Measuring the Credit Gap

James D. Hamilton (jhamilton@ucsd.edu)
Daniel Leff (dleffyaf@ucsd.edu)
University of California at San Diego

March 6, 2020

Abstract

We revisit the analysis by Drehmann and Yetman (2018) and conclude that measuring the credit gap based on the 5-year growth rate of the credit-to-GDP ratio produces a more reliable and robust predictor of financial crises than does the Hodrick-Prescott filtered series. We also conclude that estimating the credit gap based on the forecast error of a 5-year-ahead regression can be even more useful, provided a sufficiently long sample is available to estimate coefficients of the regression.

A recent paper by Drehmann and Yetman (2018) defended the use of the Hodrick-Prescott filter as a tool for identifying deviations from trend in the credit-to-GDP ratio for purposes of setting countercyclical capital buffers under Basel Banking Supervision guidelines. Their conclusions have been cited by over a dozen other studies, including Beltran, Jahan-Parvar, and Paine (2019), Reichlin, Ricco and Hasenzagl (2019), Schüler (2019), Sprincean (2019), and De Jong and Sakarya (2020). In this paper we revisit their findings and reach a different conclusion.

Hamilton (2018) proposed interpreting the cyclical component of a time series as the component that could not be predicted two years in advance based on the variable's own lagged values. He showed that this component can typically be estimated from the residuals of a simple linear regression, though he noted that for many economic and financial time series we could accomplish something very similar simply by taking the 2-year change in the variable. He recommended a 2-year horizon on the grounds that "the primary reason that we would be wrong in predicting the value of most macro and financial variables at a horizon of $h = 8$ quarters ahead is cyclical factors such as whether a recession occurs over the next two years and the timing of recovery from any downturn." He further raised a practical reason for not using a larger value of h: "a bigger sample size T will be needed the bigger is h. The information in a finite data set about very long-horizon forecasts is quite limited."

Credit cycles are typically viewed as evolving much slower than business cycles. For analyzing credit cycles, Hamilton recommended, "For such an application I would use $h = 5$ years, with the regression-free implementation $(y_{t+5} - y_t)$ having particular appeal given the length of datasets available."

Drehmann and Yetman (2018) compare Hamilton's two approaches with the popular Hodrick-Prescott characterization of trends for purposes of identifying the cyclical component of the credit-to-GDP ratio. They use quarterly observations on y_{it}, the ratio of credit to GDP for country i in quarter t for $i = 1, ..., 42$ different countries[1]. The first observations on y_{it} come as early as 1970:Q1 for some countries while data for others like Peru only begin in 2001:Q4. They use data through date T_i to estimate a linear 5-year-ahead predictive regression for country i,

$$y_{it} = \beta_{i0} + \beta_{i1} y_{i,t-20} + \beta_{i2} y_{i,t-21} + \beta_{i3} y_{i,t-22} + \beta_{i4} y_{i,t-23} + v_{i,t} \quad \text{for } t = 1, 2, ..., T_i \tag{1}$$

for T_i the observation 10 years after the data for country i begins. They are interested in the usefulness of the estimated residual $\hat{v}_{i,T_i}$ for purposes of predicting a financial crisis in that country at date $T_i + s$ for different forecasting horizons $s = 1, ..., 12$ quarters. They then re-estimate the regression using data now through date $T_i + 1$ in order to predict a financial crisis at $T_i + 1 + s$, and so on through the end of the sample. The average area under the Receiver

[1] This is measured as credit to the private non-financial sector, capturing total borrowing from all domestic and foreign sources. See Dembiermont, Drehmann and Muksakunratana (2013).

Operating Characteristic curve (denoted AUC) for $\hat{v}_{i,T_i}$ as a predictor of financial crisis at $T_i + s$ is then calculated across T_i and across countries. Panel A of Figure 1 reproduces some of their key results. The dot-dashed blue line plots the value of AUC when the predictor is the residual from (1) as a function of the horizon s. Drehmann and Yetman compared this with the signal provided by v^*_{i,T_i}, the terminal residual from an HP filter applied to country i through date T_i using a smoothing parameter of $\lambda = 400,000$ (shown in solid black in Panel A) and with the simple five-year growth rate $v^\dagger_{i,T_i} = \log(y_{i,T_i}) - \log(y_{i,T_i-20})$ (dashed red).

