

**Finance and Economics Discussion Series  
Divisions of Research & Statistics and Monetary Affairs  
Federal Reserve Board, Washington, D.C.**

**Credit Spreads as Predictors of Real-Time Economic Activity: A  
Bayesian Model-Averaging Approach**

**Jon Faust, Simon Gilchrist, Jonathan H. Wright, and Egon  
Zakrajsek**

**2012-77**

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

# Credit Spreads as Predictors of Real-Time Economic Activity: A Bayesian Model-Averaging Approach

Jon Faust\*    Simon Gilchrist†    Jonathan H. Wright‡    Egon Zakrajšek§

October 23, 2012

Forthcoming in the *Review and Economics and Statistics*

## Abstract

Employing a large number of financial indicators, we use Bayesian Model Averaging (BMA) to forecast real-time measures of economic activity. The indicators include credit spreads based on portfolios—constructed directly from the secondary market prices of outstanding bonds—sorted by maturity and credit risk. Relative to an autoregressive benchmark, BMA yields consistent improvements in the prediction of the cyclically-sensitive measures of economic activity at horizons from the current quarter out to four quarters hence. The gains in forecast accuracy are statistically significant and economically important and owe almost exclusively to the inclusion of credit spreads in the set of predictors.

JEL CLASSIFICATION: C11, C53

KEYWORDS: forecasting, real-time data, Bayesian Model Averaging, credit spreads

---

This is a revised version of the paper that appeared under the same title as the NBER Working Paper No. 16725, January 2011. We are grateful to two anonymous referees and Mark Watson (the Editor) for their detailed comments and suggestions. We also thank John Duca, Lutz Kilian, Michael McCracken, Emanuel Moench, Demien Pouzo, Barbara Rossi, and participants at the 2011 NYU Economics Alumni Conference, the 2012 “Day Ahead” Conference on Financial Markets and Institutions, the 2012 Texas Econometrics Camp, and the 2012 Macro Financial Modeling Meeting at NYU for helpful comments. Samuel Haltenhof, Robert Kurtzman, Michael Levere, and Ben Rump provided outstanding research assistance at various stages of the project. All errors and omissions are our own. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System.

\*BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM, JOHNS HOPKINS UNIVERSITY, AND NBER.

†BOSTON UNIVERSITY AND NBER.

‡JOHNS HOPKINS UNIVERSITY AND NBER.

§BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM.

# 1 Introduction

One area of agreement among economists at universities, central banks, and on Wall Street is that forecasting economic activity is hard. While the existing methods give us some ability to forecast economic developments for the current quarter and perhaps the quarter after that, their predictive power is modest at best and deteriorates rapidly as the forecast horizon extends beyond the very near term. Moreover, what little predictability there seems to be appears to be captured about as well by simple models—such as a univariate autoregression—as by the large number of complex statistical and DSGE forecasting methods that have been proposed in the literature; see, for example, Sims [2005]; Tulip [2005]; Faust and Wright [2009]; and Edge and Gürkaynak [2010].

Economists have long sought to improve on this record by using information from financial markets. Because they are inherently forward looking, the argument goes, financial market prices should impound information about investors’ expectations of future economic outcomes.<sup>1</sup> From a theoretical perspective, default-risk indicators such as credit spreads—the difference in yields between various corporate debt instruments and government securities of comparable maturity—are particularly well suited for forecasting economic activity. Philippon [2009], for example, presents a model in which the decline in investment fundamentals, owing to a reduction in the expected present-value of corporate cash flows, leads to a widening of credit spreads prior to a cyclical downturn. As emphasized by Bernanke, Gertler, and Gilchrist [1999] and Gilchrist and Zakrajšek [2012], increases in credit spreads can also signal disruptions in the supply of credit resulting from the worsening in the quality of corporate balance sheets or from the deterioration in the health of financial intermediaries that supply credit.<sup>2</sup>

The empirical success of default-risk indicators as predictors of economic activity is decidedly mixed, however, with results varying substantially across various credit spread indexes and different time periods. For example, the “paper-bill” spread—the difference between yields on non-financial commercial paper and comparable-maturity Treasury bills—had substantial forecasting power for economic activity during the 1970s and the 1980s, but its predictive ability vanished in the subsequent decade. In contrast, credit spreads based on indexes of speculative-grade (i.e., “junk”) corporate bonds, which contain information from markets that were not in existence before the mid-1980s, did particularly well at forecasting output growth during the 1990s, according to Gertler and Lown [1999] and Mody and Taylor [2004]. Stock and Watson [2003], however, show

---

<sup>1</sup>Financial indicators considered in this vast literature include stock prices (Fama [1981] and Harvey [1989]); spreads between long and short-term risk-free interest rates (Harvey [1988]; Estrella and Hardouvelis [1991]; Estrella and Mishkin [1998]; and Hamilton and Kim [2002]); the term structure of interest rates more generally (Ang, Piazzesi, and Wei [2006]); spreads between rates on short-term commercial paper and rates on Treasury bills (Bernanke [1990]; Friedman and Kuttner [1992, 1998]; and Emery [1999]); and yield spreads on longer-term corporate debt (Gertler and Lown [1999]; King, Levin, and Perli [2007]; Mueller [2009]; Gilchrist, Yankov, and Zakrajšek [2009]; and Gilchrist and Zakrajšek [2012]).

<sup>2</sup>In fact, in a recent paper, Del Negro and Schorfheide [2012] show that by incorporating credit spreads into relatively simple DSGE models noticeably improved—relative to the Blue Chip Economic Consensus forecast—the ability of such models to forecast the “Great Recession.”

that the forecasting ability of this default-risk indicator is quite uneven.

In a recent paper, Gilchrist, Yankov, and Zakrajšek [2009] (GYZ hereafter) argue that these mixed results may be due to the fact that the credit spread indexes used by researchers tend to be based on aggregates of returns on a mishmash of bonds with different duration, credit risk, and other characteristics. In part to address these problems, GYZ constructed 20 monthly credit spread indexes for different maturity and credit risk categories using secondary market prices of individual senior unsecured corporate bonds.<sup>3</sup> Their findings indicate that these credit spread indexes have substantial predictive power, at both short- and longer-term horizons, for the growth of payroll employment and industrial production. Moreover, they significantly outperform the predictive ability of the standard default-risk indicators, a result that suggests that using “cleaner” measures of credit spreads may, indeed, lead to more accurate forecasts of economic activity.

This paper extends the analysis of GYZ in several dimensions. Most importantly, we provide a thorough evaluation of the marginal information content of credit spreads in real-time economic forecasting. Given the extensive and ongoing search for consistent predictors of U.S. economic activity, the macroeconomics profession runs a substantial risk that results like those of GYZ are due to researchers stumbling on variables that just happen to fit the existing sample, but which, in reality, have no predictive power. The regular breakdown of new forecasting relationships soon after they are documented confirms that this risk is real. Thus, it is especially important that any such analysis takes into account model search and selection issues.

To guard against the problem of selecting financial indicators that just happen to fit our sample, we adopt a Bayesian Model Averaging (BMA) approach and evaluate it in a pseudo out-of-sample forecasting exercise. As explained more fully below, we add the new credit spread indexes to a predictor set containing over 100 asset market indicators, as well as a large number of real variables, and begin with a prior that each predictor is equally likely to be useful in forecasting future economic activity. The posterior weight assigned to each predictor in period  $t$  is then based on a Bayesian updating scheme that uses only the information available at time  $t$ . While our BMA scheme has, under certain conditions, a formal Bayesian justification, we follow a large and growing literature that takes a frequentist perspective and relies on the BMA framework as a pragmatic approach to data-based weighting of a large number of competing prediction models<sup>4</sup>. The combination of the BMA framework and out-of-sample forecast evaluation mitigates—though does not completely eliminate—the problem of data mining.

While following GYZ’s basic approach for constructing credit spread indexes, we also improve on their methodology by adjusting the underlying micro-level credit spreads for the call option embedded in many of the underlying securities. As pointed out by Duffee [1998] and Duca [1999],

---

<sup>3</sup>GYZ measure the underlying credit risk by the issuer’s expected default frequency (EDF™), a market-based default-risk indicator calculated by Moody’s/KMV that is more timely than the issuer’s credit rating.

<sup>4</sup>See, for example, Min and Zellner [1993]; Fernandez, Ley, and Steel [2001b]; Avramov [2002]; Cremers [2002]; Sala-i-Martin, Doppelhofer, and Miller [2004]; Koop and Potter [2004]; King, Levin, and Perli [2007]; and Wright [2008].

fluctuations in the value of embedded options—reflecting shifts in the term structure of risk-free rates—can substantially alter the information content of movements in corporate bond yields at business cycle frequencies.

Our results indicate that the new credit spread indexes have considerable marginal predictive power for real-time measures of economic activity, especially those of the cyclically sensitive nature. When using the entire predictor set to forecast a wide array of economic activity indicators, the gains in the root mean-square prediction error (RMSPE)—relative to a univariate autoregressive benchmark—are statistically significant and often substantial in economic terms. BMA forecasts consistently generate reductions in out-of-sample RMSPEs on the order of 10 percent when forecasting the cumulative growth of cyclically sensitive economic indicators four quarters into the future. Consumption growth is the main exception to this general result—there are no gains in predictive accuracy relative to our benchmark for this measure of economic activity.

When we omit the credit spread indexes from the predictor set and redo the analysis, we obtain the standard result, namely, that the predictive accuracy of the BMA method—like that of most other documented forecasting methods—is statistically indistinguishable from that of the univariate autoregressive benchmark. This result indicates that there is something different about the information content of credit spreads and that our BMA weighting scheme is able to pick out this difference in real-time from a large number of predictors, all of which were treated equally *ex ante*. Indeed, the analysis of the evolution of posterior weights that the BMA scheme assigns to various variables in the predictor set shows that it is economic downturns that lead to the majority of the posterior weight being placed on the credit spreads. This finding suggests that corporate bond spreads—when properly measured—may be one of the earliest and clearest aggregators of accumulating evidence of incipient recession.

The remainder of the paper is organized as follows. Section 2 describes our bond-level data and the construction of portfolios based on the option-adjusted credit spreads. In Section 3, we outline the econometric methodology used to combine forecasts by BMA. Section 4 contains our main empirical results. In Section 5, we compare the performance of BMA forecasts at different stages of the business cycle—that is, in economic recessions and expansions. And lastly, Section 6 concludes.

## 2 Data Sources and Methods

### 2.1 Credit Spreads

The key information for our analysis comes from a large sample of fixed income securities issued by U.S. corporations.<sup>5</sup> Specifically, from the Lehman/Warga (LW) and Merrill Lynch (ML) databases,

---

<sup>5</sup>Securities issued by government-sponsored entities, such as Fannie Mae and Freddie Mac, were excluded from our analysis.

we extracted month-end prices of outstanding long-term corporate bonds traded in the secondary market between January 1986 and September 2011.<sup>6</sup> To guarantee that we are measuring borrowing costs of different firms at the same point in their capital structure, we restricted our sample to senior unsecured issues with a fixed coupon schedule only. For such securities, we spliced the month-end prices across the two data sources.

We exploit the micro-level nature of our data to construct credit spreads that are not contaminated by the maturity/duration mismatch that is a bane of most commonly-used credit spread indexes. Specifically, for each individual bond issue in our sample, we construct a theoretical risk-free security that replicates exactly the promised cash-flows of the corresponding corporate debt instrument. For example, consider a corporate bond  $k$  issued by firm  $i$  that at time  $t$  is promising a sequence of cash-flows  $\{C_s\}_{s=1}^S$ , which consists of the regular coupon payments and the repayment of the principle at maturity. The price of this bond in period  $t$  is given by

$$P_{it}[k] = \sum_{s=1}^S C_s D(t_s),$$

where  $D(t) = e^{-r_t t}$  is the discount function in period  $t$ . To calculate the price of a corresponding risk-free security—denoted by  $P_t^f[k]$ —we discount the promised cash-flow sequence  $\{C_s\}_{s=1}^S$  using continuously-compounded zero-coupon Treasury yields in period  $t$ —denoted by  $r_t$ —derived from the daily estimates of the U.S. Treasury yield curve estimated by Gürkaynak, Sack, and Wright [2007]. The resulting price  $P_t^f[k]$  can then be used to calculate the yield—denoted by  $y_t^f[k]$ —of a hypothetical Treasury security with exactly the same cash-flows as the underlying corporate bond. The credit spread  $S_{it}[k] = y_{it}[k] - y_t^f[k]$ , where  $y_{it}[k]$  denotes the yield of the corporate bond  $k$ , is thus free of the “duration mismatch” that would occur were the spreads computed simply by matching the corporate yield to the estimated yield of a zero-coupon Treasury security of the same maturity.

To ensure that our results are not driven by a small number of extreme observations, we eliminated all bond/month observations with credit spreads below 5 basis points and with spreads greater than 3,500 basis points.<sup>7</sup> In addition, we dropped from our sample very small corporate

---

<sup>6</sup>These two data sources are used to construct benchmark corporate bond indexes used by market participants. Specifically, they contain secondary market prices for a vast majority of dollar-denominated bonds publicly issued in the U.S. corporate cash market. The ML database is a proprietary data source of daily bond prices that starts in 1997. The LW database of month-end bond prices is available from 1973 through mid-1998 (see Warga [1991] for details).

<sup>7</sup>These cutoffs resulted in dropping about 3 percent of observations from the sample of bonds issued by nonfinancial firms and about 1 percent from the sample of bonds issued by their financial counterparts. A significant majority of these observations involved *negative* credit spreads, economically a nonsensical value. Although our method for calculating credit spreads does not suffer from the duration mismatch, it nevertheless yielded a negative spread in a small number of cases. Such implausible observations were likely a result of stale secondary market prices, reflecting the relatively poor liquidity of the underlying securities—while corporate bonds are actively traded, the volume of transactions is far lower than in the Treasury market (Edwards, Harris, and Piwowar [2007]).

Table 1: Corporate Bond Characteristics by Type of Firm

<i>Nonfinancial Firms</i>					
Bond Characteristic	Mean	StdDev	Min	Median	Max
No. of bonds per firm/month	3.14	3.89	1.00	2.00	76.0
Mkt. value of issue (\$mil.)	349.2	342.6	1.22	255.1	5,628
Maturity at issue (years)	12.8	9.2	1.0	10.0	50.0
Term to maturity (years)	10.4	8.3	1.0	7.4	30.0
Duration (years)	6.28	3.30	0.91	5.72	17.1
Credit rating (S&P)	-	-	D	BBB1	AAA
Coupon rate (pct.)	7.18	2.03	0.75	6.95	17.5
Nominal yield to maturity (pct.)	7.07	3.08	0.42	6.79	44.3
Credit spread (bps.)	222	290	5	132	3,499
<i>Financial Firms</i>					
Bond Characteristic	Mean	StdDev	Min	Median	Max
No. of bonds per firm/month	3.06	3.50	1.00	2.00	26.0
Mkt. value of issue (\$mil.)	486.8	566.9	9.11	274.4	4,351
Maturity at issue (years)	10.4	8.0	2.0	10.0	40.0
Term to maturity (years)	8.7	7.8	1.0	5.9	30.0
Duration (years)	5.54	3.25	0.90	4.82	15.3
Credit rating (S&P)	-	-	C	A2	AAA
Coupon rate (pct.)	6.80	1.94	2.00	6.60	15.8
Nominal yield to maturity (pct.)	6.54	2.78	0.69	6.27	41.2
Credit spread (bps.)	185	250	5	117	3,499

NOTE: Sample period: 1986:M1–2011:M9. No. of nonfinancial firms/bonds = 1,156/6,404 (Obs. = 334,685); No. of financial firms/bonds = 202/942 (Obs. = 46,135). The market value of the bond issues is deflated by the CPI (2000 = 100). Sample statistics are based on trimmed data; see text for details.

issues—those with a par value of less than \$1 million—and all observations with a remaining term-to-maturity of less than one year or more than 30 years.<sup>8</sup> These selection criteria yielded a sample of 6,404 individual securities issued by firms in the nonfinancial sector and 942 securities issued by financial firms. We matched these corporate securities with their issuer’s quarterly income and balance sheet data from Compustat and daily data on equity valuations from CRSP, yielding a matched sample of 1,156 nonfinancial firms and 202 financial firms.

