Forecasting Industry Employment for a Resource-based Economy Using Bayesian Vector Autoregressive Models^{*}

Chang K. Seung and Sung K. Ahn⁺

ABSTRACT. Bayesian vector autoregressive (BVAR) models are developed to forecast industry employment for a resource-based economy. Two different types of input-output (I-O) information are used as priors: (i) a reduced-form I-O relationship and (ii) an economic-base version of the I-O information. Out-of-sample forecasts from these two I-O-based BVAR models are compared with forecasts from an autoregressive model, an unconstrained VAR model, and a BVAR model with a Minnesota prior. Results indicate most importantly that overall the model version with economic base information performs the best in the long run.

Keywords: Bayesian vector autoregressive models, industry employment forecast, economic base, input-output information, natural resource sectors

JEL Classification Codes: R15, C11, C53

1. INTRODUCTION

Since Doan, Litterman, and Sims (1984) first used the Bayesian vector autoregressive (BVAR) approach to forecast macroeconomic variables, numerous studies have been conducted for national macroeconomic time series studies (e.g., Todd, 1984; Litterman, 1986; LeSage and Magura, 1991) or regional time-series studies (e.g., Amirizadeh and Todd, 1984; Magura, 1990; Partridge and Rickman, 1998; Puri and Soydemir, 2000; Rickman, 2001; Rickman, 2002). For a national-level analysis, Litterman (1986), for example, used a BVAR model to show that prior means and variances can improve the forecasting accuracies for macroeconomic variables. For an example of a regional analysis, Partridge and Rickman (1998) used the approach to forecast industry employment for the state of Georgia. In their study, the authors incorporated regional employment-based input-output (I-O) coefficients to specify prior means in one model and to weight the variances of a Minnesota-type prior in another model. They considered final demand effects and relationships to national and world economies.

Most of existing BVAR studies for a regional analysis have compared the forecasting performance among various alternative models employing different assumptions, including those on hyperparameters, and provided comparative discussion of the results from these models. For example, Magura (1990) found that BVAR models with I-O information produce greater forecast accuracy than autoregressive models (AR), unrestricted vector autoregressive models (UVAR), or more naïve BVAR models. Partridge and Rickman (1998) found that AR models and Minnesota prior models (MVAR) are more accurate in the short run, whereas most I-O-oriented models are more accurate in the long run. Up to this point, however, regional applications of the BVAR approach are limited to a few regions or states [e.g., Georgia (Partridge and Rickman,

^{*} The authors acknowledge funding and support from the National Marine Fisheries Service, Alaska Fisheries Science Center.

⁺ Seung is an Economist with the Alaska Fisheries Science Center, National Marine Fisheries Service. Ahn is a Professor of Statistics, Department of Finance and Management Science, College of Business, Washington State University.

Contact author: Chang K. Seung, Alaska Fisheries Science Center, National Marine Fisheries Service, 7600 Sand Point Way NE, Seattle, WA 98115-6349, USA; E-mail: <u>Chang.Seung@noaa.gov</u>.

1998), Oklahoma (Rickman, 2001), Ohio (Magura, 1998), and California (Puri and Soydemir, 2000)].¹ Their conclusions about the forecasting performance of a specific type of model can not be generalized to other regions or states unless similar results from the model are found for the regions. Although BVAR models are promising for regional economic forecasting, their applicability to other regions and time periods needs to be examined.

