending.

In the forecast sample, the Great Recession of 2007–2009 might play a significant role. Because of the sharp changes in consumer spending and in credit measures in that period and its aftermath, it might be that all of our analysis has been overly affected by that unique event. So, we can try to look at samples that end prior to the recession.

5.3 Exercise 3 — Excluding the Great Recession

If we end the sample in 2007Q4 before financial conditions deteriorated in the economy, we might see more evidence that credit bureau variables could help forecast consumption spending. Table 3 shows the results of carrying out the exercise. Note that we cannot use the variables INQUIRE and SCORE because of insufficient data.

The results in Table 3 show that credit bureau variables have some forecasting power once we eliminate the impact of the Great Recession. Most notably, credit bureau variables help significantly in forecasting consumption of services, with as much as a 13.2 percent improvement in the RMSFE (for using the level and change of AVGDEBT).

An element that hampers our ability to test the forecasting ability of the credit bureau variables is the fairly short sample period, relative to most macroeconomic variables, for which we have data.[15] To get around that, we can vary the pre-forecasting period and forecasting periods.

5.4 Exercise 4 — Using a Shorter Forecast Period

In forecast evaluation, we face a tradeoff between the length of the pre-forecasting period and the length of the (out-of-sample) forecasting period. If we shorten the forecasting period, then we

[15] Data that span more than two business cycles are considered a relatively long time series in terms of consumer credit information but not in the macroeconomic forecasting literature.

have a longer sample on which to base our forecasts, so we might get more accurate coefficient estimates and thus better forecasts. But because the forecasting evaluation period will be shorter, we are less likely to find statistically significant differences in the forecasts. We run the same type of exercise as in Table 3, but we use data from 1992Q1 to 2006Q4, then start the forecast period at 2007Q1, rolling through one quarter at a time, finishing in 2015Q3. The results are shown in Table 4. As we expected, there are few significant results, except for the utilization variable, making the forecasts of services consumption significantly worse in the statistical sense.

5.5 Exercise 5 — Using a Longer Forecast Period, Pre-Recession

Our next possibility is that we can shorten the sample period and start forecasting earlier. So, rather than using data from 1992Q1 to 2001Q4, then starting to forecast, we instead just use data from 1992Q1 to 1996Q4, starting the forecasting period at 1997Q1, rolling through one quarter at a time, finishing in 2007Q4 to avoid the recession effects. The disadvantage of the shorter estimation period is that the shorter sample might yield less precise coefficient estimates, so the forecasts may not be as good. Table 5 shows the results.

The results in Table 5 are interesting; they show the most significant results of any of our exercises. For overall consumption, none of the results are statistically significant. But for the consumption of durables and nondurables, in many cases, the use of credit bureau variables makes the forecasts significantly worse. However, for the consumption of services, three of the credit bureau variables lead to a statistically significant forecast improvement.

5.6 Exercise 6 — Using a Longer Forecast Period, but Include the Great Recession

Our last possibility is to keep the same shorter sample period, start forecasting earlier, and forecast over the entire period from 1997Q1 to 2015Q3. The results are shown in Table 6.

Table 6 shows results similar to what we have seen before in other cases. For overall consumption, the results are mixed with nothing significant. For consumption of durables and nondurables, generally the forecasts are worse when using credit bureau variables and, in one case, statistically significantly worse. For services consumption, generally credit bureau variables help, and in one case, statistically significantly so.

5.7 Exercise 7 — Using a Weaker Baseline Model

In the forecasting literature, many of the results for whether a variable helps improve forecasts depend on the choice of an appropriate baseline model. In a sense, our procedures have made it more difficult for credit bureau variables to matter because we chose a really good set of baseline models for the different consumption variables. Had we chosen a worse baseline model, it is more likely that the credit bureau variables would matter.

To illustrate this point, consider the SCORE variable. In Tables 1 and 2, the use of the SCORE variable to forecast made the forecasts worse in all but one case (for services, using the change in SCORE), and in one case, it was statistically significantly worse. But how does that result vary with the baseline model? We examine three different baseline models: one with one lag of overall consumption spending, another using three lags of overall consumption spending (as shown in Table 1), and a third variation that uses five lags of overall consumption spending. Then we run the exercise shown in Table 1, varying only the baseline. The results are shown in Table 7.