Table 1 contains summary statistics for the key characteristics of bonds in our sample by the type of firm (nonfinancial vs. financial). Note that a typical firm has only a few senior unsecured issues outstanding at any point in time—the median firm in both sectors, for example, has two such

<sup>8</sup>We also eliminated a very small number of puttable bonds from our sample. In contrast, a significant fraction of the securities in our sample is callable, which raises an important issue of how to separate time-varying prepayment risk from the default risk premium. We address this issue in detail later in the paper.

issues trading at any given month. The size of bond issues, measured by their market value, tend to be somewhat larger, on average, in the financial sector. Not surprisingly, the maturity of these debt instruments is fairly long, with the average maturity at issue of more than 10 years in both sectors. Because corporate bonds typically generate significant cash flow in the form of regular coupon payments, their effective duration is considerably shorter.

According to the S&P credit ratings, our sample spans the entire spectrum of credit quality, from “single D” to “triple A.” At A2, the median bond/month observation in the financial sector is somewhat above that in the nonfinancial sector (i.e., BBB1), though they are both solidly in the investment-grade category. Turning to returns, the (nominal) coupon rate on the bonds issued by nonfinancial firms averaged 7.18 percent during our sample period, compared with 6.89 percent for bonds issued by their financial counterparts. The average expected total return was 7.29 percent per annum in the nonfinancial sector and 6.80 percent in the financial sector. Relative to Treasuries, an average bond issued by a nonfinancial firm has an expected return of about 222 basis points above the comparable risk-free rate. Reflecting their generally higher credit quality—at least as perceived by the ratings agencies—the average credit spread on a bond issued by a financial intermediary is 185 basis points.

## 2.2 Default Risk

The measurement of firm-specific default risk is the crucial input in the construction of our bond portfolios. To measure an issuer’s probability of default at each point in time, we employ the contingent claims approach to corporate credit risk developed in the seminal work of Merton [1973, 1974]. The key insight of this “distance-to-default” (DD) framework is that the equity of the firm can be viewed as a call option on the underlying value of the firm with a strike price equal to the face value of the firm’s debt. Although neither the underlying value of the firm nor its volatility is directly observable, they can, under the assumptions of the model, be inferred from the value of the firm’s equity, the volatility of its equity, and the firm’s observed capital structure.

Formally, the distance-to-default—essentially, a volatility-adjusted measure of leverage—is given by

$$DD = \frac{\ln(V/D) + (\mu_V - 0.5\sigma_V^2)}{\sigma_V},$$

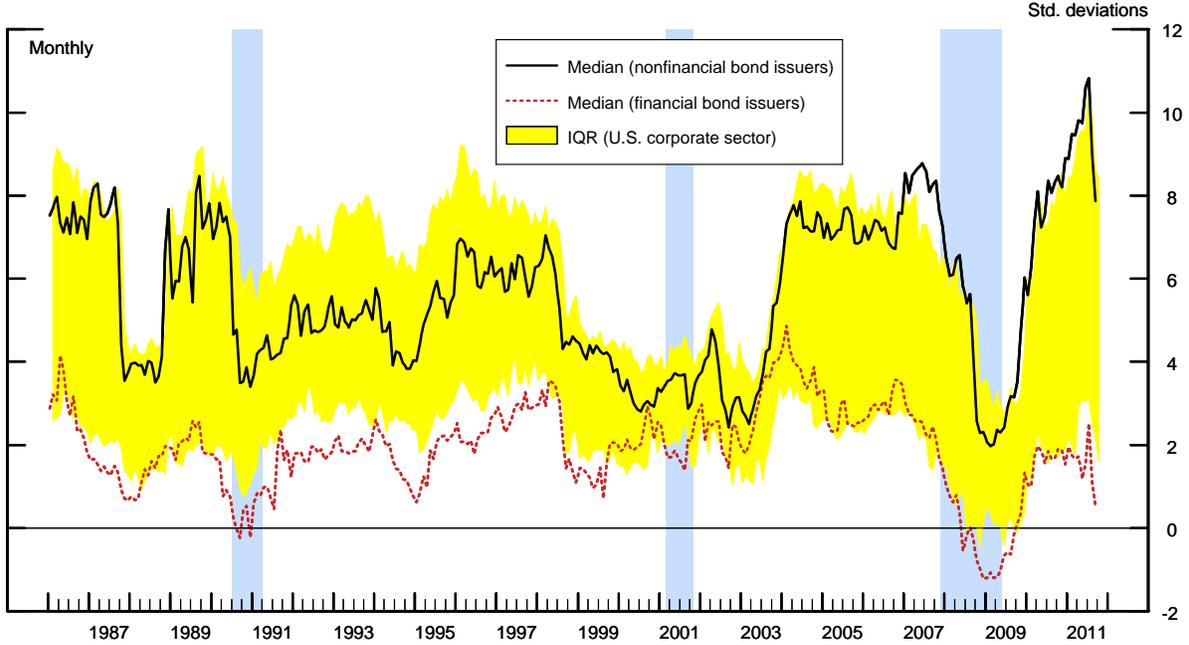
where  $V$  is the total value of the firm (i.e., the market value of its debt and equity),  $D$  is the book-value of firm’s liabilities,  $\mu_V$  is the expected return on  $V$ , and  $\sigma_V$  is the volatility of firm value. Theoretically, default occurs when the ratio of the value of assets to debt falls below one (or its log is negative)—in effect, distance-to-default measures the number of standard deviations the log of this ratio must deviate from its mean for default to occur.<sup>9</sup>

The numerical procedure used to construct this market-based measure of default risk is based on

---

<sup>9</sup>In this context, the probability of default is given by  $\Phi(-DD)$ , where  $\Phi$  denotes the standard normal CDF.

Figure 1: Distance-to-Default by Type of Firm



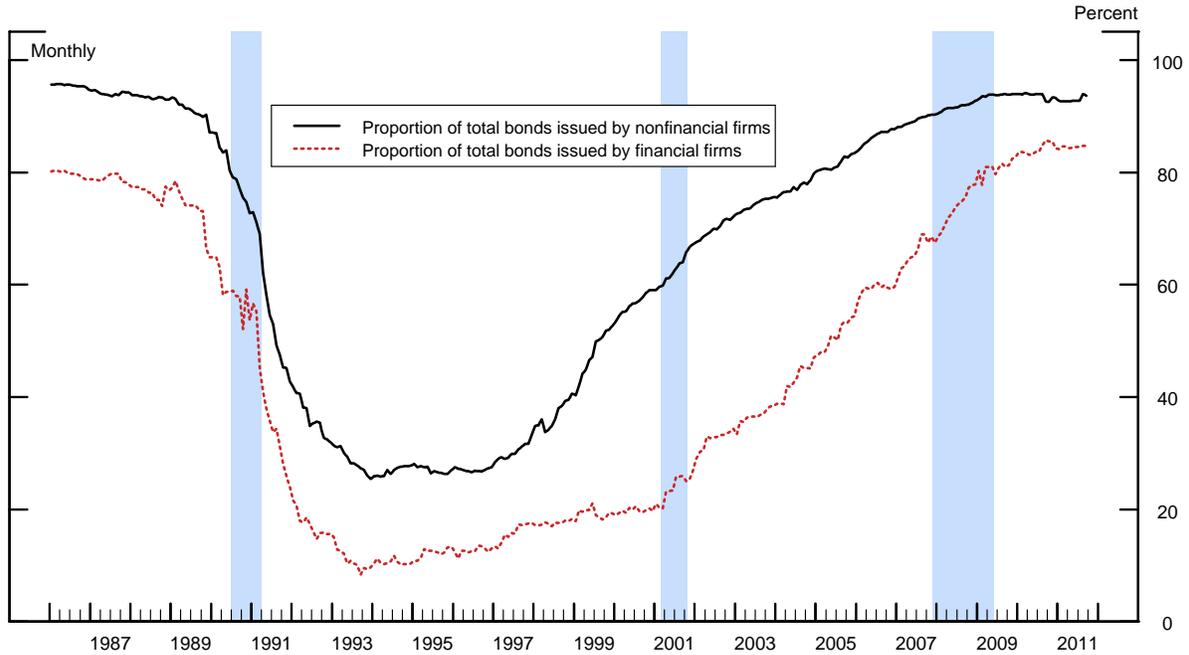
NOTE: Sample period: 1986:M1–2011:M9. The solid line depicts the weighted median DD of the 1,156 nonfinancial bond issuers in our sample; the dotted line depicts the weighted median DD of the 202 financial bond issuers. The shaded band depicts the weighted interquartile range of the DDs for the entire U.S. corporate sector; all percentiles are weighted by the firm’s outstanding liabilities. The shaded vertical bars represent the NBER-dated recessions.

Bharath and Shumway [2008]. Employing their methodology, we calculate the distance-to-default for all U.S. corporations covered by S&P’s Compustat and CRSP over the 1986:M1–2011:M9 period. Figure 1 plots the cross-sectional median of the DDs for the 1,156 nonfinancial and 202 financial bond issuers in our sample. As a point of comparison, the figure also depicts the cross-sectional interquartile range (IQR) of the DDs for the entire Compustat-CRSP matched sample.<sup>10</sup>

According to this metric, the credit quality of the median nonfinancial bond issuer in our sample is, on average, appreciably higher than that of the median financial issuer, a result that is primarily due to the fact that financial firms tend to have higher leverage than their nonfinancial counterparts. More importantly, the median DD for both sets of firms is strongly procyclical, implying that equity market participants anticipate corporate defaults to increase during economic downturns. In addition, this indicator of default risk worsened significantly in periods of financial market stress, such as those associated with the stock market crash in October 1987 and the collapse

<sup>10</sup>To ensure that our results were not driven by a small number of extreme observations, we eliminated from our sample all firm/month observations with a DD of more than 20 or less than -2, cutoffs corresponding roughly to the 99th and 1st percentiles of the DD distribution, respectively.

Figure 2: Callable Corporate Bonds



NOTE: Sample period 1986:M1–2011:M9. The figure depicts the proportion of bonds in our sample that are callable. The shaded vertical bars represent the NBER-dated recessions.

of the Long-Term Capital Management hedge fund in the early autumn of 1998. In fact, during the height of the recent financial crisis in the latter part of 2008 and early 2009, the IQR of the distribution of the DDs across the entire U.S. corporate sector shifted noticeably lower, with the median DD of our sample of 202 financial firms falling to a historic low.

### 2.3 Call-Option Adjustment

Figure 2 shows the proportion of bonds in our sample that are callable—that is, the issuer has, under certain pre-specified conditions, the right to “call” (i.e., redeem) the security prior to its maturity. The share of senior unsecured bonds with embedded call options is, on average, substantial in both sectors.<sup>11</sup> Moreover, the proportion of callable debt has changed considerably over the course of our sample period, with almost all bonds being subject to a call provision at the start of our sample. In the late 1980s, however, the composition of debt began to shift noticeably toward noncallable debt, and by the mid-1990s, the majority of senior unsecured debt traded in the secondary market

<sup>11</sup>The proportions and the U-pattern of the two series are very similar if the shares are weighted by the amount issued. The fact that financial firms, on average, tend to issue less callable debt than their nonfinancial counterparts primarily reflects the noticeably shorter average duration of bonds issued by financial intermediaries (see Table 1).

was in the form of noncallable securities. Over the past decade or so, this trend has been reversed, as firms resumed issuing large amounts of callable long-term debt.

As shown by Duffee [1998], if a firm’s outstanding bonds are callable, movements in the risk-free rates—by changing the value of the embedded call option—will have an independent effect on bond prices, complicating the interpretation of the behavior of credit spreads. For example, as the general level of interest rates in the economy increases, the option to call becomes less valuable, which accentuates the price response of callable bonds relative to that of noncallable bonds. As a result, a rise in interest rates will, *ceteris paribus*, compress the credit spreads of callable bonds more than the credit spreads of their noncallable counterparts. In addition, prices of callable bonds are more sensitive to uncertainty regarding the future course of interest rates. On the other hand, to the extent that callable bonds are, in effect, of shorter duration, they may be less sensitive to changes in default risk.

To deal with this issue, we utilize the micro-level aspect of our bond data to adjust directly for the value of embedded options in callable bonds. Ideally, we would correct for the callability of each bond using option price theory. However, our bond-level data set does not contain any information regarding the terms of the underlying call provisions—it just indicates whether the bond is callable or not. Accordingly, we consider the following empirical credit-spread pricing model:

$$\begin{aligned} \ln S_{it}[k] = & CALL_i[k] \times (\beta_{0c} + \beta_{1c}DD_{it} + \beta_{2c}DD_{it}^2 + \lambda'_c \mathbf{Z}_{it}[k]) \\ & + (1 - CALL_i[k]) \times (\beta_{0n} + \beta_{1n}DD_{it} + \beta_{2n}DD_{it}^2 + \lambda'_n \mathbf{Z}_{it}[k]) \\ & + CALL_i[k] \times (\theta_1 LEV_t + \theta_2 SLP_t + \theta_3 CRV_t + \theta_4 VOL_t) + RIG_{it}[k] + \epsilon_{it}[k], \end{aligned} \quad (1)$$

where  $CALL_i[k]$  is an indicator variable that equals one if bond  $k$  (issued by firm  $i$ ) is callable and zero otherwise;  $DD_{it}$  denotes the issuer’s distance-to-default; and  $\epsilon_{it}[k]$  represents the credit-spread “pricing error.”<sup>12</sup> In our framework, credit spreads on outstanding bonds are allowed to depend (possibly nonlinearly) on the issuer’s distance-to-default  $DD_{it}$  and a vector of bond-specific control variables  $\mathbf{Z}_{it}[k]$  that could influence credit spreads through term or liquidity premiums.<sup>13</sup> Specifically, these variables include the log of the bond’s duration, the log of the amount outstanding, the log of the bond’s age, and the log of the bond’s (fixed) coupon rate. The regression also includes credit rating fixed effects ( $RIG_{it}[k]$ ), which capture the “soft information” regarding the firm’s financial health that is complementary to our market-based measures of default risk (Löffler [2004, 2007]).

The distance-to-default and bond-specific controls are allowed to have differential effects on the credit spreads of callable and noncallable bonds. Because shifts in the Treasury term structure

<sup>12</sup>Because the distribution of credit spreads is highly skewed, log transformation of credit spreads ameliorates the problem of heteroskedasticity in the estimation of equation (1).

<sup>13</sup>Specification 1 allows for the nonlinear relationship between credit spreads and leverage documented by Levin, Natalucci, and Zakrajšek [2004]. As a robustness check, we also considered higher-order polynomials of the distance-to-default, but the inclusion of cubic and quartic terms had virtually no effect on our results.

affect the value of the embedded call option, the spreads of callable bonds are also allowed to depend separately on the level ( $LEV_t$ ), slope ( $SLP_t$ ), and curvature ( $CRV_t$ ) of the Treasury yield curve.<sup>14</sup> Note that movements in risk-free interest rates should, *ceteris paribus*, affect the credit spreads of noncallable bonds only insofar as they change firms' expected future cash flows and, as a result, their distance-to-default; thus, our specification does not allow these term structure variables to directly affect the spreads of noncallable bonds. Likewise, the value of the embedded call option will change in response to fluctuations in interest rate uncertainty, so we allow the credit spreads on callable bonds to respond to the option-implied volatility on the 30-year Treasury bond futures ( $VOL_t$ )—again, this variable is not allowed to affect the spreads of noncallable bonds.

We estimate the credit-spread regression (1) separately for the sample of securities issued by nonfinancial firms and those issued by financial firms. Assuming normally distributed pricing errors, the option-adjusted spread on a callable bond  $k$  (i.e.,  $CALL_i[k] = 1$ )—denoted by  $\tilde{S}_{it}[k]$ —is given by

$$\tilde{S}_{it}[k] = \exp \left[ \ln S_{it}[k] - (\hat{\beta}_{0c} + \hat{\beta}_{1c}DD_{it} + \hat{\beta}_{2c}DD_{it}^2 + \hat{\lambda}'_c \mathbf{Z}_{it}[k]) - (\hat{\theta}_1 LEV_t + \hat{\theta}_2 SLP_t + \hat{\theta}_3 CRV_t + \hat{\theta}_4 VOL_t) - \frac{\hat{\sigma}_t^2}{2} \right],$$

where  $\hat{\beta}_{0c}, \hat{\beta}_{1c}, \hat{\beta}_{2c}, \hat{\lambda}_c$  and  $\hat{\theta}_1, \dots, \hat{\theta}_4$  denote the OLS estimates of the corresponding parameters from equation (1) and  $\hat{\sigma}_t^2$  denotes the sample variance of the residuals in period  $t$ .<sup>15</sup> With this option adjustment, credit spreads are now constructed variables.