The estimated regression does significantly worse than either HP or the 5-year growth. HP does slightly better than the 5-year growth, though the differences are not statistically significant, and the ordering is sensitive to the sample period used. For example, if we start the evaluation with $T_i = 1995$:Q1 instead of the earliest possible date $T_i = 1980$:Q1, the 5-year growth does slightly better than HP (see Panel B).

These results raise two interesting questions. First, why does the estimating regression (1) do so poorly, and second, why does HP do so well?

The answer to the first question is that this exercise uses very short horizons to estimate the parameters in regression (1). For example, the first regression for the United States uses $T_i = 1980$:Q1. Because the regression is estimating a 5-year-ahead forecast and includes 4 lags beyond those 5 years, the dependent variable in regression (1) only runs from $t =$ 1975:Q4 through 1980:Q1, 18 quarterly observations to estimate 5 coefficients. Moreover, the residuals $v_{i,t}$ from this regression are very highly serially correlated due to the 5-year overlapping structure. One would expect regression estimates to perform extremely poorly in this setting, and indeed Hamilton (2018) recommended using the full sample to estimate the trend parameters rather than expanding windows as in the Drehmann and Yetman analysis for exactly this reason. Panel C shows the AUC for the three approaches when the parameters of regression (1) are estimated using the full sample of observations on each country.[2] When implemented this way, the regression approach turns out to do a little better than either of the other two.

To answer the second question, note that the smoothing parameter $\lambda = 400,000$ that Drehmann and Yetman use in their exercise is two orders of magnitude bigger than the value $\lambda = 1600$ that is typically used for quarterly data. Indeed if one were to use $\lambda = 1600$ the performance of HP turns out to be far worse, as seen in Panel D.

When applied to quarterly data, HP with $\lambda = 400,000$ is almost equivalent to just estimating a time trend. Figure 2 plots credit/GDP for the U.S. along with full-sample estimates resulting from HP with $\lambda = 400,000$ and the estimate of a deterministic time trend. HP and the linear trend differ little. Indeed, if we estimate a deterministic time trend for each

[2] For most countries this goes through 2017:Q3.

country with expanding samples as in regression (1),

$$y_{it} = \alpha_i + \delta_i t + \tilde{v}_{i,t} \quad \text{for } t = 1, 2, ..., T_i, \tag{2}$$

the deterministic time trend if anything performs slightly better than does HP (see Figure 3).

Is a deterministic time trend a plausible description of these data? An augmented Dickey Fuller test[3] fails to reject the null hypothesis of a unit root in y_{it} at the 10% level for 39 out of the 42 countries. The KPSS test[4] of Shin and Schmidt (1992), which takes the deterministic time trend as the null, rejects that null in favor of a unit root alternative for 40 out of the 42 countries. This makes it hard to justify (2) as a characteristic of the data-generating process that could be of interest.

What then is the justification for using $\lambda = 400,000$? Drehmann and Yetman cite Borio and Lowe (2002), though that study appears to use annual data. Ravn and Uhlig (2002) argued that the value of $\lambda = 1600$ that is used in almost all applications of the HP filter to quarterly data is analogous to using $\lambda = 129,600$ on monthly or $\lambda = 6.25$ on annual data. One is tempted to conclude that the choice of $\lambda = 400,000$ in the present context is motivated at least in part by the fact that the resulting estimate of the gap was found to be correlated with financial crises by other researchers using earlier data sets. To the extent this is the case, it calls into question whether Drehmann and Yetman's exercise can truly be characterized as a clean out-of-sample forecast evaluation.

The simple 5-year growth is a robust, model-free and estimation-free alternative that performs almost as well or better than HP with $\lambda = 400,000$ and far better than HP with conventional smoothing weights. Furthermore, estimating parameters with a linear regression as in (1) also appears to be a useful approach in this setting, though we would probably want 40 years of data before trusting those estimates for purposes of identifying the cyclical component in the credit-to-GDP ratio.

[3] We used the Case 4 test described in Hamilton (1994, p. 529) with $p = 4$ lags and a time trend included in the regression.

[4] Again we use $p = 4$ lags and include a trend.

References

Beltran, Daniel O., Mohammad R. Jahan-Parvar, and Fiona A. Paine (2019). "Designing Gaps as Early Warning Indicators: Accuracy, Predictive Power, and Robustness," working paper, Federal Reserve Board.