Table 7 shows that the results of the exercise are sensitive to the number of lags of consumption spending used in the baseline model. If we had used a baseline model with just one lag of overall consumption spending, the SCORE variable significantly improves the forecasts using the change in SCORE, and the RRMSFE is less than 1 for the level of SCORE or including the level and change. However, as we change the number of lags of overall consumption spending in the equation, the SCORE variable becomes less useful and makes the forecasts worse, though not significantly so.

These results highlight an important concept in the forecasting literature — the apparent strength of a forecast depends on what it is compared with. If we had picked a worse baseline model, the credit bureau variables might have all shown up as significant. But with the best baseline model possible (among the baseline models we tested), the credit bureau variables are seldom successful at significantly improving the forecasts compared with the baseline model.

6. Interpretation and Conclusions

Our analysis reveals some evidence that credit bureau data could potentially help improve forecasts of consumer spending growth. However, most of the improvements are not statistically significant. And in a few cases, the use of credit bureau data makes the forecasts worse, sometimes significantly so.

This stands in contrast to prior work in other countries, which found suggestive results that the use of credit or payments data helped improve forecasts of macroeconomic variables. However, few of those studies used real-time data, which is required in this type of analysis, and few evaluated forecast improvement in terms of out-of-sample precision, often because of data limitations. In some instances, the baseline model used for comparison could have been stronger. In others, the focus was on nowcasting, which presumes even less macroeconomic data available to the forecaster; therefore, the value of alternative data should be greater.

For reasons of space, we did not describe all of the exercises we ran in this paper. We began with a number of different baseline models before we settled on a method to determine the ideal baseline model. We checked a number of other credit bureau variables (including by type of credit), but most of the results were similar to those of the variables we show in the paper. We used a number of different measures of "actual" in an earlier version of this paper, including the initial release, the first annual release, and the final revised data. There were no systematic differences between those results and the results we present in this paper.

We initially did all of our analysis with four lags of the consumption variable and four lags of the credit bureau variable, until we discovered that both the baseline model and the model including the credit bureau variables made better forecasts with fewer lags. We initially used four lags of durables consumption in the model for durables consumption before discovering that it was better to use three lags of overall consumption instead. Similar results were obtained for nondurables consumption and services consumption.

No doubt, we could try many alternative specifications, but we are skeptical that the results would change in a striking manner. If we had wanted to show that the credit bureau variables were useful for forecasting, we could have used a very poor baseline model, and then the credit bureau variables would have helped in forecasting quite a bit. But it is a more convincing exercise for researchers and forecasters to start with the best possible baseline forecasting model and then see if credit bureau variables help reduce RMSFEs.

Our analysis uses real-time data, out-of-sample forecast tests, a strong parsimonious benchmark model, and data that span more than two business cycles, and reveals that credit bureau variables can be useful in improving forecasts of consumption spending in some subperiods and for some categories of consumption spending, especially for services.

References

Bram, Jason, and Sydney Ludvigson. “Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race.” Federal Reserve Bank of New York *Economic Policy Review* (June 1998), pp. 59–78.

Croushore, Dean. “Do Consumer Confidence Indexes Help Forecast Consumer Spending in Real Time?” *North American Journal of Economics and Finance* 16 (2005), pp. 435–450.

Croushore, Dean. “Frontiers of Real-Time Data Analysis.” *Journal of Economic Literature* 49 (March 2011), pp. 72–100.

Croushore, Dean, and Tom Stark. “A Real-Time Data Set for Macroeconomists.” *Journal of Econometrics* 105 (November 2001), pp. 111–130.

Diebold, Francis X. *Elements of Forecasting*. Cincinnati: Cengage (2007), 4th edition.

Duarte, Cláudia, Paulo M.M. Rodrigues, and António Rua. “A Mixed Frequency Approach to the Forecasting of Private Consumption with ATM/POS Data.” *International Journal of Forecasting* 33 (2017), pp. 61–75.