There are several innovative features in the present study that distinguish it from previous regional BVAR studies. First, this study develops an economic-base² version of BVAR model for a resource-dependent economy. Standard I-O-based BVAR models may be more appropriate for manufacturing-based economies (Ohio as in LeSage and Magura, 1991; and Georgia as in Partridge and Rickman, 1998). However, this type of standard IO-based BVAR may not be suitable for the Alaska economy, which is characterized by heavy dependence on natural resources as an economic base. There are some BVAR studies which consider the relationship between basic and nonbasic sectors. For example, Rickman (2001) specified a model in which manufacturing and mining industries are classified as basic industries and non-manufacturing industries as nonbasic industries, and estimated the equations with Minnesota prior imposed. However, the study did not include an economic base relationship as Bayesian priors. In contrast, Rickman, Miller, and McKenzie (2009) developed a model in which a prior of proportionality between basic and nonbasic sectors is imposed in estimating the equations. However, none of these studies or other previous regional BVAR studies used a framework in which the economic base relationship is specified as Bayesian prior for a "resource-based" economy. Thus, the present study represents the first attempt to develop an economic-base version BVAR model for analyzing an economy which depends to a large extent on natural resources.

Second, unlike previous studies which used as priors IO relationships based mostly on IMPLAN (IMpact Analysis for PLANing; Minnesota IMPLAN Group, 2001) data, this study uses both (1) IMPLAN data for non-seafood industries and (2) non-IMPLAN data for the seafood industry. In the present study, the non-IMPLAN data for the seafood industry was estimated independently using available federal and state government sources and informal interviews with key industry informants. This study thereby overcomes a serious weakness in studies that use IMPLAN seafood data, which is based on national average production functions that do not adequately characterize the scale and mix of products in the Alaska seafood industry.

Therefore, in this paper, the BVAR method is applied to the State of Alaska's economy for forecasting the industry employment in the state. Five different model variants are used:(1) an autoregressive model of degree 1 (AR), (2) an unrestricted vector autoregressive model (UVAR), (3) a BVAR model with a Minnesota prior (MVAR), (4) a BVAR with employment-

¹ Other regional applications of BVAR approach include Dua and Ray (1995), Fullerton (2001), and Gupta and Das (2008). There are some studies in the literature which use vector autoregressive (VAR) models for identifying the sources of industry employment fluctuations. These studies include Coulson (1993); Carlino, DeFina, and Sill (2001); and Chang and Coulson (2001). However, these studies do not use BVAR approach. There is also a study claiming that BVAR model with Minnesota prior does not perform better than judgmentally adjusted large-scale structural models in forecasting US economic activity (Bischoff, Belay, and Kang, 2000).

 $^{^{2}}$ Over the years, there have been numerous empirical applications of the economic base model. For example, LeSage and Reed (1989) used VAR techniques to identify the dynamic properties of the relationship between exports and local employment for eight metropolitan areas in Ohio. They used dynamic location quotient methods to decompose an area's employment into export and local components. In contrast to their study, the present study incorporated economic base information as priors within BVAR framework.

based I-O relationships [i.e., reduced-form employment relationships as in Partridge and Rickman (1998) and Rickman (2001)] (IO1_VAR), and (5) a BVAR which uses employmentbased I-O relationships with economic base relationships incorporated (IO2 VAR). This is the first example of BVAR application to a regional level analysis for an economy dependent on a natural resource economic base. Out-of-sample forecasts are conducted for these models and the results are compared in terms of forecasting accuracy. The remainder of this paper is organized as follows. Section 2 presents the methods used in this study, including a discussion of Bayesian priors and I-O information used. Section 3 provides descriptions of different model variants and data used. Section 4 describes how the forecasting experiments are conducted, and presents the results. The final section provides summary and conclusions.

2. BAYESIAN VECTOR AUTOREGRESSIVE ESTIMATION

This section describes the Bayesian VAR and Minnesota prior and provides descriptions of how the I-O information is derived.