Table 2 translates the selected coefficients from the estimated credit-spread pricing equation into the impact of variation in default risk (the sum of the linear and quadratic  $DD$  terms), the shape of the term structure, and interest rate uncertainty on the *level* of credit spreads. For callable bonds issued by nonfinancial firms, the effect of the distance-to-default on credit spreads is significantly attenuated by the call-option mechanism: A one standard deviation increase in the distance-to-default—a signal of improving credit quality—implies a decrease of 21 basis points in the spreads of noncallable bonds, compared with a 14 basis points decline in the spreads of their callable counterparts. The same call-option mechanism, however, does not seem to be as important for bonds issued by financial intermediaries. In that case, a one standard deviation increase in the distance-to-default implies a narrowing of spreads of about 13 basis points for both types of bonds.

The estimates in Table 2 also indicate that movements in the shape of the Treasury term

---

<sup>14</sup>The level, slope, and curvature factors correspond, respectively, to the first three principal components of nominal Treasury yields at 3-month, 6-month, 1-, 2-, 3-, 5-, 7-, 10-, 15-, and 30-year maturities. All yield series are monthly (at month-end) and with the exception of the 3- and 6-month bill rates are derived from the smoothed Treasury yield curve estimated by Gürkaynak, Sack, and Wright [2007].

<sup>15</sup>The  $\hat{\sigma}_t^2/2$  term is a correction for Jensen's inequality. All of the results reported in the paper are virtually identical if the Jensen's inequality correction imposes that the conditional variance of the error term is constant over time, or if no such correction is implemented.

Table 2: Selected Marginal Effects for Callable vs. Noncallable Bonds

Marginal Effect	<i>Nonfinancial Firms</i> <sup>a</sup>		<i>Financial Firms</i> <sup>b</sup>	
	<i>CALL</i> = 0	<i>CALL</i> = 1	<i>CALL</i> = 0	<i>CALL</i> = 1
Distance-to-default: $DD_{it}$	-0.209 (0.011)	-0.136 (0.008)	-0.125 (0.028)	-0.133 (0.015)
Term structure: $LEV_t$	-	-0.473 (0.041)	-	-0.421 (0.081)
Term structure: $SLP_t$	-	-0.288 (0.037)	-	-0.218 (0.047)
Term structure: $CRV_t$	-	-0.074 (0.038)	-	-0.123 (0.034)
Term structure: $VOL_t$	-	0.144 (0.013)	-	0.147 (0.017)
Adjusted $R^2$	0.746		0.615	
Pr > $W^c$	0.000		0.000	

NOTE: Sample period: 1986:M1–2011:M9. Entries in the table denote the estimated marginal effects of a one-unit change in the specified variable on the level of credit spreads (in percentage points) for noncallable ( $CALL = 0$ ) and callable ( $CALL = 1$ ) bonds. The estimates of marginal effects are based on the OLS estimates of parameters of the credit-spread regression (1). All marginal effects are evaluated at their respective sample means (not reported). Robust asymptotic standard errors reported in parentheses are double clustered in the firm ( $i$ ) and time ( $t$ ) dimensions (see Cameron, Gelbach, and Miller [2011] for details) and are computed according to the delta method.

<sup>a</sup> No. of firms/bonds = 1,156/6,404; Obs. = 334,685.

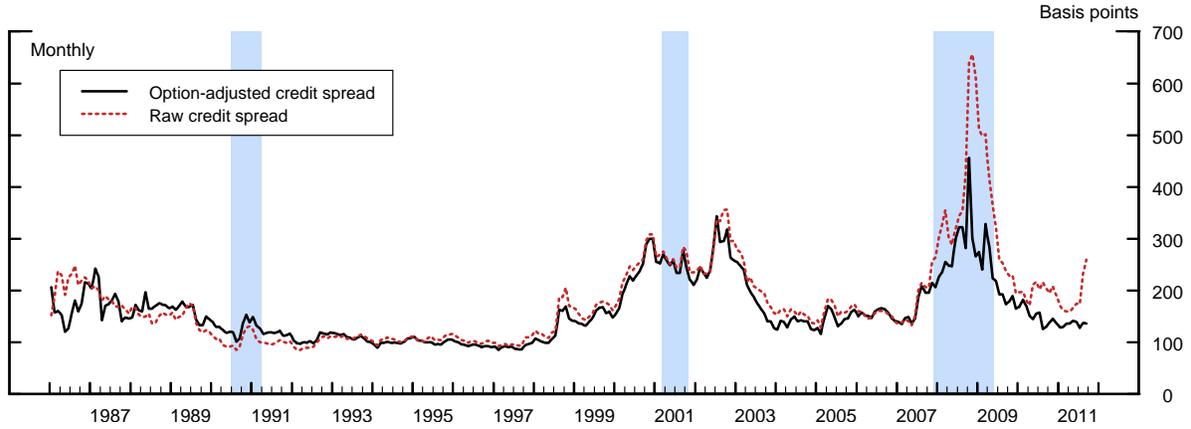
<sup>b</sup> No. of firms/bonds = 202/942; Obs. = 46,135.

<sup>c</sup>  $p$ -value for the robust Wald test of the exclusion of credit rating fixed effects.

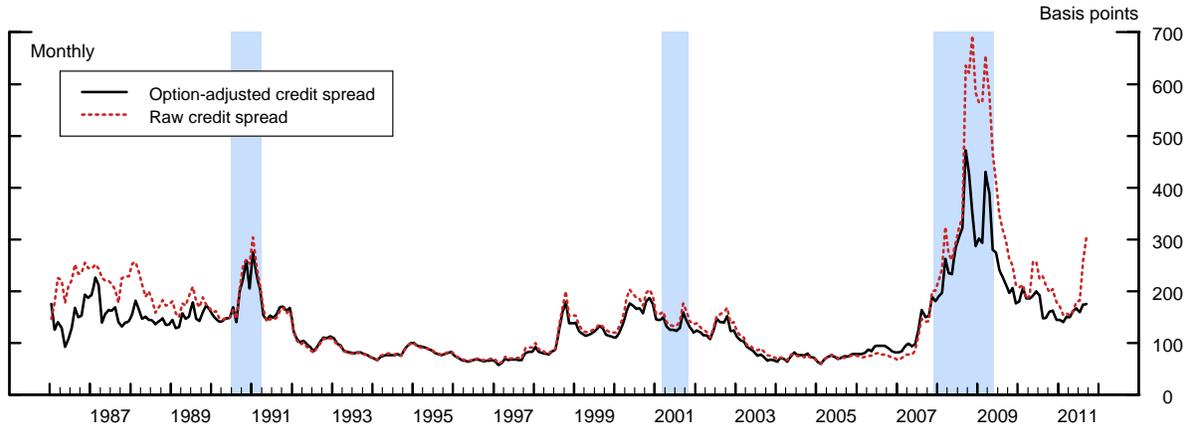
structure and interest rate uncertainty have first-order effects on the credit spreads of callable bonds, which are consistent with the theoretical predictions. For example, a one standard deviation increase in the level factor implies a narrowing of about 45 basis points in the credit spreads on callable bonds in both sectors. Similarly, an increase in the option-implied volatility on the long-term Treasury bond futures of one percentage point implies a widening of callable credit spreads of about 15 basis points because the rise in interest rate uncertainty lowers the prices of callable bonds by boosting the value of the embedded call option.

The importance of the option-adjustment procedure over the entire sample period is illustrated in Figure 3, which shows the time path of the average credit spread in our two data sets, calculated using both the raw and option-adjusted spreads. Although the two series in each sector are clearly highly correlated ( $\rho = 0.90$  for nonfinancial issuers and  $\rho = 0.95$  for financial issuers) and are all strongly countercyclical, there are a number of noticeable differences. First, the option-adjusted credit spreads are, on average, lower than their unadjusted counterparts, reflecting the positive value of the embedded call options. By eliminating, at least in part, fluctuations in the call option values,

Figure 3: Credit Spreads on Corporate Bonds by Type of Firm



(a) Nonfinancial Firms



(b) Financial Firms

NOTE: Sample period: 1986:M1–2011:M9. The solid line in each panel depicts the time-series of the weighted cross-sectional average of the option-adjusted credit spreads for our sample of bonds (see text for details); the dotted line depicts the time-series of the weighted cross-sectional average of the raw credit spreads. In all cases, the weights are equal to the market values of the underlying bond issues. The shaded vertical bars represent the NBER-dated recessions.

the option-adjusted credit spreads are also less volatile, on average, than the raw credit spreads. Lastly, the largest differences between the two series occurred in the mid-1980s and during the recent financial crisis. The former period was characterized by a high general level of interest rates and relatively high uncertainty regarding the future course of long-term interest rates, whereas the difference during the latter period owes primarily to the plunge in interest rates and the steepening of the term structure that began with the onset of the financial crisis in the summer of 2007, two factors that more than offset the spike in interest rate volatility that occurred during that period.

## 2.4 Distance-to-Default Portfolios

We summarize the information contained in credit spreads, DDs, and excess equity returns for the sample of bond issuers by constructing portfolios based on expected default risk—as measured by our estimate of the distance-to-default—at the beginning of the period. These conditional DD-based portfolios are constructed by sorting the three financial indicators in month  $t$  into bins based on the percentiles of the distribution of the distance-to-default in month  $t - 1$ . Separate portfolios are formed for the financial and nonfinancial issuers.

The distance-to-default portfolios are constructed by computing a weighted average of DDs in month  $t$  for each bin, with the weights equal to the book value of the firm’s liabilities at the end of month  $t - 1$ . Similarly, the stock portfolios are computed as a weighted average of excess equity returns in month  $t$  for each bin, with the weights equal to the market value of the firm’s equity at the end of month  $t - 1$ .<sup>16</sup> Given the relatively large number of nonfinancial issuers, the bins for nonfinancial portfolios are based on the quartiles of the DD distribution, yielding four credit-risk categories, denoted by NFIN-DD1, NFIN-DD2, NFIN-DD3, and NFIN-DD4. The financial bond issuers, by contrast, are sorted into two credit-risk categories—denoted by FIN-DD1 and FIN-DD2—based on the median of the DD distribution.

To control for maturity, we further split each DD-based bin of nonfinancial credit spreads into four maturity categories: (1) NFIN-MTY1: credit spreads of bonds with the remaining term-to-maturity of more than 1 year but less than (or equal) to 5 years; (2) NFIN-MTY2: credit spreads of bonds with the remaining term-to-maturity of more than 5 years but less than (or equal) 10 years; (3) NFIN-MTY3: credit spreads of bonds with the remaining term-to-maturity of more than 10 years but less than (or equal) to 15 years; (4) NFIN-MTY4: credit spreads of bonds with the remaining term-to-maturity of more than 15 years. Given the substantially smaller sample of bonds issued by firms in the financial sector, we split the two credit-risk categories in this sector into two maturity categories: (1) FIN-MTY1: credit spreads of bonds with the remaining term-to-maturity of more than 1 year but less than (or equal) to 5 years; and (2) FIN-MTY2: credit spreads of bonds with the remaining term-to-maturity of more than 5 years. All told, this gives us a total of 16 nonfinancial and 4 financial DD/maturity bond portfolios. Within each of these portfolios, we compute a weighted average of option-adjusted credit spreads in month  $t$ , with the weights equal to the market value of the outstanding issue.

The DD-based portfolios considered thus far were based on asset prices of a subset of U.S. corporations, namely firms with senior unsecured bonds that are traded in the secondary market. We also consider a broader set of DD-based financial indicators by constructing the same type of portfolios using the distance-to-default estimates and excess equity returns for the entire matched CRSP-Compustat sample of U.S. corporations. Given the large number of firms in any given month,

---

<sup>16</sup>Excess equity returns, which include dividends and capital gains, are measured relative to the yield on 1-month Treasury bills.

we increase the number of bins by sorting—for both nonfinancial and financial firms separately—the DDs and excess equity returns in month  $t$  into 10 deciles based on the distribution of the distance-to-default in month  $t - 1$ . As before, the conditional DD portfolios are constructed by computing a weighted average of DDs in month  $t$  for each DD decile, whereas the stock portfolios are computed as a weighted average of excess equity returns in month  $t$ . This procedure yields a total of 20 additional DD-based portfolios for the nonfinancial sector and another 20 portfolios for the financial sector.<sup>17</sup>

### 3 Econometric Methodology

We examine the predictive content of the DD-based portfolios, as well as a large number of other predictors, within the Bayesian Model Averaging (BMA) framework, an approach that is particularly well-suited to deal with model uncertainty. Initially proposed by Leamer [1978], BMA has been used extensively in the statistics literature; see, for example, Raftery, Madigan, and Hoeting [1997] and Chipman, George, and McCulloch [2001]. The BMA approach to model uncertainty has also found numerous econometric applications, including the forecasting of output growth (Min and Zellner [1993] and Koop and Potter [2004]); the forecasting of recession risk (King, Levin, and Perli [2007]); cross-country growth regressions (Fernandez, Ley, and Steel [2001b] and Sala-i-Martin, Doppelhofer, and Miller [2004]); exchange rate forecasting (Wright [2008]); and the predictability of stock returns (Avramov [2002] and Cremers [2002]).

#### 3.1 Bayesian Model Averaging

We begin with a brief review of the formal Bayesian justification for our model-averaging approach. The researcher starts with a set of  $n$  possible models, where the  $i$ -th model, denoted by  $M_i$ , is parametrized by  $\theta_i$ . The researcher has prior beliefs about the probability that the  $i$ -th model is true—denoted by  $P(M_i)$ —observes data  $D$ , and updates her beliefs to compute the posterior probability that the  $i$ -th model is the true model, according to

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{\sum_{j=1}^n P(D|M_j)P(M_j)}, \quad (2)$$

where

$$P(D|M_i) = \int P(D|\theta_i, M_i)P(\theta_i|M_i)d\theta_i \quad (3)$$

---

<sup>17</sup>Note that the DDs and excess equity returns for our sample of bond issuers are also included in these additional DD-based portfolios. Although firms that issue bonds tend to be somewhat bigger than a typical publicly-traded corporation, their inclusion has a very small effect on the time-series properties of these more-representative portfolios. For example, the correlations between the stock and DD portfolios that include bonds issuers and those that do not range from 0.85 to 0.98—depending on the decile and the type of portfolio—suggesting that any information content of these broader portfolios is not unduly influenced by the inclusion of firms that have bonds traded in the secondary market.

is the marginal likelihood of the  $i$ -th model;  $P(\theta_i|M_i)$  is the prior density of the parameter vector  $\theta_i$  associated with the  $i$ -th model; and  $P(D|\theta_i, M_i)$  is the likelihood function.

Each model also implies a forecast. In the presence of model uncertainty, the BMA forecast weights each of the individual forecasts by their respective posterior probabilities. To operationalize a BMA forecasting scheme, the researcher needs only to specify the set of models, the model priors  $P(M_i)$ , and the parameter priors  $P(\theta_i|M_i)$ . In this paper, we follow a growing literature that considers a large set of very simple models. Specifically, the models are all linear regression models, with each model adding a single regressor to the baseline specification. More formally, the  $i$ -th model in our framework is given by

$$y_{t+h} = \beta_i X_{it} + \gamma_i' Z_t + \epsilon_{t+h}, \quad (4)$$

where  $y_{t+h}$  is the variable that the researcher wishes to forecast at a horizon of  $h$  periods;  $X_{it}$  is the (scalar) predictor specific to model  $i$ ;  $Z_t$  is a  $(p \times 1)$ -vector of predictors that are common to all models; and  $\epsilon_{t+h} \stackrel{iid}{\sim} N(0, \sigma_i^2)$  is the forecast error. Without loss of generality, the model-specific predictor  $X_{it}$  is assumed to be orthogonal to the common predictors  $Z_t$ . In our setup, the vector of parameters characterizing the  $i$ -th model is thus given by  $\theta_i = [\beta_i \ \gamma_i' \ \sigma_i^2]'$ . The  $n$ -th model, which is included in every BMA application in this paper, imposes the restriction  $\beta_i = 0$  and so has  $\theta_i = [\gamma_i' \ \sigma_i^2]'$ .

In setting the model priors, we assume that all models are equally likely—that is,  $P(M_i) = 1/n$ . For the parameter priors, we follow the general trend of the BMA literature (e.g., Fernandez, Ley, and Steel [2001a]) in specifying that the prior for  $\gamma_i$  and  $\sigma_i^2$ , denoted by  $p(\gamma_i, \sigma_i)$ , is uninformative and is proportional to  $1/\sigma_i$ , for all  $i$ , while using the  $g$ -prior specification of Zellner [1986] for  $\beta_i$  conditional on  $\sigma_i^2$ . The  $g$ -prior is given by  $N(0, \phi \sigma_i^2 (X_i' X_i)^{-1})$ , where the shrinkage hyperparameter  $\phi > 0$  measures the strength of the prior—a smaller value of  $\phi$  corresponds to a more dogmatic prior.