Borio, Claudio EV, and Philip William Lowe (2002). "Asset prices, financial and monetary stability: exploring the nexus," BIS working paper 114.

De Jong, Robert M., and Neslihan Sakarya (2020). "The one-sided Hodrick-Prescott filter," working paper, Ohio State University.

Dembiermont, Christian, Mathias Drehmann, and Siriporn Muksakunratana (2013). "How Much Does the Private Sector Really Borrow? A New Database for Total Credit to the Private Non-financial Sector," *BIS Quarterly Review*, March 2013.

Drehmann, Mathias, and James Yetman (2018). "Why You *Should* Use the Hodrick-Prescott Filter– At Least to Assess Credit Gaps," working paper, Bank for International Settlements.

Hamilton, James D. (1994). *Time Series Analysis,* Princeton University Press.

Hamilton, James D. (2018). "Why You Should Never Use the Hodrick-Prescott Filter," *Review of Economics and Statistics,* 100: 831-843.

Ravn, Morten O., and Harald Uhlig (2002). "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics* 84: 371-376.

Reichlin, Lucrezia, Giovanni Ricco, and Thomas Hasenzagl (2019). "Financial Variables as Predictors of Real Growth Vulnerability," working paper, London Business School.

Schüler, Yves Stephan (2019). "How should we filter economic time series?" Working paper, Deutsche Bundesbank.

Shin, Yongcheol, and Peter Schmidt (1992). "The KPSS stationarity test as a unit root test," *Economics Letters* 38: 387-392.

Sprincean, Nicu (2019). "Early warning indicators for macrofinancial activity in Romania," *Review of Economic and Business Studies* 12: 137-162.

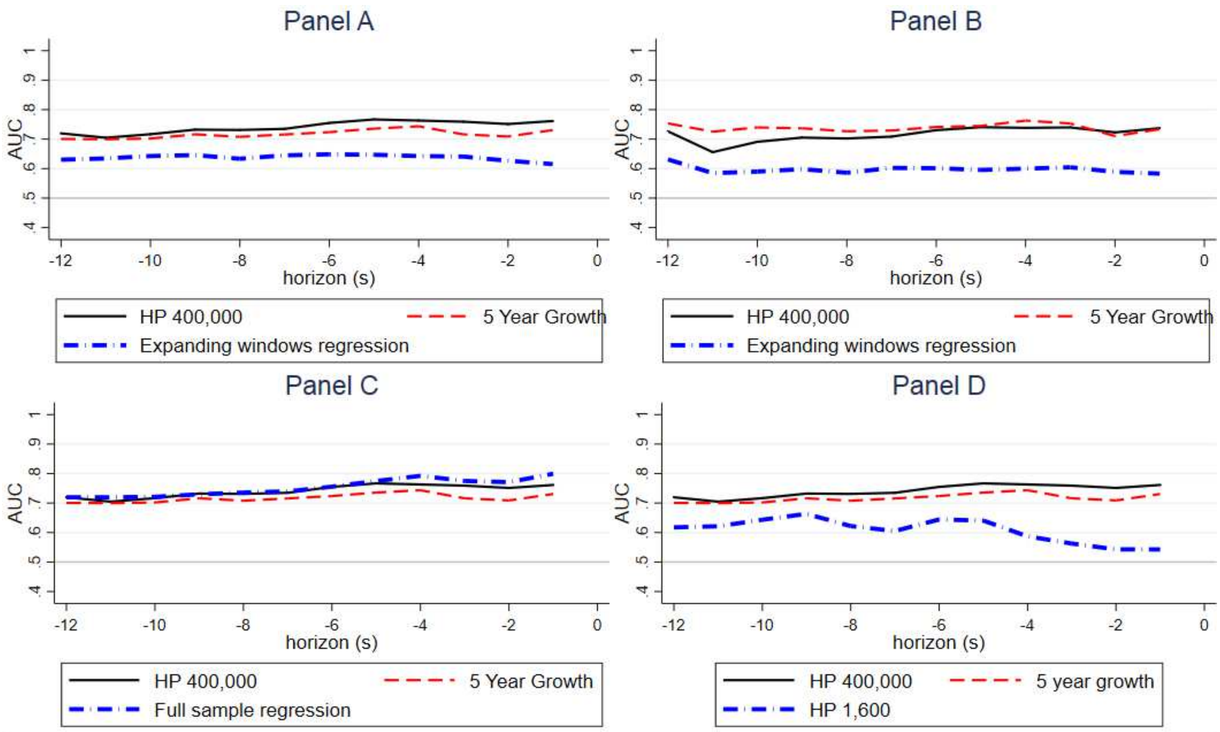

Figure 1. AUC as a function of horizon *s* based on cyclical component of credit/GDP estimated in different ways.