2.1. Bayesian VAR

The unrestricted vector autoregressive (UVAR) model suffers from degrees of freedom problems, as the number of endogenous variables increases when the model is used for forecasting. A more typical problem with UVAR models is the inefficiency of estimates, which occurs due to over parameterization and bias from nonstationarity (Bewley, 2002). Because of these problems, Litterman (1980) introduced a ridge-type modification of UVAR technique and specified the prior means of the UVAR coefficients as zero except for the coefficient for the first-own lag, whose prior mean is specified as one. To estimate the BVAR, Litterman (1980) extended Theil's (1963) mixed-estimation method. Specifically, let Y be the vector of industry employment, X the corresponding matrix of lagged dependent variables, and β the true parameter vector. To conduct mixed estimation, stochastic restrictions are imposed on the estimation of β . This results in a prior distribution of β with mean b and variance-covariance matrix ψ . To derive the mixed estimator, dummy observations are used to augment the data vectors as follows:

(1)
$$\begin{pmatrix} Y \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix} \beta + \begin{pmatrix} u \\ v \end{pmatrix}$$

Here, r and R are a vector of prior means and an identity matrix, respectively; E(u) = E(v) = 0; and $E(vv') = \psi$.

. .

The BVAR parameter estimates are then calculated as:

(2)
$$\beta_{Theil} = \left(XX + \sigma_u^2 R' \Psi R\right)^{-1} \left(XY + \sigma_u^2 R' \Psi r\right)$$

To compute β_{Theil} for each equation *i*, Litterman (1980) estimated σ_{μ}^2 by using the corresponding equation in the UVAR. To specify Ψ , Litterman (1980) used a set of hyperparameters as follows:

(3)
$$\lambda^{2}(i,j,k) = \left[\theta f(i,j) g(k) \frac{S_{i}}{S_{j}}\right]^{2}$$

© Southern Regional Science Association 2011.

In this equation, $\hat{\mathcal{X}}^{(i,j,k)}$ is the variance of the prior for the coefficient on variable *j* in equation *i* at lag length *k*. θ is a parameter representing overall tightness. Smaller values θ reflect less uncertainty around the prior means. In Equation (3), f(i,j) reflects relative tightness around the prior for variable *j* in equation i. If f(i,i) = 1, then a value of 0.5 for f(i,j) indicates that the lags of variable *j* in equation i receives half the weights of own variable *i*'s lags. g(k) is a lag decay function that tightens the distribution around the prior means for greater lag lengths of the each variable in the right hand side of the equation. Typically, a harmonic lag decay function is specified such that $g(k) = k^{-\phi}$ where *k* is lag length and ϕ is a parameter reflecting the decay rate. S_i is standard error for univariate autoregression for variable *i*. Therefore, S_i/S_j is a scaling factor, and adjusts for varying magnitudes of the variables across equations *i* and *j*. For more detailed descriptions of the BVAR approach, see Litterman (1980) and Partridge and Rickman (1998).

2.2. Minnesota Prior

In the so-called Minnesota prior approach, which was introduced by Litterman (1980), the coefficient of the first-own lag is set equal to one and all the other coefficients, including all the other own-lag coefficients and the coefficients for the other variables in the BVAR system, are set equal to zero. In most previous studies, the hyperparameters are typically specified as follows. First, the overall tightness (θ) is set equal to 0.1. Studies generally show that the results are not very sensitive to the values of θ (e.g., Doan, Litterman, and Sims, 1984; Magura 1990; Rickman, 2001). Smaller values θ reflect less uncertainty around the prior means while larger values produce results closer to those from UVAR. f(i,i) is often set equal to unity while f(i, j), $i \neq j$ is set equal to 0.5. ϕ is often set equal to unity. In this study, the values of hyperparameters are set as follows. $\theta = 0.2$ following LeSage and Magura (1991), f(i,i) = 1, f(i, j) = 0.5, and $\phi = 1$. In models with I-O information, f(i, j) is replaced by employment-based I-O relationships (in IO1_VAR) or by economic base version of the I-O relationships (in IO2_VAR).

2.3. Input-Output Information

In previous studies, two different types of I-O information have been incorporated to specify prior variances in the BVAR model. In the first approach, as in LeSage and Magura (1991), only national or regional direct I-O coefficients are included as prior information. In the second approach, as with more recent studies such as Partridge and Rickman (1998) and Rickman (2001), final demand effects and links to national and world economies as well as regional I-O coefficients are incorporated. In both cases the output-based I-O prior information is converted into employment-based information using an employment-output ratio before estimating the BVAR model.