Ermişoğlu, Ergun, Yasin Akçelik, and Arif Oduncu. “Nowcasting GDP Growth with Credit Data: Evidence from an Emerging Market Economy.” *Borsa Istanbul Review* 13 (2013), pp. 93–98.

Federal Trade Commission. Accuracy in Consumer Reporting: An FTC/CFPB Workshop. (December 10, 2019).

Galbraith, John W., and Greg Tkacz. “Electronic Transactions as High-Frequency Indicators of Economic Activity.” Bank of Canada Working Paper 2007-58 (December 2007).

Galbraith, John W., and Greg Tkacz. “Nowcasting GDP with Electronic Payments Data.” European Central Bank Statistics Papers Series 10 (August 2015).

Gil, Maria, Javier J. Pérez, A. Jesús Sánchez, and Alberto Urtasun. “Nowcasting Private Consumption: Traditional Indicators, Uncertainty Measures, Credit Cards and Some Internet Data.” Banco de Espana Working Paper 1842 (2018).

Gill, Troy, Dilhan Perera, and David Sunner. “Electronic Indicators of Economic Activity.” Reserve Bank of Australia *Bulletin* (June Quarter 2012), pp. 1–12.

Haberler, Gottfried. *Consumer Instalment Credit and Economic Fluctuations*. New York: National Bureau of Economic Research (1942).

Hall, Robert E. “Stochastic Implications of the Life Cycle — Permanent Income Hypothesis: Theory and Evidence.” *Journal of Political Economy* 86 (1978), pp. 971–987.

Harvey, David, Stephen Leybourne, and Paul Newbold, “Testing the Equality of Prediction Mean Squared Errors.” *International Journal of Forecasting* 13 (1997), pp. 281–291.

Hendricks, Gary, and Kenwood C. Youmans. *Consumer Durables and Installment Debt: A Study of American Households*. Ann Arbor: Survey Research Center, University of Michigan (1973).

Maki, Dean M. “The Growth of Consumer Credit and the Household Debt Service Burden,” in Thomas A. Durkin and Michael E. Staten (eds.), *The Impact of Public Policy on Consumer Credit*. Boston: Kluwer Academic Publishers, pp. 43–68 (2002).

Verbaan, Roy, Wilko Bolt, and Carin van der Cruijsen. “Using Debit Card Payments Data for Nowcasting Dutch Household Consumption.” DeNederlandscheBank Working Paper 571 (September 2017).

Wiermanski, Chet, and Stephanie Wilshusen. “Exploring the Use of Anonymized Consumer Credit Information to Estimate Economic Conditions: An Application of Big Data.” Federal Reserve Bank of Philadelphia Payment Cards Center Discussion Paper 15-05 (November 2015).