Letting  $\hat{\beta}_i$  and  $\hat{\gamma}_i$  denote the OLS estimates of the corresponding parameters in equation (4), the Bayesian  $h$ -period-ahead forecast made from model  $M_i$  at time  $T$  is given by

$$\tilde{y}_{T+h|T}^i = \tilde{\beta}_i X_{it} + \hat{\gamma}_i' Z_t, \quad (5)$$

where  $\tilde{\beta}_i = \left(\frac{\phi}{\phi+1}\right) \hat{\beta}_i$  denotes the posterior mean of  $\beta_i$ . In our framework, the marginal likelihood of the  $i$ -th model reduces to

$$P(D|M_i) \propto \left[\frac{1}{1+\phi}\right]^{-\frac{1}{2}} \times \left[\frac{1}{1+\phi} SSR_i + \frac{\phi}{1+\phi} SSE_i\right]^{-\frac{(T-p)}{2}}, \quad (6)$$

where  $SSR_i$  is the sum of squares from the  $i$ -th regression, and  $SSE_i$  is the associated sum of squared errors. The posterior probabilities of the models can then be worked out from equation (2), and

the final BMA forecast that takes into account model uncertainty is given by

$$\tilde{y}_{T+h|T} = \sum_{i=1}^n P(M_i|D)\tilde{y}_{T+h|T}^i. \quad (7)$$

Clearly, the BMA forecast in equation (7) will depend on the value of the shrinkage hyperparameter  $\phi$ . A low value of  $\phi$  implies that the model likelihoods are roughly equal, and so the BMA forecast will resemble equal-weighted model averaging (Bates and Granger [1969]). In contrast, a high value of  $\phi$  amounts to weighting the models by their in-sample  $R^2$  values, a procedure that is well known to generate poor out-of-sample forecasting performance. Because the relationship between the out-of-sample root mean square prediction error and the parameter  $\phi$  is often U-shaped, the best out-of-sample forecasts are obtained when  $\phi$  is neither too small nor too big. Our baseline results are based on a standard value ( $\phi = 4$ ) taken from the aforementioned literature, but we also conduct sensitivity analysis, which shows that our results are robust with respect to this choice.

We apply BMA to forecasting various indicators of economic activity using standard macroeconomic variables and asset market indicators as predictors. The common predictors  $Z_t$  in the predictive regression (4) are a constant and lags of the dependent variable. It is worth emphasizing that we view the forecasting scheme proposed above as a pragmatic approach to data-based weighting of models and make no claim to its Bayesian optimality properties.<sup>18</sup>

### 3.2 The Forecasting Setup

We focus on forecasting real economic activity, as measured by real GDP, real personal consumption expenditures (PCE), real business fixed investment, industrial production, private payroll employment, the civilian unemployment rate, real exports, and real imports over the period from 1986:Q1 to 2011:Q3. All of these series are in quarter-over-quarter growth rates (actually 400 times log first-differences), except for the unemployment rate, which is simply in first differences. Our objective is to forecast the cumulative growth rate (or the cumulative change in the case of the unemployment rate) for each of these economic indicators from quarter  $t - 1$  through quarter  $t + h$ .

Specifically, let  $y_t$  denote the growth rate in the variable from quarter  $t - 1$  to quarter  $t$ . (In the case of the unemployment rate,  $y_t$  denotes the first difference.) The average value of  $y_t$  over the forecast horizon  $h$  is denoted by  $y_{t+h}^C = \frac{1}{h+1} \sum_{i=0}^h y_{t+i}$ . The  $i$ -th forecasting model in our setup

---

<sup>18</sup>As noted by a number of papers that employ the same data-based model averaging approach, several of the conditions for strict optimality are not met in typical macro time-series applications. First, the regressors are assumed to be strictly exogenous. And second, the forecasts are overlapping  $h$ -step ahead forecasts, so the forecast errors less than  $h$  periods apart are bound to be serially correlated, even though it is assumed that they are i.i.d. normal. Nevertheless, BMA, like other methods that combine a large number of predictors to generate a forecast, may still have good forecasting properties, even if the premises underlying their theoretical justification are false (e.g., Stock and Watson [2005]). In fact, ability to provide accurate out-of-sample forecasts is a stringent test of the practical usefulness of BMA in forecasting.

is given by:

$$y_{t+h}^C = \alpha + \beta_i x_{it-1} + \sum_{j=1}^{p-1} \gamma_{ij} y_{t-j} + \epsilon_{t+h}, \quad (8)$$

where  $x_{it}$  is one of the predictors listed in Table 3 and  $p - 1$ , the number of lags, is determined recursively by the Bayes Information Criterion (BIC), conditional on  $\beta_i = 0$ .

The set of possible predictors listed in Table 3 includes 15 different real-time macroeconomic series and 119 asset market indicators. Among the latter set are our 20 bond portfolios of option-adjusted credit spreads, as well as average DDs and excess equity returns for different default-risk portfolios; in addition, we consider the predictive content of the three Fama-French risk factors (i.e., the excess market return and the SMB and HML factors), stock return on industry portfolios, a range of standard interest rates and interest rate spreads, implied volatilities from options quotes, commodity prices, and commonly-used credit spreads.<sup>19</sup> As noted above, the set of models considered also includes a model in which we restrict  $\beta_i = 0$ . All told, our BMA forecasting scheme includes 135 different models.

The timing convention in the forecasting regression (8) is as follows. We think of forecasts as being made in the middle month of each quarter. For macroeconomic variables, we use the February, May, August, and November vintages of data from the real-time data set compiled and maintained by the Federal Reserve Bank of Philadelphia; this includes data through the previous quarter for all the macroeconomic series that we consider. All asset market indicators are as of the end of the month from the *first* month of the current quarter and would have been available to forecasters as of the middle month of the quarter.

Importantly, our empirical option-adjustment procedure is also implemented in real-time—that is, the parameters of the credit-spread regression (1) are estimated each month using only data available at that time. The resulting real-time coefficient estimates are used to compute the option-adjusted credit spreads, which are then sorted into the DD-based bond portfolios.<sup>20</sup> With these fully real-time data in hand, we then use BMA to construct forecasts of the values of the dependent variable for the current and next four quarters (i.e.,  $h = 0, 1, \dots, 4$ ). Thus, we are considering both “nowcasting” (i.e.,  $h = 0$ ) and prediction at horizons up to one year ahead.

We evaluate the accuracy of these BMA forecasts in a recursive out-of-sample forecast evaluation exercise, starting with the forecasts made in 1992:Q1 and continuing through to the end of the sample period in 2011:Q3. The implementation of BMA in this recursive forecasting scheme uses the  $g$ -prior (see the discussion above) with the same shrinkage hyperparameter  $\phi$  for each date. This means that the prior is tighter for the forecasts that are made later in the forecast period—the data,

---

<sup>19</sup>The three risk factors and returns on industry portfolios were obtained from the Kenneth R. French’s website (<http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>).

<sup>20</sup>Note that the real-time implementation of the option-adjustment procedure generates spreads that differ from the option-adjusted spreads underlying Figure 3, where the option-adjustment procedure was implemented using the full data set.

Table 3: Predictors Used in BMA Forecasts

Predictor (No. of series)	Data Transformation
<i>Macroeconomic Indicators</i> (15)	
GDP	log difference
PCE	log difference
PCE (durable goods)	log difference
Residential investment	log difference
Business fixed investment	log difference
Government spending	log difference
Exports	log difference
Imports	log difference
Nonfarm private payrolls	log difference
Civilian unemployment rate	difference
Industrial production	log difference
Single-family housing starts	log difference
GDP price deflator	log difference
Consumer price index	log difference
M2	log difference
<i>Asset Market Indicators</i> (119)	
Credit spreads in DD-based bond portfolios (nonfinancial) (16)	level
Credit spreads in DD-based bond portfolios (financial) (4)	level
Avg. DD by DD percentile (nonfinancial bond issuers) (4)	level
Avg. DD by DD percentile (nonfinancial firms) (10)	level
Excess stock returns by DD percentile (nonfinancial bond issuers) (4)	level
Excess stock returns by DD percentile (nonfinancial firms) (10)	level
Avg. DD by DD percentile (financial bond issuers) (2)	level
Avg. DD by DD percentile (financial firms) (10)	level
Excess stock returns by DD percentile (financial bond issuers) (2)	level
Excess stock returns by DD percentile (financial firms) (10)	level
3-month nonfinancial commercial paper rate	level
3-month nonfinancial commercial paper rate	less 3-month Tbill rate
3-month Eurodollar rate	level
3-month Eurodollar rate	less 3-month Tbill rate
3-month Treasury bill rate	level
Federal funds rate	level
1- to 10-year Treasury yields (10) <sup>a</sup>	level
1- to 10-year Treasury yields (10)	less 3-month Tbill rate
Fama-French risk factors (3)	level
S&P 100 futures implied volatility (VXO)	level
Treasury futures implied volatility (10- and 30-year)	level
Gold price	2nd difference of logs
Oil price	2nd difference of logs
CRB commodity price index	2nd difference of logs
S&P 500 dividend yield	log
Moody's Baa-Aaa credit spread	level
Returns on industry portfolios (10)	level

NOTE: All macroeconomic series come from the real-time data set maintained by the Federal Reserve Bank of Philadelphia. The NIPA series are in real terms (c-w, \$2000).

<sup>a</sup> The nominal Treasury yields between maturities of 1- and 10-years are taken from the Treasury yield curve estimated by Gürkaynak, Sack, and Wright [2007].

of course, are also more informative at that point—a choice that is consistent with what a researcher would have done, if she was applying BMA as a shrinkage device at each point in time.

An important issue in this type of real-time forecasting exercise is the definition of what constitutes the “actual” values with which to compare our forecasts. The macroeconomic indicators that we are forecasting are subject to benchmark revisions, and some of the series are also subject to definitional and conceptual changes. None of these changes seem sensible to predict in a real-time forecasting exercise. Accordingly, we follow a standard convention (e.g., Tulip [2005] and Faust and Wright [2009]), which is to measure actual realized values from the data as recorded in the real-time data set of the Philadelphia Fed two quarters after the quarter to which the data refer. In general, this implies that we are comparing our forecasts to the “first final” vintage in series from the National Income and Product Accounts (NIPA).

### 3.3 Inference

The accuracy of the BMA forecasts is evaluated by comparing the RMSPE (root mean-square prediction error) of the BMA forecast to that obtained from a univariate autoregression:<sup>21</sup>

$$y_{t+h}^C = \alpha + \sum_{j=1}^{p-1} \gamma_j y_{t-j} + \epsilon_{t+h}. \quad (9)$$

Unfortunately, evaluating the statistical significance of the difference in RMSPEs from BMA and the direct autoregression is complicated by the fact that the forecasts are generated by nested models. As shown by Clark and McCracken [2001], the distribution of the Diebold and Mariano [1995] test statistic under the null hypothesis of equal forecast accuracy has a nonstandard distribution in this case. Accordingly, we use a bootstrap to approximate the limiting distribution of the Diebold-Mariano statistic under the null hypothesis. In the bootstrap, the predictors are, by construction, irrelevant—nevertheless, they have time-series and cross-sectional dependence properties that are designed to mimic those of the underlying data. The resulting bootstrapped  $p$ -values hence allow us to test the null hypothesis of no improvement in forecast accuracy, relative to the autoregressive benchmark.

The specific bootstrap re-sampling scheme used follows Gonçalves and Perron [2011] and Clark and McCracken [2012]. Specifically, we estimate two models: (1) a *restricted* model that involves estimating an AR(4) process for  $y_t$ ; and (2) an *unrestricted* model that consists of a regression of  $y_t$  on four lags of itself and the first three principal components of the entire predictor set. In each bootstrap replication, we then re-sample the residuals of the unrestricted model using a wild bootstrap and then construct a bootstrap sample of  $y_t$  using these re-sampled residuals, together

---

<sup>21</sup>Note that this is a direct autoregression that projects  $y_{t+h}^C$  onto  $p - 1$  lags of  $y_t$ . An alternative would be to estimate an AR( $p - 1$ ) model for  $y_t$  and then iterate it forward to construct the forecasts. This approach yielded very similar results.

with the coefficients from the restricted model; see Clark and McCracken [2012] for details. The predictor set meanwhile, is held fixed, implying that the predictors are, by construction, irrelevant for the forecasting of the dependent variable in all samples. As implemented, this bootstrap preserves any conditional heteroskedasticity in the data.<sup>22</sup>

## 4 Results

Table 4 contains the relative out-of-sample RMSPEs of the BMA forecasts, based on the benchmark value of the shrinkage hyperparameter  $\phi = 4$ . Bootstrapped  $p$ -values testing the null hypothesis that the relative RMSPE is equal to one are shown in brackets. For real GDP growth, the RMSPEs from the BMA forecasts, relative to those from the direct autoregression, are around 0.9 at all forecast horizons beyond the current quarter. As evidenced by the the associated  $p$ -values, these economically meaningful improvements in forecast accuracy are all statistically significant, at least at the 5 percent level.

The relative accuracy of BMA in forecasting output growth appears to reflect, in part, its ability to predict the growth of business fixed investment. In addition, BMA also does well in forecasting the external dimension of U.S. economic performance, namely the growth of both exports and imports. Personal consumption expenditures, in contrast, are considerably less predictable. Although BMA is noticeably more accurate than the direct autoregression in forecasting consumption growth over the very near term, the relative RMSPEs are statistically indistinguishable from one at the two- to four-quarter-ahead horizons. This result may reflect the fact that optimal consumption smoothing generally implies very little predictability of consumption growth, even if other more cyclical components of aggregate demand are forecastable.

Our BMA setup also implies economically and statistically significant gains in accuracy when predicting the growth of industrial production and changes in labor market conditions at both the near- and longer-term forecast horizons. In the case of industrial production, the relative RMSPEs lie between 0.9 and 0.96, improvements that are borderline statistically significant. The relative RMSPEs in the case of employment growth and changes in the unemployment rate are mostly around 0.88, values that are all significantly below one at a 5 percent significance level.<sup>23</sup>

Note that the benchmark to which we compared the accuracy of our BMA forecasts is a simple

---

<sup>22</sup>As a robustness check, we also computed  $p$ -values using a different bootstrap re-sampling scheme, which involved fitting an AR model for  $y_t$  and a dynamic factor model to the predictors. We constructed bootstrap samples of  $y_t$  and the predictors separately, so that the predictors were again irrelevant for forecasting the dependent variable. The exact details and the actual  $p$ -values for this alternative re-sampling scheme are available in Appendix E. All told, the two sets of bootstrap  $p$ -values are quite close to each other.

<sup>23</sup>In light of the gains in predictive accuracy documented in Table 4, a natural question to ask is whether these gains apply to measures of real economic activity only, or are also evident when forecasting inflation. We investigated this question for inflation as measured by the GDP deflator and the CPI (the results of this exercise are reported in Appendix G). It turns out that the BMA forecasts of inflation do not outperform the univariate autoregression benchmark, a result that suggests that movements in credit spreads provide more accurate and timely signals of changes of future real economic activity, rather than of prices.

Table 4: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(*Predictor Set: All Variables*)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.96 [0.03]	0.90 [0.01]	0.87 [0.01]	0.88 [0.02]	0.87 [0.02]
Personal consumption expenditures	0.88 [0.01]	0.90 [0.02]	0.97 [0.12]	1.03 [0.28]	1.06 [0.35]
Business fixed investment	0.94 [0.01]	0.84 [0.00]	0.90 [0.01]	0.90 [0.02]	0.88 [0.01]
Industrial production	0.96 [0.02]	0.92 [0.02]	0.93 [0.04]	0.92 [0.05]	0.91 [0.05]
Private employment	0.92 [0.00]	0.86 [0.00]	0.90 [0.02]	0.90 [0.02]	0.88 [0.01]
Unemployment rate	0.95 [0.01]	0.87 [0.00]	0.86 [0.00]	0.87 [0.01]	0.88 [0.02]
Exports	0.98 [0.00]	0.95 [0.00]	0.93 [0.00]	0.98 [0.03]	1.00 [0.10]
Imports	0.94 [0.00]	0.92 [0.00]	0.93 [0.02]	0.93 [0.04]	0.94 [0.05]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see text for details).

univariate autoregression. Another, even simpler, possible benchmark is given by a constant forecast, which is equal to the real-time sample average of the variable being predicted at all horizons. As it turns out, for sample periods dominated by the “Great Moderation,” a constant forecast is a relatively good benchmark because the macroeconomic indicators considered in our exercise seem to have contained little forecastable variation during this period. The counterpart to Table 4 using this alternative benchmark is in Appendix D. The upshot of those results is that our BMA forecasts of cyclically-sensitive indicators of economic activity still deliver economically and statistically significant reductions in RMSPEs relative to the constant forecast benchmark.