The present study uses employment-based I-O prior information based on an Alaska social accounting matrix (SAM). The Alaska SAM was constructed using non-IMPLAN data for the seafood industry and IMPLAN data for non-seafood industries. Based on this SAM, two different types of I-O information as prior variance³ are specified. In one model version

³ In this study, prior means are not used. We plan to use them for a future study.

(IO1_VAR), reduced-form employment relationships are used as in Partridge and Rickman (1998) and Rickman (2001). The reduced-form employment relationships, f(i, j), are derived from the SAM following the procedures in Partridge and Rickman (1998), and include both endogenous interindustry transactions (reflected in I-O direct coefficients) and endogenous final demand responses. More specifically, the following procedures are used to derive the reduced-form I-O relationships.

First, the row equation in the Alaska SAM model is used as follows:

(4)
$$Q_i = \sum_{j=1}^n \alpha_{ij} Q_j + \alpha_{Ci} C + \alpha_{Ii} I + \alpha_{SLi} G_{SL} + \alpha_{FDi} G_{FD} + \alpha_{Xi} X$$

where Q_i is industry *i*'s output; *C*, *I*, G_{SL} , G_{FD} , and *X* are, respectively, consumer demand, investment demand, state and local government demand, federal government demand, and exports; *n* is the number of industries in the model including federal and state/local government sectors; α_{ij} is the technical coefficient that relates output of industry *i* to output of industry *j*; α_{Ci} , α_{Ii} , α_{SLi} , α_{FDi} , and α_{Xi} are the proportions of the corresponding final demands which consist of industry *i*'s output. Thus Equation (4) above relates industry *i*'s output (Q_i) to output in other industries through intermediate input demands and final demands (*C*, *I*, G_{SL} , G_{FD} , and *X*). Next, to *derive* reduced-form employment I-O relationships, Q_i is replaced by E_i/β_i where E_i is employment in industry *i* and β_i is employment to output ratio in industry *i*. Substituting Q_i for E_i/β_i , and ignoring the export term for the moment,

(5)
$$E_{i} = \beta_{i} \left[\sum_{j=1}^{n} \left(\frac{\alpha_{ij}}{\beta_{j}} \right) E_{j} + \alpha_{Ci} C + \alpha_{Ii} I + \alpha_{SLi} G_{SL} + \alpha_{FDi} G_{FD} \right]$$

Then, local final demand components (C, I, and G_{SL}) are expressed in terms of employment as follows.

$$(6) C = \gamma_C \sum_{h=1}^n E_h W_h$$

$$I = \gamma_I \sum_{h=1}^n E_h W$$

(8)
$$G_{SL} = \gamma_{SL} \sum_{h=1}^{n} E_h W_h$$

where W_h is the average annual wage rate for industry h; γ_C , γ_I , and γ_{SL} are respective ratios of local final demand to total wage and salary income in the region. Federal government spending is assumed to be exogenous. Substituting Equations (6), (7), and (8) into Equation (5) and calculating, in the resulting equation, partial derivative of employment of industry *i* with respect to employment in industry *j* yields:

(9)
$$\frac{\partial E_i}{\partial E_j} = \beta_i \left[\frac{\alpha_{ij}}{\beta_j} + W_j \left(\alpha_{Ci} \gamma_C + \alpha_I \gamma_I + \alpha_{SL} \gamma_{SL} \right) \right]$$

Equation (9) presents long-run change in employment in industry *i* in response to a unit change in employment in industry *j*. For more detailed procedures to derive reduced-form employment I-O relationships, see Partridge and Rickman (1998). The reduced-form employment I-O relationships thus derived are used as f(i, j) in the I-O-based BVAR (IO1_VAR) below.