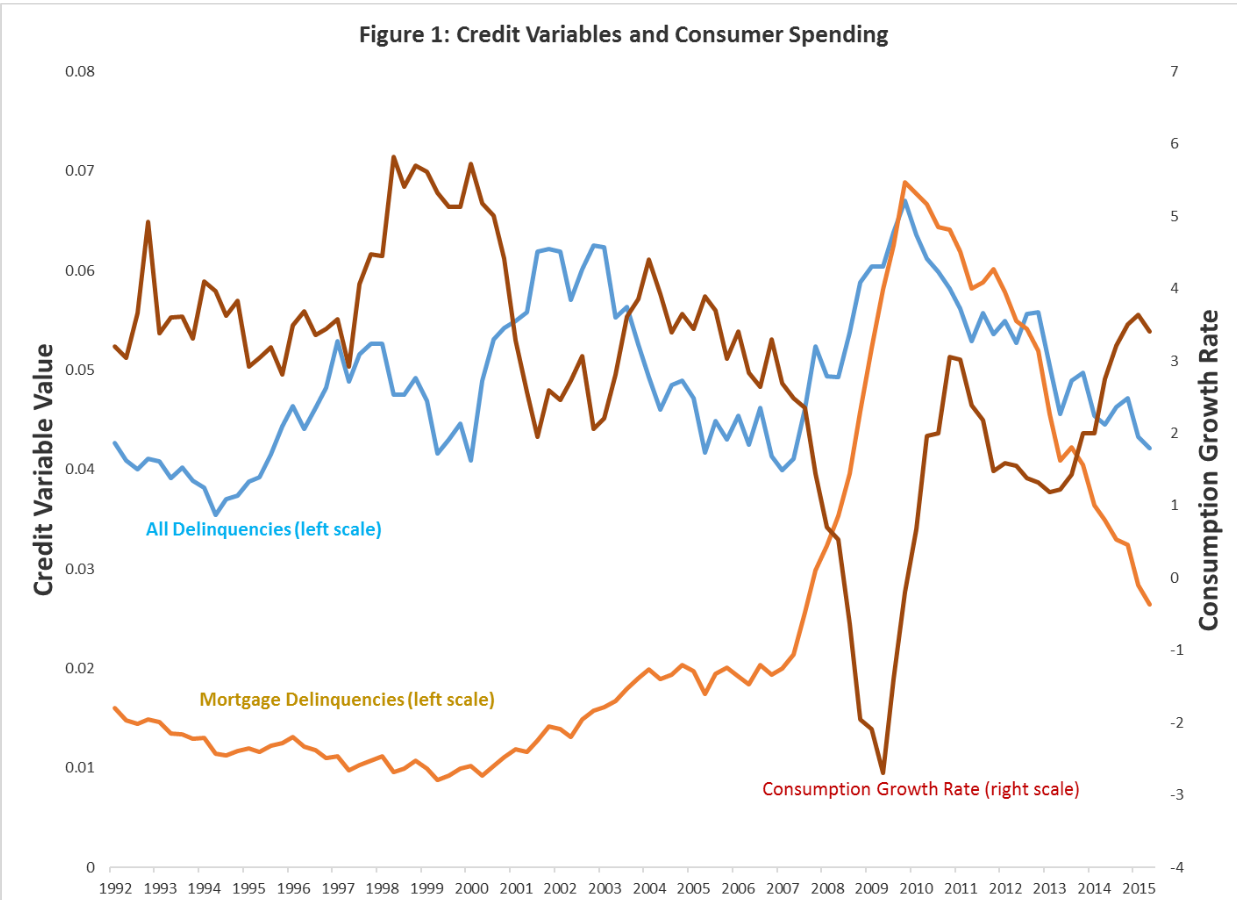

Figure 1: Credit Variables and Consumer Spending
Source: Authors' calculations using data from TransUnion and the Real-Time Data Set for Macroeconomists