Overall, our first set of results indicates that for forecasting a range of real economic activity indicators, BMA—with (option-adjusted) portfolio credit spreads in the set of predictors—yields improvements relative to the univariate autoregressive benchmark that are both economically and statistically significant. The gains in forecasting accuracy are most pronounced for cyclically-sensitive indicators of economic activity, such as the growth of business fixed investment and in-

Table 5: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(Predictor Set: All Variables Except Credit Spreads in DD-Based Portfolios)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.98 [0.11]	0.98 [0.13]	0.98 [0.14]	0.99 [0.15]	0.99 [0.14]
Personal consumption expenditures	0.97 [0.14]	0.96 [0.11]	0.99 [0.21]	1.03 [0.32]	1.06 [0.40]
Business fixed investment	0.95 [0.01]	0.95 [0.04]	0.95 [0.05]	0.97 [0.09]	0.94 [0.05]
Industrial production	0.99 [0.09]	1.02 [0.52]	1.05 [0.67]	1.05 [0.53]	1.04 [0.33]
Private employment	0.98 [0.07]	1.00 [0.26]	1.05 [0.58]	1.06 [0.51]	1.03 [0.24]
Unemployment rate	0.97 [0.01]	0.97 [0.03]	1.02 [0.36]	1.05 [0.51]	1.04 [0.31]
Exports	0.97 [0.00]	1.02 [0.61]	1.02 [0.32]	1.02 [0.27]	1.01 [0.19]
Imports	0.96 [0.00]	0.97 [0.04]	1.01 [0.16]	1.04 [0.28]	1.04 [0.26]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see text for details).

dustrial production, as well as for indicators measuring changes in labor market conditions.<sup>24</sup>

To gauge more precisely the information content of credit spreads in predicting economic activity, we repeat the above analysis, except that we exclude the 20 models that utilize the credit spreads in the DD-based bond portfolios from the pool of prediction models. As shown in Table 5, very few of the entries are less than 0.97, and, especially at longer forecast horizons, most relative RMSPEs are greater than one. This finding is consistent with the standard result that a majority of forecasting methods perform about as well as a univariate autoregression. These results also illustrate a sense of how the information content of our portfolio-based credit spread indexes differs from that of the other real and asset market indicators in the predictor set: When assigning the weight to a predictor using only information available at the time of the forecast, the BMA method

<sup>24</sup>As an additional robustness check, we also considered other methods for forecasting in a data-rich environment, including a factor-augmented autoregression and an equally-weighted average of OLS-based forecasts. In general, our BMA forecasting scheme outperformed these methods.

Table 6: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
*(Predictors: Credit Spreads in DD-Based Portfolios Only)*

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.94 [0.00]	0.91 [0.00]	0.90 [0.01]	0.94 [0.02]	0.97 [0.06]
Personal consumption expenditures	0.87 [0.00]	0.86 [0.00]	0.95 [0.07]	0.98 [0.13]	1.04 [0.30]
Business fixed investment	0.92 [0.00]	0.83 [0.00]	0.90 [0.01]	0.90 [0.02]	0.91 [0.02]
Industrial production	0.94 [0.01]	0.91 [0.01]	0.94 [0.05]	0.94 [0.07]	0.92 [0.07]
Private employment	0.91 [0.00]	0.86 [0.00]	0.90 [0.00]	0.89 [0.01]	0.86 [0.01]
Unemployment rate	0.91 [0.00]	0.86 [0.00]	0.85 [0.00]	0.86 [0.00]	0.87 [0.00]
Exports	0.98 [0.00]	0.97 [0.00]	0.99 [0.02]	1.05 [0.36]	1.10 [0.56]
Imports	0.93 [0.00]	0.89 [0.00]	0.90 [0.01]	0.92 [0.05]	0.93 [0.08]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see text for details).

singles out the portfolio-based credit spreads and is able to exploit their predictive ability for future economic activity to improve significantly upon the benchmark forecast.

Another way to highlight the predictive ability of credit spreads is shown in Table 6, which contains the results of the forecasting exercise based *only* on models that include portfolio credit spreads as predictors. These results are very similar to those reported in Table 4, which utilize the information content of the entire predictor set. Although restricting the predictor set to only DD-based portfolios of credit spreads leads to some loss of predictive accuracy for real GDP growth, it actually improves the accuracy of the BMA forecasts for labor market indicators and business fixed investment. Because the autoregressive benchmark is embedded in all of these forecasting exercises, the results in Tables 5–6 together imply that any forecasting gains over the univariate autoregression are due predominantly to the information content of credit spreads in our DD-based portfolios.

## 4.1 Which Predictors are the Most Informative?

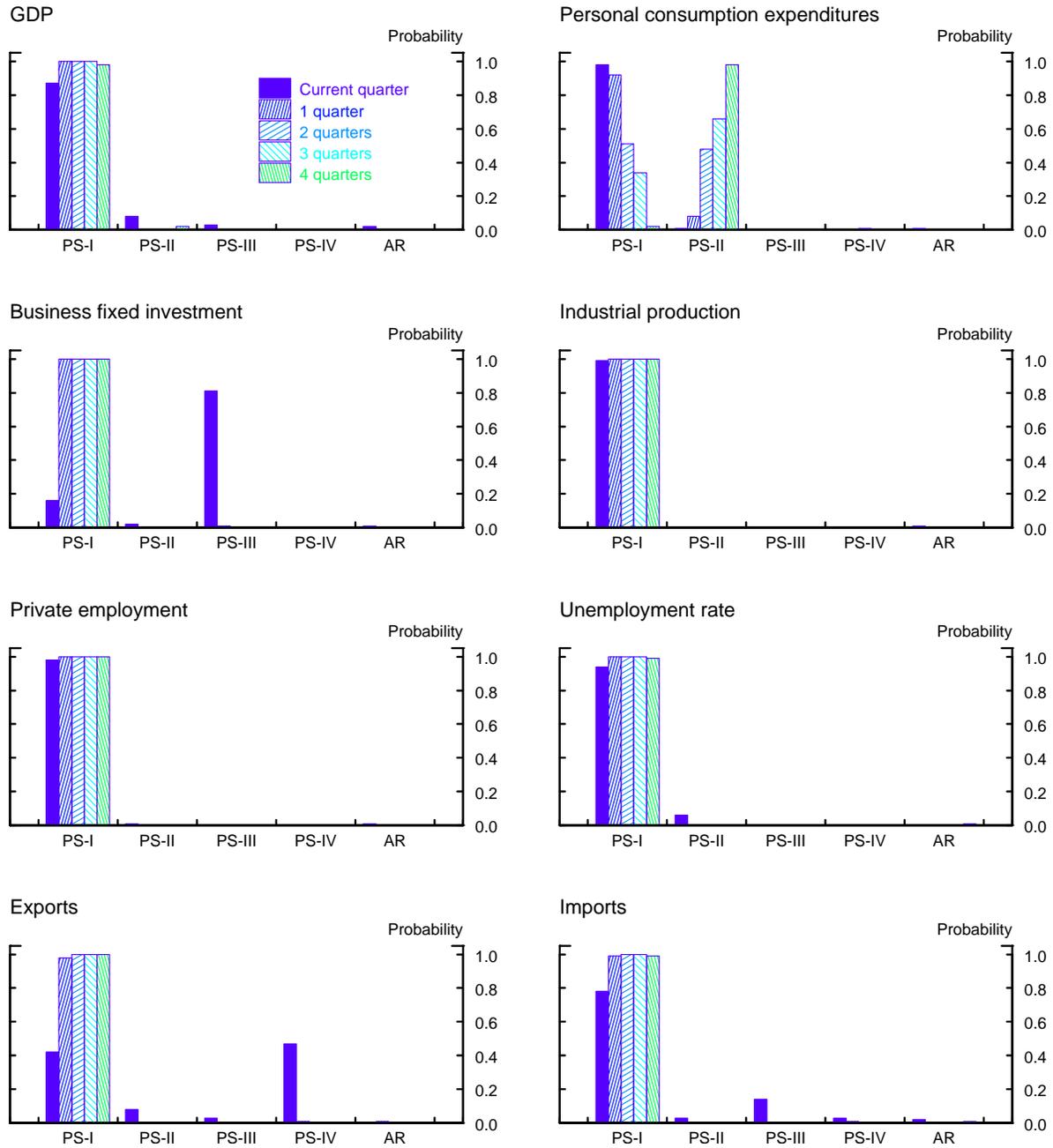
The vertical bars in the panels of Figure 4 depict the final total weights—that is, the sum of posterior probabilities—that BMA assigns to variables in the following predictor subsets: PS-I = option-adjusted credit spreads in the 20 DD-based bond portfolios; PS-II = macroeconomic variables; PS-III = other interest rates and interest rate spreads; and PS-IV = all other asset market indicators. Lastly, the AR bucket in each panel represent the final posterior probabilities that BMA assigned to the autoregressive benchmark. Results are shown for all the forecast horizons considered and for each of the eight different indicators of economic activity. Note that, by construction, these probabilities sum up to one at each forecast horizon.

These results provide a visual confirmation of the information content of the option-adjusted credit spreads in our DD-based bond portfolios. With the exception of consumption growth, BMA assigns the vast majority of the posterior weight to credit spreads in the DD-based portfolios. But even in that case, most of the posterior weight for the near-term forecasts of the growth in real PCE (i.e.,  $h = 0, 1, 2$ ) is assigned to the portfolio credit spreads; at longer horizons (i.e.,  $h = 3, 4$ ), BMA forecasts of consumption growth assign some weight to the macroeconomic variables, but the accuracy of these forecasts is, according to Table 4, statistically indistinguishable from those made by the benchmark autoregression.

It should be emphasized, however, that Figure 4 shows the posterior probabilities for the different subsets of predictors as of 2011:Q3, that is, at the *end* of our sample period. In our real-time forecasting exercise, these posterior probabilities were updated each time a new forecast was made and thus, in principle, could have changed over time. Figure 5 illustrates how these probabilities evolved over time. Specifically, for each indicator of economic activity, the figure plots the total posterior weight attributed to the option-adjusted credit spreads in the 20 DD-based portfolios against the time that the forecast was made. (To conserve space, we show the posterior probabilities for the four-quarter-ahead forecast horizon only.)

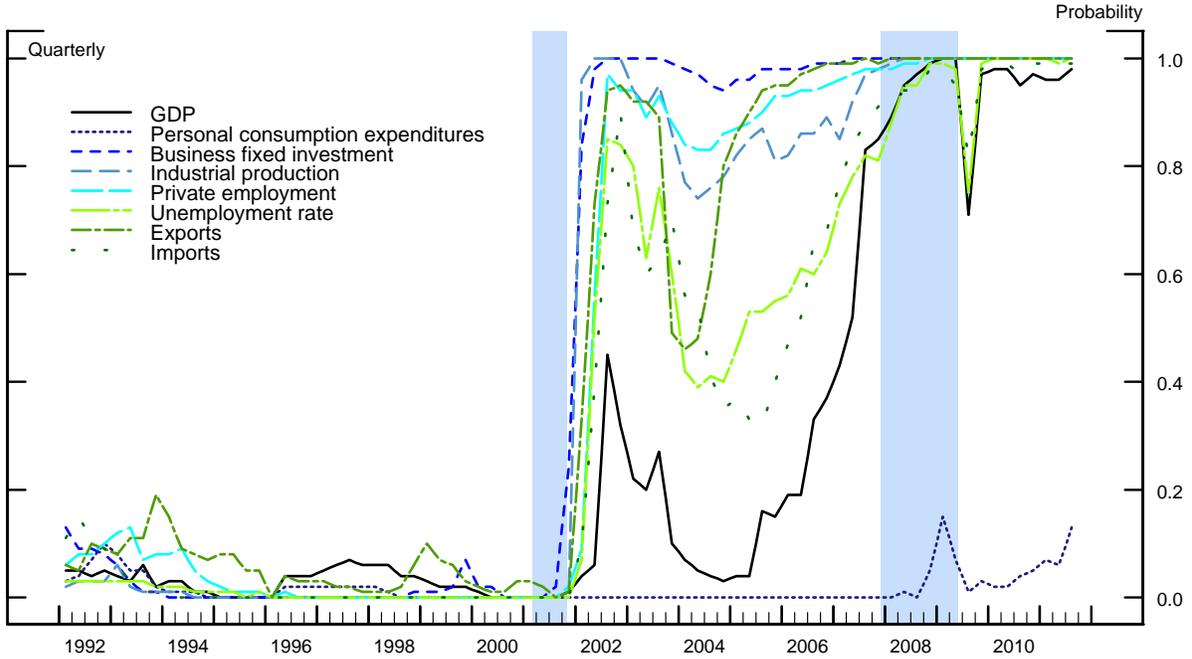
In line with the specified prior, forecasts made in the 1990s assigned very little weight to the portfolio credit spreads. The macroeconomic outcomes during the 2000–01 cyclical downturn led BMA to significantly increase—relative to other predictors—the posterior weight on the portfolio credit spreads, a pattern that was further reinforced by the 2007–09 financial crisis. In fact, by the end of our sample period, BMA assigns the vast majority of the posterior weight to the information content of credit spreads in the DD-based portfolios, a result consistent with those shown in Figure 4. However, it is important to note that during the 1990s—a portion of the sample period that is included in the forecast evaluation—the real-time BMA forecasts of economic activity based on the entire predictor set would have differed markedly from those based only on the credit spreads.

Figure 4: Final Posterior Probabilities by Predictor Subset



NOTE: The figure depicts the sum of final (as of 2011:Q3) posterior probabilities that BMA assigns to variables in the following predictor subsets: PS-I = option-adjusted credit spreads in the 20 DD-based bond portfolios; PS-II = macroeconomic variables; PS-III = other interest rates and interest rate spreads; and PS-IV = all other asset market indicators; the bars in the AR bucket represent the final posterior probabilities that BMA assigned to the autoregressive benchmark.

Figure 5: Real-Time Posterior Probabilities for Credit Spreads in DD-Based Portfolios  
*(Four-Quarter-Ahead Forecast Horizon)*

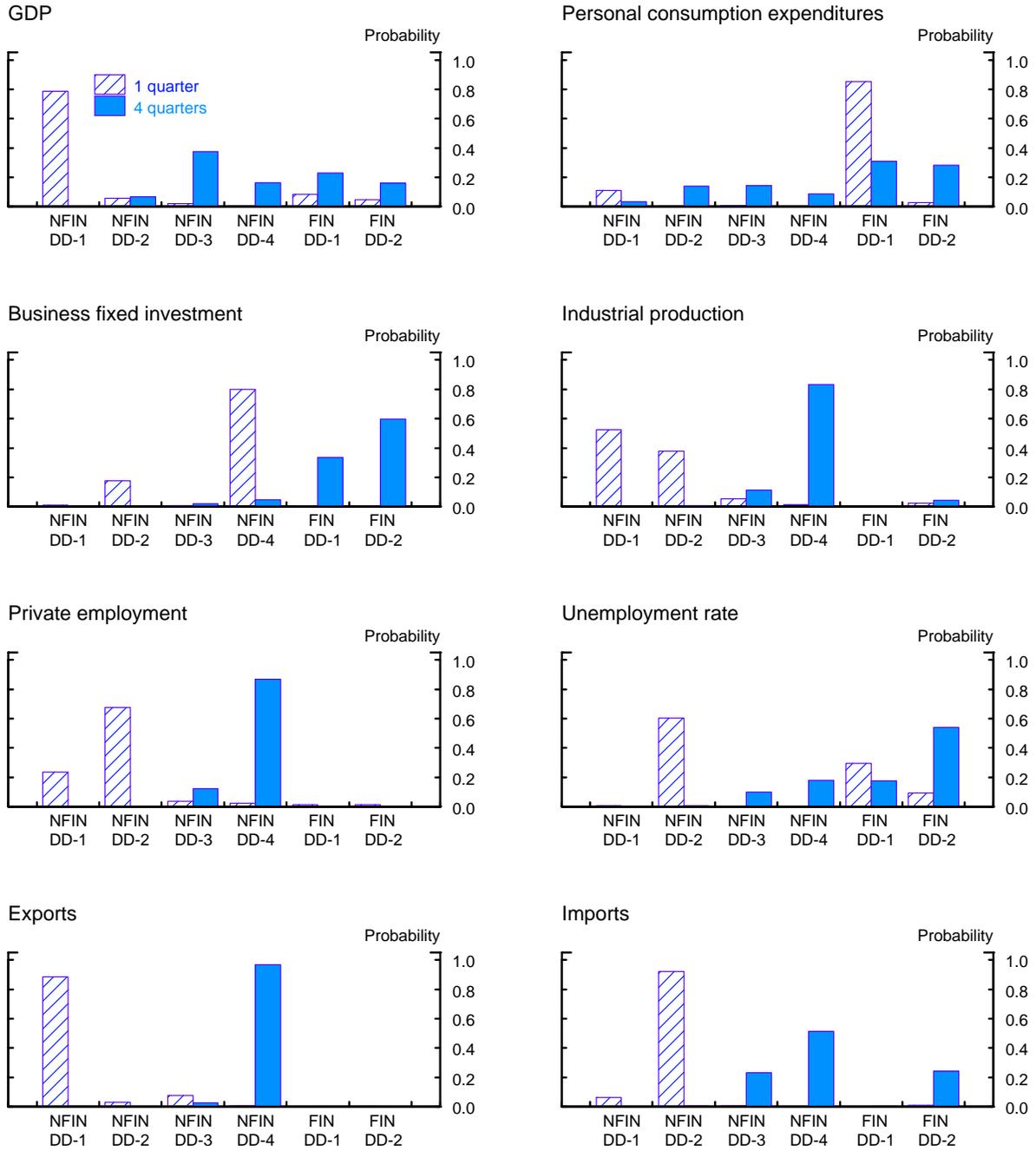


NOTE: Sample period: 1992:Q1–2011:Q3. The figure depicts the real-time evolution of the sum of posterior probabilities that BMA assigns to the option-adjusted credit spreads in the 20 DD-based bond portfolios. The results shown are for the four-quarter-ahead forecast horizon (i.e.,  $h = 4$ ). In each quarter, the posterior probabilities for the 20 portfolios—16 in the case of nonfinancial portfolios and four in the case of financial portfolios—have been added together. The shaded vertical bars represent NBER-dated recessions.