© Southern Regional Science Association 2011.

In the other I-O-based model version (IO2_VAR), economic base relationships are incorporated into the reduced-form employment relationship f(i, j). Specifically, natural resource sectors (oil and gas and seafood) and government are specified as the basic industries in IO2_VAR. Next, elements in the basic sectors' column and nonbasic sectors' row in f(i, j) are replaced by one, while the elements in nonbasic sectors' column and basic sectors' row are derived by dividing the corresponding elements in the f(i, j) used in IO1_VAR by 10. By specifying an economic base BVAR this way, this study generalizes the BVAR approach to resource-based regional economies. Standard I-O-based BVAR models may be more appropriate for manufacturing-based economies (Ohio as in LeSage and Magura, 1991; and Georgia as in Partridge and Rickman, 1998). However, this type of standard I-O-based BVAR may not be as suitable for the Alaska economy, which is characterized by heavy dependence on natural resources as an economic base.

3. MODEL SPECIFICATION AND DATA

In this study, five different models and associated model assumptions (outlined below) are specified for the state of Alaska and the results from these models are compared in terms of forecasting performance:

- (1) AR model: note that only first-own lag of the dependent variable is included on the righthand side.
- (2) UVAR model: in this model there are no constraints.
- (3) MVAR model: the prior mean of the first-own lag is set at unity while the prior means for all the other variables and lags are set equal to zeroes (Litterman 1986). A mixed estimation method is used. f(i,i) is set equal to unity and f(i, j) equal to 0.5. The tightness parameter (θ) is set to 0.2.
- (4) IO1_VAR model: reduced-form employment I-O relationships (Partridge and Rickman, 1998; Rickman, 2001) derived from the Alaska SAM are used to specify prior variances. In other words, the reduced-form employment relationships are used as the values for the weighting function f(i, j) elements. f(i, i) is set equal to unity. Value of θ is set equal to 0.2.
- (5) IO2_VAR model: the economic-base relationships are incorporated in the reduced-form employment relationships in f(i, j) to specify the prior variances. The elements in basic sectors' column and nonbasic sectors' row in f(i, j) are replaced by one while the elements in nonbasic sectors' column and basic sectors' row in f(i, j) are derived by dividing the corresponding elements in the f(i, j) used in IO1_VAR by 10. f(i, i) is set equal to unity. Value of θ is set to 0.2.

This study used the U.S. Bureau of Labor Statistics' (BLS) monthly NAICS-based employment data for the State of Alaska's industries for the years 1990 through 2005. The data linkages between industries within Alaska were not seasonally adjusted and were aggregated to 15 sectors for analysis (see Table 1). The 15 sectors are oil and gas (OIL), seafood (SEA), government (GOV), other natural resource sector (ONR), transportation (TRN), construction (CON), manufacturing (MAN), wholesale trade (WHO), retail trade (RET), information (INF), finance and real estate (FIN), professional service (PRO), educational service (EDU), leisure

Industry Name	NAICS Code	IMPLAN Sector	Employment Share
Oil and Gas (OIL)	211	19	3.4
Seafood (SEA)	1141, 3117	16, 71	3.4
Government (GOV)	NA	495-499, 503-506	27.9
Other Natural Resource (ONR)	111-113, 1142, 115, 212, 213	1-15, 17, 18, 20-29	0.9
Transportation and Utilities (TRN)	221, 481-488, 491-493	30-32, 391-400	7.0
Construction (CON)	23	33-45	4.9
Manufacturing (MAN)	3111-3116, 3118, 3119, 312-	46-70, 72-389	
	316, 321-327, 331-337, 339		1.5
Wholesale Trade (WHO)	42	390	2.3
Retail Trade (RET)	441-448, 451-454	401-412	11.6
Information (INF)	511-514	413-424	2.3
Financial Activities (FIN)	521-525, 531-533	425-436	4.8
Professional and Business Services (PRO)	541, 55, 561, 562	437-460	8.1
Education and Health Services (EUD)	611, 621-624	461-470	8.5
Leisure an Hospitality	711-713, 721, 722	471-481	9.3
Other Services (OSV)	811-814	482-494	4.0