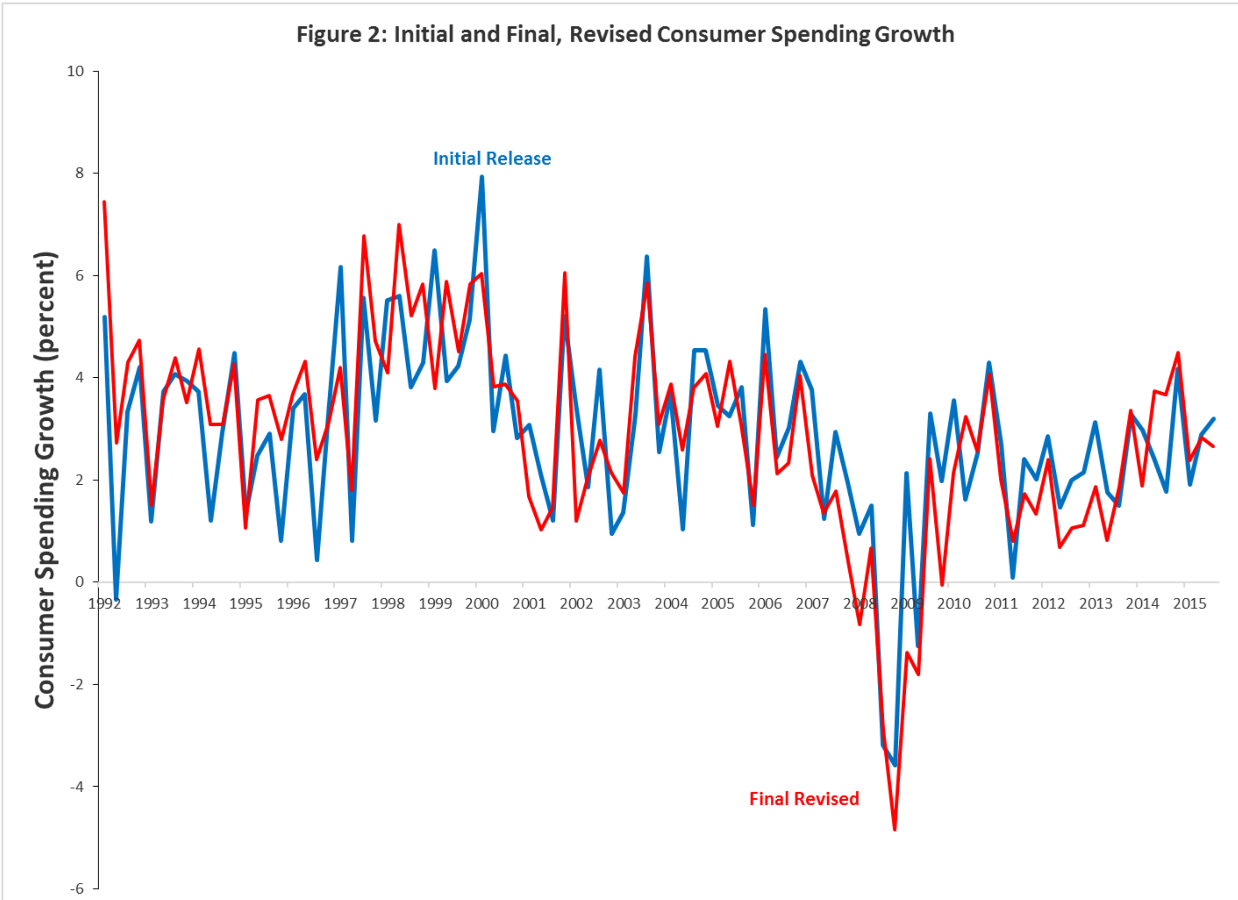

Figure 2: Initial and Final, Revised Spending Growth
Source: Real-Time Data Set for Macroeconomists produced by the Federal Reserve Bank of Philadelphia

Table 1: RRMSEs for Credit Bureau Variables						
	Level		Change		Level and Change	
Variable	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
INQUIRE	1.072	0.46	1.067	0.48	1.080	0.50
DELINQ	0.990	0.76	0.982	0.61	0.994	0.88
MDELINQ	0.992	0.95	0.989	0.90	1.025	0.85
AVGDEBT	0.970	0.50	0.990	0.74	0.980	0.63
UTILIZE	0.997	0.93	0.989	0.73	1.002	0.96
SCORE	1.049	0.37	1.036	0.46	1.008	0.88

Notes: The actual value used is the pre-benchmark value. The forecast period is 2002Q1–2015Q3 for variables other than INQUIRE and SCORE, for which it is 2007Q1–2015Q3.