The time-series evolution of posterior weights is important because the prediction of cyclical turning points is of special interest in many forecasting applications. As emphasized by Philippon [2009], the anticipation of rising defaults associated with economic downturns may make corporate bond spreads a particularly timely indicator of an incipient recession. The result is also consistent with the recent work by Gertler and Kiyotaki [2010], Gertler and Karadi [2011], Brunnermeier and Sannikov [2011], and He and Krishnamurthy [2012], who introduce macroeconomic models in which shocks to the value of assets held by financial intermediaries—by reducing the supply of credit—have independent effects on the real economy.

Next, we examine the posterior weights implied by the forecasting exercise shown in Table 5, a case in which the predictor set includes only the option-adjusted credit spreads in the 20 DD-based bond portfolios. Figure 6 depicts the total final posterior probabilities that BMA assigns to nonfinancial portfolios in each DD quartile (NFIN-DD1, NFIN-DD2, NFIN-DD3, and NFIN-DD4) and the posterior probabilities assigned to the financial portfolios in the two halves of the DD distribution (FIN-DD1 and FIN-DD2). Results are shown for the one-quarter-ahead and four-

Figure 6: Final Posterior Probabilities for Credit Spreads in DD-Based Portfolios



NOTE: The figure depicts the sum of final (as of 2011:Q3) posterior probabilities that BMA assigns to the option-adjusted credit spreads in the DD-based bond portfolios. The results shown are for the case in which the predictor set includes only the option-adjusted credit spreads in the 20 DD-based bond portfolios (see Table 6). The posterior probabilities for maturity categories within each DD bin—four in the case of nonfinancial portfolios and two in the case of financial portfolios—have been added together.

quarter-ahead forecast horizons only. For the ease of presentation, we also summed up the posterior probabilities across the maturity categories within each DD-based portfolio—by construction, therefore, these six posterior probabilities must sum to one.

In forecasting economic activity over the subsequent quarter (i.e.,  $h = 1$ ), BMA tends to place most posterior weight on credit spreads based on portfolios that contain bonds issued by nonfinancial firms. At the four-quarter-ahead forecast horizon, in contrast, the posterior probabilities are concentrated on credit spreads based on portfolios that contain bonds issued by financial firms in the lower half of the credit-quality spectrum; though not reported, most of that posterior probability is assigned to portfolios that contain longer maturity bonds (i.e., FIN-DD1-MTY2).

## 4.2 Robustness Checks

The appendixes at the end of the paper contain a number of robustness checks. The results reported thus far have been based on the value of the shrinkage hyperparameter  $\phi = 4$ . In Appendix H, we show that our BMA forecasts generate relative RMSPEs that are less than one for a wide range of choices of  $\phi$ .

Appendix A reports BMA forecasting results that rely on different forms of empirical option-adjustment techniques when constructing DD-based portfolios, as well as for the case when we construct the portfolios using raw (i.e., unadjusted) credit spreads. In general, we find that the BMA forecasts that use raw credit spreads continue to be more accurate—at least at shorter horizons—than the forecasts obtained from direct autoregressions. However, the gains in predictive accuracy are neither as large nor as consistent as those based on the option-adjusted credit spreads.

These results suggest that the information content of credit spreads on corporate bonds is significantly influenced by fluctuations in the values of embedded options. Given that the widely-used credit spread indexes (e.g., the Baa–Aaa and the “high-yield” spreads) are constructed using prices on both callable and non-callable bonds and that the portion of callable corporate debt is changing over time, this may also help explain the uneven forecasting performance of standard credit spread indexes for future economic activity.

## 5 When Do Credit Spreads Forecast Best?

Predictive relationships between economic variables may be episodic, performing better at some times than at others; see, for example, Stock and Watson [2009] and Rossi [2012]. Therefore, it seems natural to ask if the ability of credit spreads to forecast economic activity owes its performance to recessions, expansions, or both. To formally examine this question, we consider the relative RMSPEs of the BMA forecasts over two different subsamples: (1) forecasts made for quarters that turned out to fall into NBER-dated contractions; and (2) forecasts made for quarters that turned out to fall into NBER-dated expansions. When creating these two subsamples, we use the standard

Table 7: Out-of-Sample Predictive Accuracy of BMA Forecasts During Economic Expansions and Contractions  
(Predictors: All Predictors)

Economic Activity Indicator	<i>NBER-Dated Expansions</i>					<i>NBER-Dated Contractions</i>				
	Forecast Horizon ( $h$ quarters)					Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	1.03 [0.93]	1.05 [0.80]	1.03 [0.34]	0.91 [0.02]	0.87 [0.01]	0.88 [0.16]	0.74 [0.09]	0.71 [0.11]	0.85 [0.28]	0.92 [0.35]
Personal consumption expenditures	0.97 [0.13]	0.99 [0.24]	0.92 [0.05]	0.94 [0.08]	1.03 [0.26]	0.73 [0.06]	0.80 [0.19]	1.01 [0.48]	1.12 [0.60]	1.11 [0.59]
Business fixed investment	0.94 [0.01]	0.87 [0.01]	0.97 [0.08]	0.87 [0.02]	0.88 [0.02]	0.94 [0.20]	0.76 [0.12]	0.78 [0.12]	0.84 [0.24]	0.83 [0.26]
Industrial production	1.00 [0.19]	1.03 [0.54]	1.09 [0.82]	1.02 [0.26]	0.97 [0.11]	0.92 [0.07]	0.84 [0.10]	0.85 [0.15]	0.87 [0.24]	0.89 [0.28]
Private employment	0.98 [0.08]	1.02 [0.41]	1.07 [0.59]	1.01 [0.14]	0.93 [0.05]	0.88 [0.02]	0.70 [0.02]	0.75 [0.13]	0.77 [0.24]	0.82 [0.28]
Unemployment rate	1.02 [0.88]	1.03 [0.61]	1.06 [0.64]	0.98 [0.08]	0.90 [0.03]	0.87 [0.02]	0.77 [0.03]	0.73 [0.05]	0.77 [0.09]	0.86 [0.20]
Exports	1.01 [0.54]	1.04 [0.78]	1.03 [0.29]	1.05 [0.39]	1.06 [0.37]	0.95 [0.09]	0.88 [0.07]	0.84 [0.12]	0.86 [0.20]	0.85 [0.22]
Imports	0.95 [0.00]	0.97 [0.03]	1.02 [0.24]	0.97 [0.08]	0.98 [0.11]	0.95 [0.20]	0.89 [0.19]	0.87 [0.25]	0.91 [0.30]	0.92 [0.35]

NOTE: Overall sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. The table breaks out the forecasts into two subsamples—forecasts made for quarters that turned out to fall into NBER-dated expansions (left panel) and forecasts made for quarters that turned out to fall into NBER-dated recessions (right panel). The quarters corresponding to peaks and troughs are both coded as being part of the recession. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression for the two subsamples. The autoregression is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are shown in brackets (see text for details).

NBER business cycle dates and code quarters corresponding to peaks and troughs as being part of the recession.

Clearly, this exercise is of no use to real-time economic forecasting—no forecaster ever knows whether the quarter for which the prediction is being made will turn out to be a recession or an expansion. Nevertheless, it is still useful as a diagnostic to understand, *ex post* of course, when the predictability of credit spreads in the DD-based portfolios arose.<sup>25</sup> The results of this exercise are summarized in Table 7.

Depending on the measure of economic activity being predicted and the forecast horizon, the BMA forecast may be more accurate than the autoregressive benchmark during economic expansions alone. For example, the relative RMSPE of the BMA forecast is significantly below one when predicting real GDP growth and business fixed investment during economic expansions at the three- and four-quarter-ahead horizons; a similar result holds for the year-ahead change in the unemployment rate.

In economic terms, however, the estimated gains in predictive accuracy from our BMA forecasts appear to be greatest during economic downturns. During the two NBER-dated recessions in our sample, the relative RMSPEs are below one for all indicators and all horizons, with the sole exception being the growth of real PCE at longer horizons. At the same time, the recessions typically span such a short period of time that the gains in forecast accuracy during economic contractions alone are statistically significant only at shorter horizons. All told, the documented improvements in overall predictive accuracy from BMA forecasts based on credit spreads in our DD-based portfolios appear to accrue mainly, though not exclusively, in recessionary periods.

## 6 Conclusion

This paper has revisited the forecasting of real-time economic activity using a large number of macroeconomic and asset market indicators. Our contribution involved expanding the set of asset market indicators with credit spreads based on corporate bond portfolios sorted by the instrument’s maturity and credit risk as measured by the issuer’s distance-to-default. These portfolio credit spreads were constructed directly from the secondary market prices of a large number of senior unsecured bonds issued by U.S. financial and nonfinancial corporations. Using a flexible empirical credit-spread pricing framework, the micro-level credit spreads were adjusted for the callability of the underlying issue, a pervasive feature of the corporate cash market and one that significantly influences the information content of credit spreads for future economic activity.

To take explicitly into account model selection issues, we employed Bayesian model averaging techniques. Our results indicate that the accuracy of the BMA forecasts significantly exceeds—both economically and statistically—the accuracy of the forecasts obtained from a univariate direct

---

<sup>25</sup>It may also be helpful if the user of the forecast has a nonstandard loss function that penalizes forecast misses more during recessions than expansions (or vice-versa).

autoregression, a benchmark that has proven to be quite difficult to beat when forecasting real-time economic activity.

The gains in forecasting accuracy stem almost exclusively from the inclusion of the option-adjusted portfolio credit spreads in the set of predictors—Bayesian model averaging consistently assigns very high posterior probabilities to models that include these asset market indicators. In contrast, if the portfolio credit spreads are omitted from the predictor set, the BMA forecasts of future economic activity are generally statistically indistinguishable from the forecasts obtained from a direct autoregression. This finding highlights the rich amount of information contained in corporate bond spreads, information, as argued by Gilchrist and Zakrajšek [2012], that may be particularly useful for identifying the importance of credit supply shocks in the determination of macroeconomic outcomes.

Although the combination of BMA and out-of-sample forecasting appreciably mitigates concerns about data mining, the sample period used in the analysis contains only three distinct recessions. Economic downturns in the United States have different causes, with factors such as monetary tightenings, oil price shocks, and bursting of asset price bubbles all having played varying roles in our historical sample. Time alone will tell how our BMA forecasts do in predicting future recessions. However, to the extent that significant disruptions in credit supply may also accompany future recessions, BMA forecasts utilizing the information content of credit spreads in our DD-based bond portfolios will likely provide a timely and informative signal regarding the evolution of cyclically-sensitive indicators of real economic activity, such as growth of business fixed investment and industrial output, as well as of changes in labor market conditions.

## References

- ANG, A., M. PIAZZESI, AND M. WEI (2006): “What Does the Yield Curve Tell Us About GDP Growth?,” *Journal of Econometrics*, 131(1-2), 359–403.
- AVRAMOV, D. (2002): “Stock Return Predictability and Model Uncertainty,” *Journal of Financial Economics*, 64(3), 423–458.
- BATES, J. M., AND C. W. GRANGER (1969): “The Combination of Forecasts,” *Operational Research Quarterly*, 20(4), 451–468.
- BERNANKE, B. S. (1990): “On the Predictive Power of Interest Rates and Interest Rate Spreads,” *New England Economic Review*, November, 51–68.
- BERNANKE, B. S., M. GERTLER, AND S. GILCHRIST (1999): “The Financial Accelerator in a Quantitative Business Cycle Framework,” in *The Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, pp. 1341–1393. Elsevier Science B.V, Amsterdam.
- BHARATH, S. T., AND T. SHUMWAY (2008): “Forecasting Default with the Merton Distance to Default Model,” *Review of Financial Studies*, 21(3), 1339–1369.

- BRUNNERMEIER, M. K., AND Y. SANNIKOV (2011): “A Macroeconomic Model with a Financial Sector,” Working Paper, Dept. of Economics, Princeton University.
- CAMERON, A. C., J. B. GELBACH, AND D. L. MILLER (2011): “Robust Inference with Multi-Way Clustering,” *Journal of Business and Economic Statistics*, 29(2), 238–249.
- CHIPMAN, H., E. I. GEORGE, AND R. E. MCCULLOCH (2001): “The Practical Implementation of Bayesian Model Selection,” in *Model Selection*, ed. by P. Lahiri, pp. 65–116. IMS Lecture Notes–Monograph Series, No. 38, Beachwood, OH.
- CLARK, T. E., AND M. W. MCCRACKEN (2001): “Tests of Equal Forecast Accuracy and Encompassing for Nested Models,” *Journal of Econometrics*, 105(1), 85–110.
- (2012): “Reality Checks and Comparisons of Nested Predictive Models,” *Journal of Business and Economic Statistics*, 30(1), 53–66.
- CREMERS, M. K. J. (2002): “Stock Return Predictability: A Bayesian Model Selection Perspective,” *Review of Financial Studies*, 15(4), 1223–1249.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2012): “DSGE Model-Based Forecasting,” Forthcoming, *Handbook of Economic Forecasting*.
- DIEBOLD, F. X., AND R. S. MARIANO (1995): “Comparing Predictive Accuracy,” *Journal of Business and Economic Statistics*, 13(3), 253–263.
- DUCA, J. V. (1999): “An Overview of What Credit Market Indicators Tell Us,” *Economic and Financial Review, Federal Reserve Bank of Dallas*, Third Quarter, 2–13.
- DUFFEE, G. R. (1998): “The Relation Between Treasury Yields and Corporate Bond Yield Spreads,” *Journal of Finance*, 53(6), 225–241.
- EDGE, R., AND R. S. GÜRKAYNAK (2010): “How Useful are Estimated DSGE Model Forecasts for Central Bankers,” *Brookings Papers on Economic Activity*, 41(2), 209–259.
- EDWARDS, A. K., L. E. HARRIS, AND M. S. PIWOWAR (2007): “Corporate Bond Market Transaction Costs and Transparency,” *Journal of Finance*, 62(3), 1421–1451.
- EMERY, K. M. (1999): “The Information Content of the Paper-Bill Spread,” *Journal of Business and Economic Statistics*, 48(1), 1–10.
- ESTRELLA, A., AND G. A. HARDOUVELIS (1991): “The Term Structure as Predictor of Real Economic Activity,” *Journal of Finance*, 46(2), 555–576.
- ESTRELLA, A., AND F. S. MISHKIN (1998): “Predicting U.S. Recessions: Financial Variables as Leading Indicators,” *Review of Economics and Statistics*, 80(1), 45–61.
- FAMA, E. F. (1981): “Stock Returns, Real Activity, Inflation and Money,” *American Economic Review*, 71(4), 545–565.
- FAUST, J., AND J. H. WRIGHT (2009): “Comparing Greenbook and Reduced-Form Forecasts Using a Large Realtime Dataset,” *Journal of Business and Economic Statistics*, 27(4), 486–479.