TABLE 1. Alaskan Industry Categories for the Study(Employment Share = 1990-2003 Average)

Note: NA: not applicable

service (LEI), and other services (OSV). In IO2_VAR, OIL, SEA, and GOV are specified as economic base sectors and the other 12 sectors as nonbasic sectors.⁴ Deterministic monthly seasonal dummy variables were included in each of the models.⁵ Since the Alaska economy, as other state economies in the United States, depends on exports of goods and services (including oil and gas) to the rest of the United States (RUS), total energy-sector employment in RUS and total non-energy-sector employment in RUS are used as exogenous variables in each model.⁶

To derive the I-O information used in IO1_VAR and IO2_VAR, an Alaska SAM was constructed based on 2004 IMPLAN data for non-seafood industries and seafood industry data from diverse sources. A 15 sector SAM was developed based on IMPLAN data. IMPLAN data for seafood sectors was not used because these do not adequately characterize the Alaska seafood industry and fish harvesting employment and output is seriously underestimated. Instead, the

⁴ This study used the assignment method in which an industry is assigned into either a basic or nonbasic sector based on the analyst's knowledge of the regional economy. So we allocated each major industry relying on natural resources (Oil and Gas and Seafood in our study) to a basic sector. Since government (federal, state, and local combined) is an important economic base in Alaska, it was treated as a separate basic sector.

⁵ In this study, seasonal dummies are used along with a 12 month lag. The seasonal dummies are for the different mean levels of the different seasons, and do not necessarily correct for (stationary) autocorrelations at seasonal frequencies. (The different mean level implies non-stationarity.) After adjusting for the different seasonal mean levels, there are still autocorrelations at seasonal frequencies, especially at lag 12. There are many empirical studies published or unpublished where seasonal dummies are used along with autocorrelations at seasonal lags. See for example, Pierce (1978), Hylleberg et al. (1990), and Findley et al. (1998). Also, the X-11 procedure employed for seasonal adjustments, especially, at the US Census Bureau uses models that incorporate both seasonal dummies and (seasonal) lagged terms. It is a common practice in empirical work that seasonal dummies are used along with autocorrelations at seasonal lags.

⁶ National dummies are treated deterministically. This means that this study is based on Litterman's "circle-star" specification. In other words, it is assumed that the star (national) variables, including total US energy employment and total US non-energy employment, influence both star and circle (state) variables whereas circle variables influence only other circle variables.

seafood sectors were constructed independently using data from sources such as the Commercial Fisheries Entry Commission (CFEC), Alaska Fisheries Information Network (AKFIN), Pacific Fisheries Information Network (PacFIN), and National Marine Fisheries Service (NMFS). The Alaska Fisheries Economic Assessment Model (FEAM) was used to structure enterprise budgets, supply distributions, and regional spending profiles for vessels and processors engaged in Alaska fisheries. Industry expenditure budget items and net income were distributed to Alaska and other regions using information about the locations of industry suppliers. Budgets and trade flows were adjusted using results from interviews with key industry informants. Interviewees were asked to validate inferred information on their enterprise budgets, supplier relationships, seafood production and markets. Finally, the informal survey results were compared with seafood industry financial information, where available. Information developed for these sectors include catch and deliveries of fish, input purchases from non-fishery sectors, seafood commodity production, employment, labor earnings, capital income, tax payments, and other information needed to develop the SAM. For a more detailed description of procedures used to develop the seafood industry dataset, see The Research Group (2007). The sector aggregation scheme is shown in Table 1.