Table 2: RRMSEs for Credit Bureau Variables, Consumption Categories						
	Level		Change		Level and Change	
Variable	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
Consumption of Durable Goods						
INQUIRE	1.124	0.16	1.088	0.25	1.117	0.26
DELINQ	1.012	0.53	0.977	0.25	0.989	0.68
MDELINQ	1.049	0.68	0.999	0.99	1.061	0.63
AVGDEBT	1.002	0.95	1.003	0.86	1.003	0.90
UTILIZE	1.014	0.55	0.993	0.68	1.011	0.58
SCORE	1.043	0.22	1.089	0.02**	1.039	0.36
Consumption of Nondurable Goods						
INQUIRE	1.127	0.51	1.100	0.49	1.141	0.46
DELINQ	1.001	0.97	0.980	0.47	1.008	0.82
MDELINQ	1.025	0.65	0.994	0.88	1.038	0.53
AVGDEBT	1.019	0.54	0.967	0.13	1.007	0.82
UTILIZE	0.996	0.90	0.977	0.35	0.997	0.93
SCORE	1.076	0.56	1.054	0.64	1.029	0.79
Consumption of Services						
INQUIRE	1.091	0.09	1.041	0.32	1.103	0.06*
DELINQ	0.983	0.59	0.979	0.52	0.979	0.55
MDELINQ	1.006	0.93	0.997	0.95	1.061	0.45
AVGDEBT	0.947	0.23	0.976	0.47	0.954	0.30
UTILIZE	1.015	0.60	0.976	0.46	1.008	0.77
SCORE	1.036	0.34	0.990	0.85	1.018	0.75

Notes: The actual value used is the pre-benchmark value. The forecast period is 2002Q1–2015Q3 for variables other than INQUIRE and SCORE, for which it is 2007Q1–2015Q3.

Table 3: RRMSEs, Pre-Recession Sample (2002 to 2007)						
	Level		Change		Level and Change	
Variable	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
Overall Consumption						
DELINQ	1.003	0.94	0.955	0.16	0.990	0.84
MDELINQ	0.939	0.38	0.928	0.32	0.928	0.46
AVGDEBT	0.984	0.80	0.950	0.40	0.979	0.79
UTILIZE	0.995	0.92	1.024	0.50	1.035	0.44
Consumption of Durable Goods						
DELINQ	1.050	0.17	0.942	0.05**	0.996	0.94
MDELINQ	0.972	0.57	0.953	0.39	0.948	0.53
AVGDEBT	1.018	0.65	0.982	0.49	1.018	0.67
UTILIZE	1.005	0.72	1.042	0.22	1.011	0.66
Consumption of Nondurable Goods						
DELINQ	1.038	0.44	0.985	0.29	1.055	0.23
MDELINQ	1.050	0.40	1.012	0.77	1.078	0.28
AVGDEBT	1.079	0.13	0.927	0.08*	1.028	0.67
UTILIZE	1.018	0.71	0.969	0.21	1.009	0.87
Consumption of Services						
DELINQ	0.913	0.12	0.922	0.12	0.934	0.22
MDELINQ	0.934	0.64	0.863	0.11	0.938	0.69
AVGDEBT	0.894	0.09*	0.908	0.09*	0.868	0.10*
UTILIZE	0.901	0.08*	0.915	0.07*	0.909	0.07*

Notes: The actual value used is the pre-benchmark value. The forecast period is 2002Q1–2007Q4.

Table 4: RRMSEs, Shorter Forecast Period (2007 to 2015)						
	Level		Change		Level and Change	
Variable	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
Overall Consumption						
DELINQ	0.980	0.61	0.981	0.67	0.983	0.72
MDELINQ	0.989	0.94	0.983	0.88	1.024	0.88
AVGDEBT	0.952	0.37	0.992	0.82	0.967	0.47
UTILIZE	0.993	0.89	0.972	0.46	0.986	0.77
Consumption of Durable Goods						
DELINQ	0.988	0.52	0.989	0.67	0.977	0.45
MDELINQ	1.076	0.67	1.007	0.95	1.100	0.58
AVGDEBT	0.985	0.69	1.009	0.59	0.987	0.71
UTILIZE	1.018	0.60	0.989	0.60	1.011	0.69
Consumption of Nondurable Goods						
DELINQ	0.980	0.50	0.972	0.53	0.972	0.52
MDELINQ	1.014	0.86	0.993	0.91	1.027	0.76
AVGDEBT	0.985	0.67	0.984	0.53	0.997	0.91
UTILIZE	0.980	0.67	0.979	0.57	0.988	0.79
Consumption of Services						
DELINQ	1.019	0.60	1.000	0.99	1.001	0.98
MDELINQ	1.021	0.75	1.033	0.63	1.083	0.33
AVGDEBT	0.985	0.74	1.003	0.93	0.988	0.80
UTILIZE	1.064	.02**	1.000	1.00	1.049	0.10*

Notes: The actual value used is the pre-benchmark value. The forecast period is 2007Q1–-2015Q3.