- FERNANDEZ, C., E. LEY, AND M. STEEL (2001a): “Benchmark Priors for Bayesian Model Averaging,” *Journal of Econometrics*, 100(2), 381–427.
- (2001b): “Model Uncertainty in Cross-Country Growth Regressions,” *Journal of Applied Econometrics*, 16(5), 563–576.
- FRIEDMAN, B. M., AND K. N. KUTTNER (1992): “Money, Income, Prices, and Interest Rates,” *American Economic Review*, 82(3), 472–492.
- (1998): “Indicator Properties of the Paper-Bill Spread: Lessons From Recent Experience,” *Review of Economics and Statistics*, 80(1), 34–44.
- GERTLER, M., AND P. KARADI (2011): “A Model of Unconventional Monetary Policy,” *Journal of Monetary Economics*, 58(1), 17–34.
- GERTLER, M., AND N. KIYOTAKI (2010): “Financial Intermediation and Credit Policy in Business Cycle Analysis,” in *Handbook of Macroeconomics*, ed. by B. M. Friedman, and M. Woodford, vol. 3, pp. 547–599. North-Holland, Elsevier, Amsterdam.
- GERTLER, M., AND C. S. LOWN (1999): “The Information in the High-Yield Bond Spread for the Business Cycle: Evidence and Some Implications,” *Oxford Review of Economic Policy*, 15(3), 132–150.
- GILCHRIST, S., V. YANKOV, AND E. ZAKRAJŠEK (2009): “Credit Market Shocks and Economic Fluctuations: Evidence From Corporate Bond and Stock Markets,” *Journal of Monetary Economics*, 56(4), 471–493.
- GILCHRIST, S., AND E. ZAKRAJŠEK (2012): “Credit Spreads and Business Cycle Fluctuations,” *American Economic Review*, 102(4), 1692–1720.
- GONÇALVES, S., AND B. PERRON (2011): “Bootstrapping Factor-Augmented Regression Models,” Working Paper, University of Montreal.
- GÜRKAYNAK, R. S., B. SACK, AND J. H. WRIGHT (2007): “The U.S. Treasury Yield Curve: 1961 to the Present,” *Journal of Monetary Economics*, 54(8), 2291–2304.
- HAMILTON, J. D., AND D. H. KIM (2002): “A Reexamination of the Predictability of Economic Activity Using the Yield Spread,” *Journal of Money, Credit, and Banking*, 34(2), 340–360.
- HARVEY, C. R. (1988): “The Real Term Structure and Consumption Growth,” *Journal of Financial Economics*, 22(2), 305–322.
- (1989): “Forecasts of Economic Growth from the Bond and Stock Market,” *Financial Analysts Journal*, 45(5), 38–45.
- HE, Z., AND A. KRISHNAMURTHY (2012): “Intermediary Asset Pricing,” Forthcoming, *American Economic Review*.
- KING, T. B., A. T. LEVIN, AND R. PERLI (2007): “Financial Market Perceptions of Recession Risk,” Finance and Economics Discussion Series Paper 2007-57, Federal Reserve Board.

- KOOP, G., AND S. POTTER (2004): “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging,” *The Econometrics Journal*, 7(2), 550–565.
- LEAMER, E. E. (1978): *Specification Searches: Ad Hoc Inference With Nonexperimental Data*. John Wiley & Sons, Inc., New York, NY.
- LEVIN, A. T., F. M. NATALUCCI, AND E. ZAKRAJŠEK (2004): “The Magnitude and Cyclical Behavior of Financial Market Frictions,” Finance and Economics Discussion Series Paper 2004-70, Federal Reserve Board.
- LÖFFLER, G. (2004): “An Anatomy of Rating Through the Cycle,” *Journal of Banking and Finance*, 28(3), 695–720.
- (2007): “The Complementary Nature of Ratings and Market-Based Measures of Default Risk,” *Journal of Fixed Income*, 17(1), 38–47.
- MERTON, R. C. (1973): “Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, 4(1), 141–183.
- (1974): “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” *Journal of Finance*, 29(2), 449–470.
- MIN, C., AND A. ZELLNER (1993): “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts With Applications to Forecasting International Growth Rates,” *Journal of Econometrics*, 56(1-2), 89–118.
- MODY, A., AND M. P. TAYLOR (2004): “Financial Predictors of Real Activity and the Financial Accelerator,” *Economic Letters*, 82(2), 167–172.
- MUELLER, P. (2009): “Credit Spreads and Real Activity,” Working Paper, London School of Economics.
- PHILIPPON, T. (2009): “The Bond Market’s  $q$ ,” *Quarterly Journal of Economics*, 124(3), 1011–1056.
- RAFTERY, A., D. MADIGAN, AND J. A. HOETING (1997): “Bayesian Model Averaging for Linear Regression Models,” *Journal of the American Statistical Association*, 92(437), 179–191.
- ROSSI, B. (2012): “Advances in Forecasting Under Instabilities,” Forthcoming, *Handbook of Economic Forecasting*.
- SALA-I-MARTIN, X., G. DOPPELHOFER, AND R. I. MILLER (2004): “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach,” *American Economic Review*, 94(4), 813–835.
- SIMS, C. A. (2005): “Limits to Inflation Targeting,” in *The Inflation-Targeting Debate*, ed. by B. S. Bernanke, and M. Woodford, vol. 32, pp. 283–310. NBER Studies in Business Cycles, Cambridge, MA.
- STOCK, J. H., AND M. W. WATSON (2003): “How Did Leading Indicators Forecasts Perform During the 2001 Recessions?,” *Federal Reserve Bank of Richmond Economic Quarterly*, 89(3), 71–90.

- (2005): “Implications of Dynamic Factor Models for VAR Analysis,” NBER Working Paper No. 11467.
- (2009): “Phillips Curve Inflation Forecasts,” in *Understanding Inflation and the Implications for Monetary Policy*, ed. by J. Fuhrer, Y. Kodrycki, J. Little, and G. Olivei, pp. 99–202. The MIT Press, Cambridge.
- TULIP, P. (2005): “Has Output Become More Predictable? Changes in Greenbook Forecast Accuracy?,” Finance and Economics Discussion Series Paper 2005-31, Federal Reserve Board.
- WARGA, A. D. (1991): “A Fixed Income Database,” Working Paper, University of Houston.
- WRIGHT, J. H. (2008): “Bayesian Model Averaging and Exchange Rate Forecasting,” *Journal of Econometrics*, 146(2), 329–341.
- ZELLNER, A. (1986): “On Assessing Prior Distributions and Bayesian Regression Analysis With  $g$ -prior Distributions,” in *Bayesian Inference and Decision Techniques*, ed. by P. K. Goel, and A. Zellner, pp. 233–243. North-Holland, Amsterdam, The Netherlands.

# Appendices

## A Different Option-Adjustment Procedures

This section reports the relative RMSPEs of BMA forecasts (as in Table 4 of the paper) using different forms of the call-option adjustment:

- (1) Table A-1 reports results with no option-adjustment—that is, the DD-based portfolios are constructed using raw credit spreads.
- (2) Table A-2 reports results without any Jensen’s inequality correction—the  $\hat{\sigma}_t^2/2$  term is deleted from the option adjustment.
- (3) Table A-3 reports results with a time-invariant Jensen’s inequality correction—the  $\hat{\sigma}_t^2/2$  term in the option adjustment is replaced by  $\hat{\sigma}^2/2$ , where  $\hat{\sigma}^2$  is the sample variance of the errors over all bonds and all time periods.

According to entries in Table A-1, BMA forecasts that use raw credit spreads continue to be more accurate than the forecasts obtained from direct autoregressions, at least at shorter horizons. Although gains in forecast accuracy are economically and statistically significant in some cases, they are neither as large nor as consistent—both across economic indicators and horizons—as those that relied on the option-adjusted credit spreads. The BMA forecasts that use option-adjusted spreads with the alternative option adjustments (Tables A-2 and A-3) have virtually identical predictive accuracy, compared with the baseline option adjustment (Table 4 of the paper). In summary, the option adjustment is important for the predictive content of credit spreads. However, how one treats the Jensen’s inequality term in our regression-based option-adjustment procedure is inconsequential.

Table A-1: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
*(Predictor Set: All Variables With Raw Credit Spreads)*

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.97 [0.05]	0.97 [0.06]	0.94 [0.05]	0.96 [0.07]	0.92 [0.04]
Personal consumption expenditures	0.96 [0.07]	0.97 [0.15]	1.00 [0.22]	1.05 [0.33]	1.09 [0.45]
Business fixed investment	0.93 [0.01]	0.89 [0.01]	0.96 [0.06]	0.95 [0.05]	0.94 [0.05]
Industrial production	0.97 [0.03]	0.98 [0.08]	1.05 [0.60]	1.05 [0.42]	1.05 [0.35]
Private employment	0.96 [0.03]	0.97 [0.07]	1.04 [0.37]	1.09 [0.49]	1.06 [0.30]
Unemployment rate	0.96 [0.00]	0.97 [0.04]	1.05 [0.46]	1.09 [0.55]	1.07 [0.34]
Exports	0.98 [0.00]	0.96 [0.00]	0.99 [0.02]	1.00 [0.07]	0.96 [0.03]
Imports	0.95 [0.00]	0.95 [0.01]	1.00 [0.11]	1.01 [0.15]	0.99 [0.13]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see paper for details).

Table A-2: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(Predictor Set: All Variables With Option-Adjusted Credit Spreads)  
(No Jensen's Inequality Correction)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.96 [0.03]	0.90 [0.01]	0.86 [0.01]	0.88 [0.02]	0.89 [0.03]
Personal consumption expenditures	0.89 [0.01]	0.91 [0.03]	0.97 [0.11]	1.02 [0.27]	1.06 [0.35]
Business fixed investment	0.94 [0.01]	0.82 [0.00]	0.88 [0.01]	0.86 [0.01]	0.86 [0.01]
Industrial production	0.97 [0.03]	0.93 [0.02]	0.94 [0.04]	0.92 [0.05]	0.91 [0.06]
Private employment	0.93 [0.00]	0.87 [0.00]	0.90 [0.01]	0.90 [0.02]	0.89 [0.02]
Unemployment rate	0.96 [0.01]	0.88 [0.00]	0.87 [0.00]	0.87 [0.01]	0.88 [0.02]
Exports	0.98 [0.00]	0.95 [0.00]	0.92 [0.00]	0.96 [0.01]	0.98 [0.06]
Imports	0.95 [0.00]	0.93 [0.00]	0.94 [0.03]	0.94 [0.05]	0.95 [0.08]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see paper for details).

Table A-3: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(Predictor Set: All Variables With Option-Adjusted Credit Spreads)  
(Time-Invariant Jensen's Inequality Correction)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.96 [0.03]	0.90 [0.01]	0.86 [0.01]	0.86 [0.02]	0.86 [0.02]
Personal consumption expenditures	0.88 [0.01]	0.88 [0.02]	0.94 [0.08]	1.01 [0.22]	1.06 [0.35]
Business fixed investment	0.94 [0.01]	0.82 [0.00]	0.87 [0.00]	0.87 [0.01]	0.87 [0.01]
Industrial production	0.96 [0.03]	0.93 [0.02]	0.91 [0.03]	0.90 [0.04]	0.90 [0.06]
Private employment	0.93 [0.00]	0.87 [0.00]	0.90 [0.01]	0.90 [0.02]	0.88 [0.01]
Unemployment rate	0.96 [0.01]	0.88 [0.00]	0.86 [0.00]	0.85 [0.00]	0.87 [0.01]
Exports	0.98 [0.00]	0.95 [0.00]	0.93 [0.00]	0.97 [0.02]	0.99 [0.06]
Imports	0.95 [0.00]	0.93 [0.00]	0.92 [0.02]	0.92 [0.03]	0.93 [0.05]

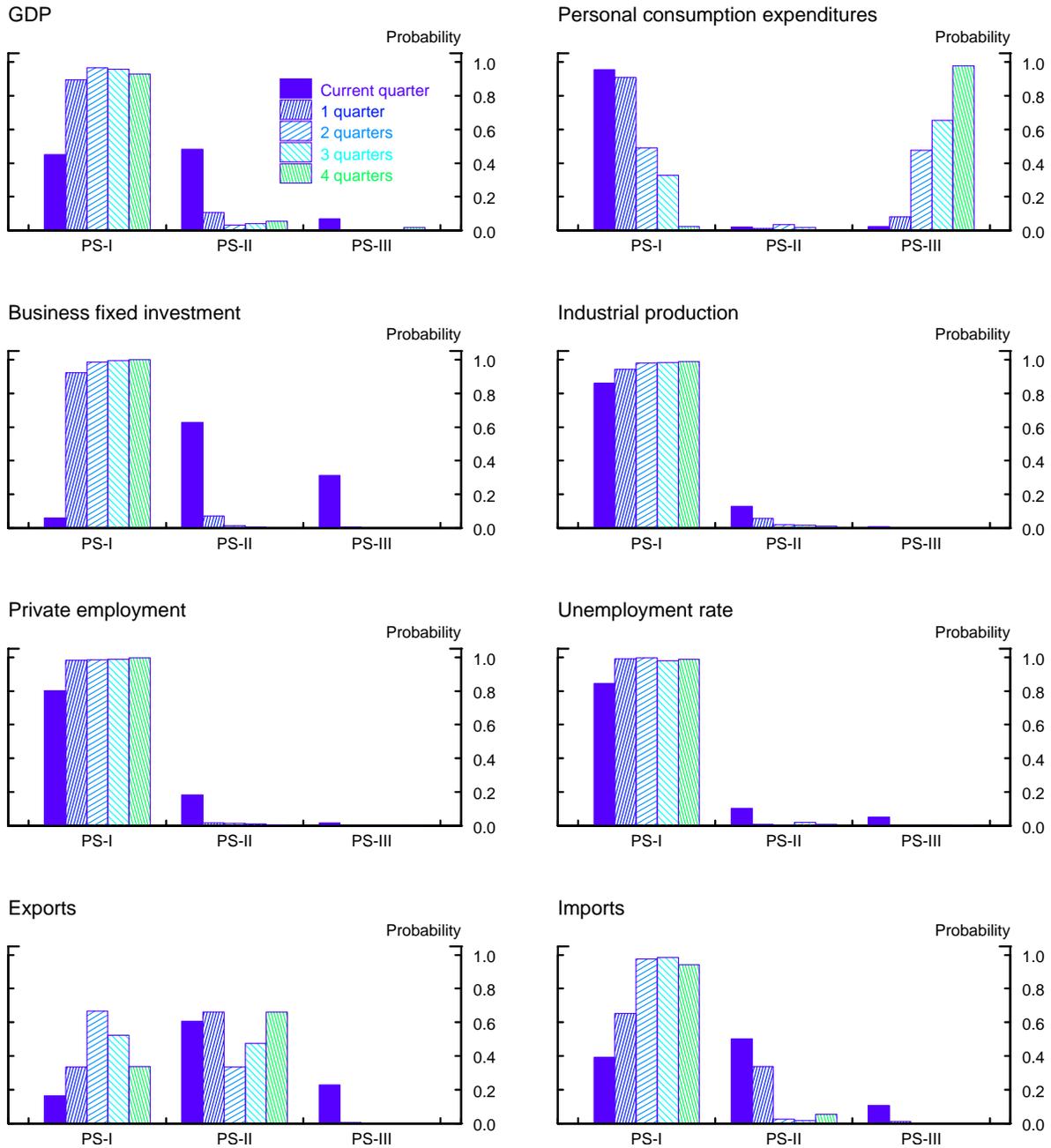
NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see paper for details).

## B Raw vs. Option-Adjusted Credit Spreads

One of the aims of the paper was to demonstrate that by using “cleaner” measures of credit spreads, one is able to improve on real-time forecasts of economic activity, as well as to point out that the well-documented uneven forecasting performance of standard credit spread indexes (i.e., the Baa-Aaa and the “junk” spreads) might reflect contamination induced by swings in the values of embedded options. This appendix compares the information content of raw credit spreads with that of their option-adjusted counterparts. In particular, the vertical bars in the panels of Figure B-1 depict the final total weights—that is, the sum of posterior probabilities—that BMA assigns to variables in the following predictor subsets: PS-I = *option-adjusted* credit spreads in the 20 DD-based bond portfolios; PS-II = *raw* credit spreads in the 20 DD-based bond portfolios; and PS-III = all other predictors (this subset of models also includes the autoregressive benchmark). Results are shown for all the forecast horizons considered and for each of the eight different indicators of economic activity. Note that, by construction, these probabilities sum up to one at each forecast horizon.