Notes to Figure 1. Panel A: credit gap as assessed using (1) HP with λ = 400,000; (2) 5-year change in the log of credit/GDP, and (3) regression estimated from credit/GDP data through date T_i to predict crisis s quarters after T_i, with evaluations beginning 10 years after the first observation of credit/GDP for each country. Panel B: same as panel A except evaluations begin 1995:Q1. Panel C: same as Panel A except regression parameters estimated using full sample of observations for each country. Panel D: (1)-(2) as in Panel A with (3) replaced by HP with λ = 1600.

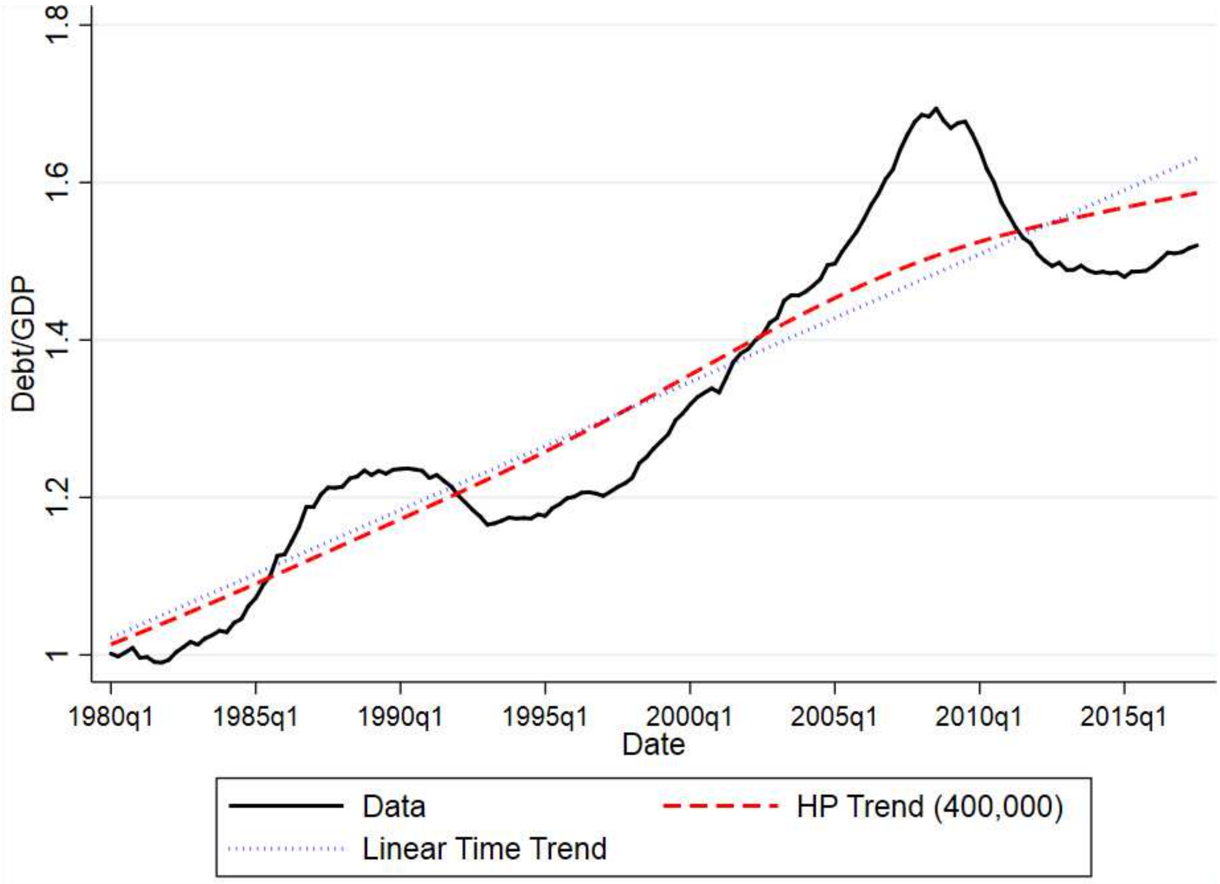

Figure 2. U.S. credit/GDP ratio 1980:Q1 to 2017:Q4 and full-sample estimates of (1) HP trend with λ = 400,000; (2) deterministic time trend.