Table 5: RRMSEs, Longer Pre-Recession Forecast Period (1997 to 2007)						
	Level		Change		Level and Change	
Variable	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
Overall Consumption						
DELINQ	0.971	0.54	0.978	0.57	0.961	0.50
MDELINQ	0.985	0.67	0.979	0.65	1.020	0.70
AVGDEBT	1.036	0.45	0.971	0.43	1.067	0.31
UTILIZE	0.991	0.83	0.983	0.72	0.995	0.92
Consumption of Durable Goods						
DELINQ	1.060	0.02**	0.995	0.89	1.032	0.34
MDELINQ	1.021	0.50	1.000	1.00	1.032	0.59
AVGDEBT	1.076	0.03**	1.030	0.49	1.088	0.13
UTILIZE	1.052	0.07*	1.025	0.33	1.058	0.08*
Consumption of Nondurable Goods						
DELINQ	1.057	0.10*	1.026	0.14	1.077	0.03**
MDELINQ	1.047	0.14	1.039	0.14	1.093	0.05**
AVGDEBT	1.083	0.01**	1.001	0.98	1.085	0.04**
UTILIZE	1.037	0.28	1.029	0.55	1.087	0.15
Consumption of Services						
DELINQ	0.883	0.10*	0.904	0.18	0.861	0.08*
MDELINQ	0.919	0.25	0.880	0.09*	0.904	0.31
AVGDEBT	0.966	0.68	0.993	0.93	1.036	0.78
UTILIZE	0.889	0.09*	0.935	0.39	0.899	0.20

Notes: The actual value used is the pre-benchmark value. The forecast period is 1997Q1–2007Q4.

Table 6: RRMSEs, Longer Forecast Period Including Recession (1997 to 2015)						
	Level		Change		Level and Change	
Variable	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
Overall Consumption						
DELINQ	0.978	0.48	0.984	0.61	0.978	0.56
MDELINQ	0.997	0.97	0.994	0.93	1.039	0.65
AVGDEBT	1.003	0.94	0.987	0.61	1.027	0.52
UTILIZE	0.994	0.85	0.980	0.52	0.992	0.83
Consumption of Durable Goods						
DELINQ	1.026	0.12	0.995	0.82	1.009	0.68
MDELINQ	1.055	0.55	1.012	0.83	1.075	0.44
AVGDEBT	1.036	0.16	1.022	0.35	1.043	0.21
UTILIZE	1.036	0.10*	1.007	0.67	1.035	0.11
Consumption of Nondurable Goods						
DELINQ	1.022	0.37	1.004	0.86	1.033	0.24
MDELINQ	1.031	0.45	1.014	0.68	1.058	0.23
AVGDEBT	1.037	0.14	0.996	0.84	1.045	0.09*
UTILIZE	1.012	0.67	1.007	0.81	1.044	0.27
Consumption of Services						
DELINQ	0.926	0.14	0.937	0.22	0.906	0.10*
MDELINQ	0.959	0.44	0.940	0.26	0.975	0.72
AVGDEBT	0.970	0.63	0.999	0.98	1.025	0.78
UTILIZE	0.948	0.26	0.958	0.43	0.950	0.36

Notes: The actual value used is the pre-benchmark value. The forecast period is 1997Q1–2015Q3.

Table 7: RRMSEs for SCORE Variable with Different Baseline Models						
	Level		Change		Level and Change	
Number of lags in baseline model	RRMSFE	*p*-value	RRMSFE	*p*-value	RRMSFE	*p*-value
1	0.936	0.18	0.870	0.10*	0.860	0.11
3	1.049	0.37	1.036	0.46	1.008	0.88
5	1.067	0.38	1.085	0.20	1.049	0.37

Notes: The actual value used is the pre-benchmark value. The forecast period is 2007Q1–2015Q3.