These results provide a visual confirmation of the fact that information content of the option-adjusted credit spreads consistently exceeds that of raw credit spreads, especially for cyclically-sensitive indicators of economic activity and at horizons extending beyond the very near term.

Figure B-1: Final Posterior Probabilities by Predictor Subset



NOTE: The figure depicts the sum of final (as of 2011:Q3) posterior probabilities that BMA assigns to variables in the following predictor subsets: PS-I = option-adjusted credit spreads in the 20 DD-based bond portfolios; PS-II = raw credit spreads in the 20 DD-based bond portfolios; and PS-III = all other predictors (this subset of models also includes the autoregressive benchmark).

## C BMA Forecasts With the Median Credit Spread

This appendix reports the relative RMSPEs of BMA forecasts (as in Table 4 of the paper) that use only the median credit spread of our 24 DD-based bond portfolios, rather than all the 24 credit spreads. According to the entries in Table C-1, the resulting BMA forecasts continue to be more accurate than the forecasts obtained from direct autoregressions. The predictive gains, however, are smaller than those from the corresponding forecasts that included all the portfolio-based credit spreads.

Table C-1: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(*Predictor Set: All Variables With the Median Option-Adjusted Credit Spread*)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.98 [0.10]	0.96 [0.06]	0.94 [0.07]	0.95 [0.08]	0.95 [0.10]
Personal consumption expenditures	0.95 [0.05]	0.96 [0.09]	0.99 [0.21]	1.03 [0.31]	1.06 [0.38]
Business fixed investment	0.95 [0.01]	0.85 [0.00]	0.85 [0.00]	0.86 [0.01]	0.86 [0.01]
Industrial production	0.97 [0.03]	0.93 [0.02]	0.92 [0.03]	0.91 [0.04]	0.92 [0.06]
Private employment	0.98 [0.03]	0.94 [0.02]	1.01 [0.24]	1.06 [0.44]	1.04 [0.26]
Unemployment rate	0.96 [0.01]	0.94 [0.00]	0.99 [0.10]	1.02 [0.27]	1.03 [0.26]
Exports	0.98 [0.00]	1.05 [0.93]	1.05 [0.73]	1.03 [0.35]	1.02 [0.23]
Imports	0.95 [0.00]	0.93 [0.00]	0.93 [0.02]	0.90 [0.02]	0.92 [0.04]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see paper for details).

## D Alternative Benchmark

This appendix reports the RMSPEs of BMA forecasts (as in Table 4 of the paper), relative to the benchmark of a constant forecast, which is equal to the real-time sample average of the variable being predicted at all horizons. According to the entries in Table D-1, the BMA forecasts incorporating the information content of option-adjusted credit spreads in the DD-based portfolios again generally do better than this very simple benchmark.

Table D-1: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(*Predictor Set: All Variables With Option-Adjusted Credit Spreads*)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.89 [0.00]	0.85 [0.01]	0.83 [0.01]	0.85 [0.03]	0.85 [0.02]
Personal consumption expenditures	0.85 [0.08]	0.85 [0.17]	0.89 [0.27]	0.98 [0.38]	1.03 [0.44]
Business fixed investment	0.81 [0.00]	0.72 [0.00]	0.79 [0.00]	0.82 [0.02]	0.84 [0.02]
Industrial production	0.90 [0.00]	0.90 [0.02]	0.93 [0.05]	0.92 [0.06]	0.90 [0.05]
Private employment	0.48 [0.03]	0.53 [0.03]	0.62 [0.05]	0.66 [0.05]	0.68 [0.05]
Unemployment rate	0.80 [0.00]	0.74 [0.00]	0.77 [0.00]	0.80 [0.00]	0.82 [0.01]
Exports	0.94 [0.03]	0.95 [0.00]	0.94 [0.00]	0.99 [0.20]	1.01 [0.29]
Imports	0.92 [0.44]	0.96 [0.22]	0.94 [0.03]	0.97 [0.07]	0.99 [0.10]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a constant forecast (i.e., a forecast that is set equal to the recursively estimated real-time sample mean at all horizons); the constant forecast is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see paper for details).

## E Alternative Bootstrap

This appendix reports the relative RMSPEs of BMA forecasts (as in Table 4 of the paper), except that the  $p$ -values of the null hypothesis of equal predictive accuracy use a different bootstrap procedure.

This alternative bootstrap involves fitting an AR(4) process to  $y_t$  and separately estimating a dynamic factor model using the set of all predictors  $X_t$ :

$$X_t = \Lambda F_t + u_t; \tag{E-1}$$

and

$$F_t = \Phi F_{t-1} + v_t, \tag{E-2}$$

where the elements of the vector  $F_t$  correspond to the first three principal components of  $X_t$ . In each bootstrap replication, we first re-sample with replacement from the residuals of the AR(4) process for  $y_t$  to construct bootstrap samples of  $y_t$ . We then independently re-sample with replacement from the residuals in equations (E-1) and (E-2), thereby constructing bootstrap samples of  $X_t$  for use in BMA; note that in this setup, the predictor set  $X_t$  is, by construction, irrelevant for the forecasting of the dependent variable. Unlike the bootstrap in the paper, this bootstrap does not preserve the conditional heteroskedasticity that may be present in the underlying data.

Entries in Table E-1 shows the relative RMSPEs of our BMA forecasts, exactly as in Table 4, except using these alternative  $p$ -values; these alternative  $p$ -values are quite similar to those reported in the paper.

Table E-1: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(Predictor Set: All Variables With Option-Adjusted Credit Spreads)  
(Alternative Bootstrap)

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	0.96 [0.01]	0.90 [0.00]	0.87 [0.00]	0.88 [0.01]	0.87 [0.01]
Personal consumption expenditures	0.88 [0.00]	0.90 [0.01]	0.97 [0.11]	1.03 [0.49]	1.06 [0.70]
Business fixed investment	0.94 [0.00]	0.84 [0.00]	0.90 [0.00]	0.90 [0.01]	0.88 [0.01]
Industrial production	0.96 [0.00]	0.92 [0.00]	0.93 [0.02]	0.92 [0.02]	0.91 [0.02]
Private employment	0.92 [0.00]	0.86 [0.00]	0.90 [0.00]	0.90 [0.01]	0.88 [0.01]
Unemployment rate	0.95 [0.00]	0.87 [0.00]	0.86 [0.00]	0.87 [0.00]	0.88 [0.00]
Exports	0.98 [0.02]	0.95 [0.00]	0.93 [0.00]	0.98 [0.07]	1.00 [0.16]
Imports	0.94 [0.00]	0.92 [0.00]	0.93 [0.02]	0.93 [0.03]	0.94 [0.04]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets; these  $p$ -values are based on the alternative bootstrap described in Appendix D.

## F Forecasting With the Term Spread and the Paper-Bill Spread

This appendix reports the relative RMSPEs of forecasts that use—in addition to the lags of the variable being forecasted—either the 10y/3m term spread or the paper-bill spread as predictor. Table F-1 reports the results using the term spread, while Table F-2 reports the results using the paper-bill spread.

Table F-1: Real-Time Out-of-Sample Predictive Accuracy of the Term Spread

Economic Activity Indicator	Forecast Horizon ( $h$ quarters)				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	1.01 [0.44]	1.00 [0.24]	1.00 [0.21]	1.00 [0.21]	1.00 [0.21]
Personal consumption expenditures	1.00 [0.39]	1.01 [0.52]	1.02 [0.48]	1.03 [0.54]	1.03 [0.52]
Business fixed investment	0.99 [0.19]	0.99 [0.16]	0.98 [0.13]	0.98 [0.12]	0.97 [0.11]
Industrial production	0.99 [0.09]	1.00 [0.15]	1.00 [0.15]	0.99 [0.09]	0.98 [0.08]
Private employment	1.00 [0.26]	1.00 [0.22]	1.01 [0.24]	1.00 [0.19]	0.99 [0.13]
Unemployment rate	1.01 [0.71]	1.01 [0.49]	1.01 [0.27]	1.00 [0.18]	0.98 [0.11]
Exports	1.02 [0.98]	1.03 [0.91]	1.04 [0.88]	1.05 [0.88]	1.07 [0.86]
Imports	0.99 [0.07]	0.99 [0.12]	0.99 [0.12]	0.98 [0.10]	0.96 [0.07]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the forecast using a direct autoregression augmented with the term spread (10-year less 3-month Treasury yield) to the forecast using the unaugmented direct autoregression. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets, using the bootstrap described in the paper.

Table F-2: Real-Time Out-of-Sample Predictive Accuracy of the Paper-Bill Spread

Economic Activity Indicator	Forecast Horizon ( $h$ quarters]				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP	1.05 [1.00]	1.04 [0.98]	1.09 [1.00]	1.08 [0.99]	1.09 [0.99]
Personal consumption expenditures	0.97 [0.04]	1.03 [0.92]	1.04 [0.92]	1.05 [0.91]	1.05 [0.88]
Business fixed investment	0.95 [0.0]	1.00 [0.32]	1.01 [0.49]	1.02 [0.61]	1.02 [0.60]
Industrial production	1.00 [0.18]	0.97 [0.11]	1.01 [0.61]	1.02 [0.60]	1.02 [0.51]
Private employment	0.98 [0.06]	0.99 [0.14]	1.03 [0.72]	1.05 [0.83]	1.05 [0.80]
Unemployment rate	0.98 [0.03]	0.98 [0.06]	1.00 [0.48]	1.02 [0.67]	1.02 [0.72]
Exports	1.01 [0.50]	1.07 [0.97]	1.07 [0.98]	1.08 [1.00]	1.07 [0.99]
Imports	0.95 [0.00]	0.94 [0.01]	0.97 [0.09]	0.99 [0.22]	0.99 [0.26]

NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate (or change in the case of unemployment rate) of each economic activity indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the forecast using a direct autoregression augmented with the 3-month paper-bill spread to the forecast using the unaugmented direct autoregression. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets, using the bootstrap described in the paper.

## G BMA Forecasts of Inflation

This appendix reports the relative RMSPEs of our BMA forecasts (as in Table 4 of the paper), but where the variable being forecasted is inflation, measured either by the GDP price deflator or the CPI. According to the entries in Table G-1, our BMA forecasts do not improve on the autoregressive benchmark.

Table G-1: Real-Time Out-of-Sample Predictive Accuracy of BMA Forecasts  
(*Predictor Set: All Variables With Option-Adjusted Credit Spreads*)

Inflation Measure	Forecast Horizon ( $h$ quarters]				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
GDP price deflator	1.00 [0.53]	1.00 [0.52]	1.01 [0.44]	1.01 [0.37]	1.02 [0.36]
CPI	1.00 [0.40]	1.02 [0.66]	1.02 [0.35]	1.01 [0.27]	1.02 [0.26]

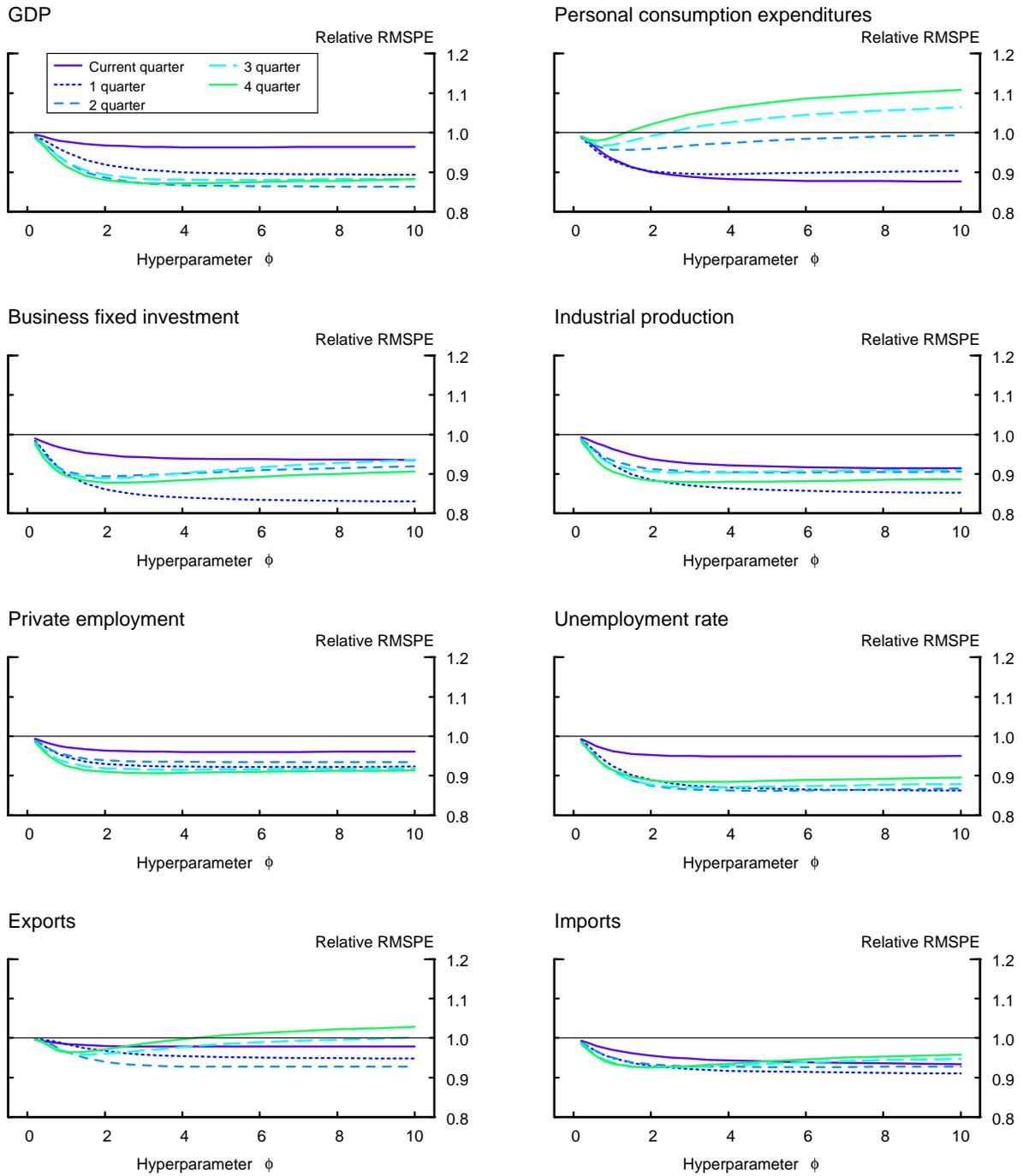
NOTE: Sample period: 1986:Q1–2011:Q3. The jump-off date for the out-of-sample recursive forecasts is 1992:Q1. The forecasted variable is the cumulative growth rate of each inflation indicator over the specified forecast horizon. Entries in the table denote the ratio of the RMSPE from the BMA forecast to the RMSPE from a direct autoregression; the autoregressive benchmark is also included in the set of candidate models. Bootstrapped  $p$ -values (500 replications) for the test of the null hypothesis that the ratio of the RMSPEs is equal to one are reported in square brackets (see paper for details).

## H Varying the Hyperparameter $\phi$

The results reported in the paper were based on the value of the shrinkage hyperparameter  $\phi = 4$ . This appendix examines the robustness of our results to different values of  $\phi$ , the parameter governing the strength of the  $g$ -prior. Figure H-1 plots the out-of-sample RMSPE of the BMA forecast—relative to the RMSPE from a direct autoregression—as a function of  $\phi$  for all six economic indicators and all five forecast horizons. Our BMA forecasting setup delivers substantial gains in forecast accuracy relative to the direct autoregression for a wide range of values of  $\phi$ ; in fact, the qualitative nature of our results appears to be fairly insensitive to the choice of the shrinkage parameter. In some cases, the relative RMSPE decreases monotonically in  $\phi$  (at least over the range of values of  $\phi$  considered). In others, the relationship between the RMSPE and  $\phi$  is U-shaped, and the best forecasts are consequently obtained with a small or intermediate value of  $\phi$ .

With a sufficiently small value of  $\phi$ —implying a very informative prior—BMA outperforms the univariate time-series benchmark in all cases considered in this paper. This is an attractive feature of BMA with a sufficiently informative prior, at least in this data set. Overall, setting  $\phi = 4$  as a benchmark seems to be a good choice because it yields relative RMSPEs that are less than one in nearly all cases, and it often yields substantial gains in forecast accuracy.

Figure H-1: BMA Forecasting Performance and the Informativeness of the  $g$ -Prior



NOTE: The figure depicts the ratio of the RMSPE of the BMA forecast to the RMSPE from a direct autoregression for the different values of the shrinkage hyperparameter  $\phi$ .