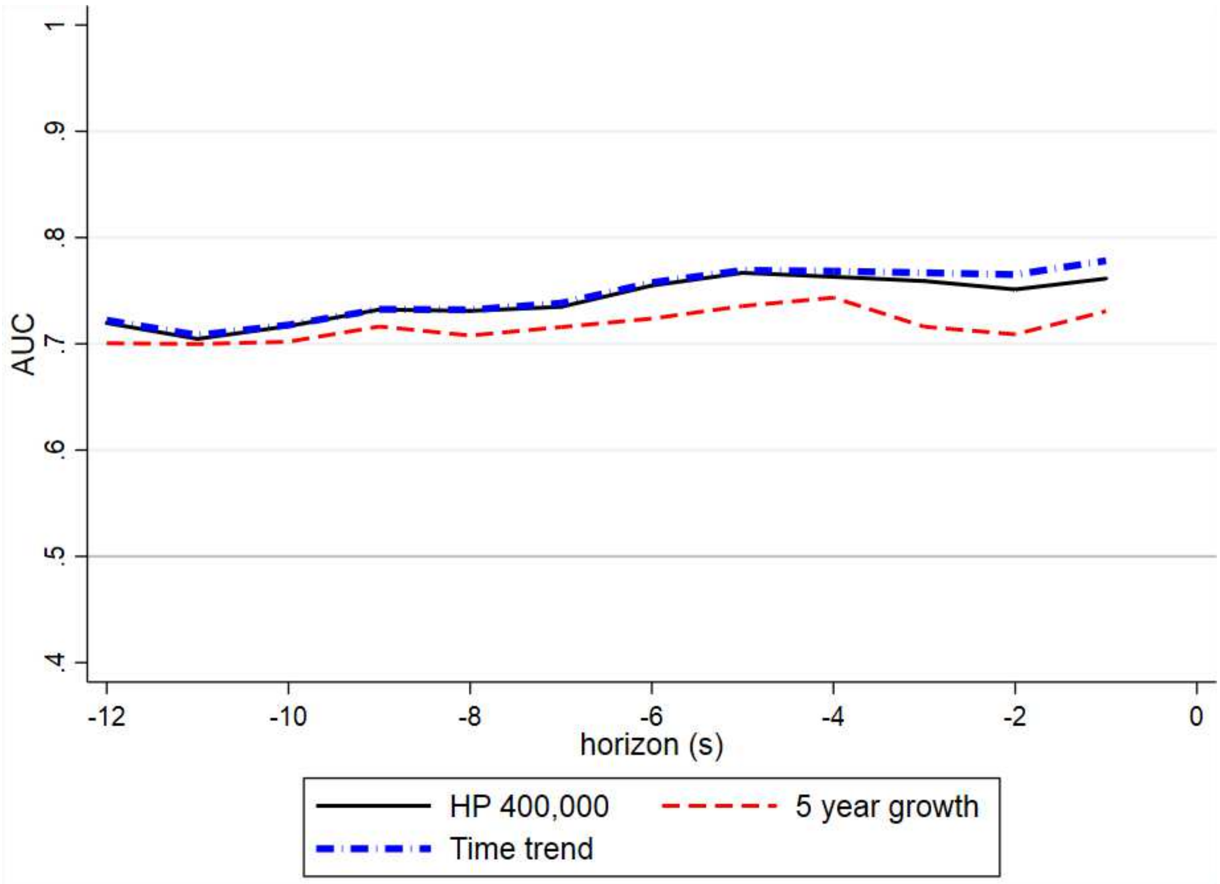

Figure 3. AUC as a function of horizon *s* based on cyclical component of credit/GDP estimated from (1) HP with λ = 400,000; (2) 5-year change in the log of credit/GDP, and (3) residuals from deterministic time trend estimated from regressions of expanding sample size.