4. MODEL IMPLEMENTATION AND RESULTS

All models in this study are estimated with data on levels.⁷ Out-of-sample forecasts were conducted for five different lengths of forecast—1 month, 6 months, 12 months, 18 months, and 24 months. The first forecast period is 2004:1 and the last forecast period is 2005:12. We conduct rolling forecasts, in that we first estimate all five models through December 2003, and forecast over a *k*-month-ahead horizon. Then, we re-estimate the models incorporating one more month of the sample data, and forecasted for the next *k*-month-horizon. We continue this process until 2005:12 becomes the last forecast period. Thus, in this study, there are 24 one-step, 19 six-step, 13 twelve-step, 7 eighteen-step, and 1 twenty-four-step forecasts.⁸ Lags of 1 month and 12 months were chosen for UVAR model because of (1) the prominent auto-correlations shown at those lags, (2) the partial autoregression matrices that are significant at those lags, (3) information criteria that support the model, and (4) parsimony.⁹ As is often done, the same lags were used for all the other model versions.

⁷ Twelve of the fifteen variables (i.e., the employment series for fifteen industries) in this study do not contain a unit root. Only three variables for employment in Information, Education, and Other Services have a unit root. It is suspected that the test result for Information is influenced by a structural break. An analysis with structural break for the variable (Information) indicated no unit root. Although there is some evidence that the two remaining variables (Education and Other services) contain unit roots, it was decided to estimate all the models in levels. This is because even in the presence of a unit root in only a few of the variables, the least-squares estimators are known to be consistent (Ahn, 2007; Ahn and Reinsel, 1990).

⁸ We devised a separate VAR model to produce "future forecast values" of the two exogenous variables (total RUS energy sector employment and total RUS non-energy sector employment), and used these values to forecast the endogenous variables.

⁹ Akaike information criterion (AIC) and corrected AIC (AICc) were used to determine the lag length. Given relatively small number of observations compared with the number of variables on the right side of the equations, we decided not to include too many lags. Thus, we tried the following eleven lag specifications: (1) lags 1 and 3, (2) lags 1 and 4, (3) lags 1 and 6, (4) lags 1 and 12, (5) lags 1, 3, and 4, (6) lags 1, 2, and 6, (7) lags 1, 3, and 12, (8) lag 1, 4, and 12, (9) lags 1, 6, and 12, (10) lags 1, 3, 4, and 12, and (11) lags 1, 2, 6, and 12. AIC and AICc suggest that the lag structure with two lags specified in (1) to (4) is appropriate given the number of observations. We decided to use lags 1 and 12 in order to pick up the effects from the previous month and the previous year, and thus to represent a monthly and an annual effect while other lags are relatively harder to interpret. Although the AIC and AICc improve significantly with lags specified in (5) through (11) above, we decided not to use these lags since almost one third of the degrees of freedom (or more in case of four-lag cases) is used up for the parameter estimation. In the sensitivity analysis in this section, we examined how the results change if we use lags in (1), (2), and (3) above instead of lags 1 and 12. The seasonal periods of 2, 3, 4, and 6 months are sub-periods of 12, and thus the possible autocorrelations at these lags were given a special attention mainly because there are mild autocorrelations at these lags. In this

The forecasting accuracy was measured by root mean square errors (RMSEs). Results for forecasting performance are summarized in Tables 2 through 5. These tables summarize the results from 75 different individual forecasts for each model; i.e., five different forecasting periods for each of 15 industries. Tables 2 through 5 are based on RMSEs. Table 2 presents the number of industries that each model forecasts most accurately for each forecasting length. Table 3 presents the number of most accurate forecasts by industry over all forecast lengths.

Table 2 shows that, overall, MVAR produces largest number of most accurate forecasts, followed by UVAR, AR, IO2_VAR, and IO1_VAR; MVAR produces 18 most accurate forecasts out of a total of 75 forecasts. IO1_VAR produces the smallest number of most accurate forecasts. Looking at the forecasting performance by length of forecast, AR model produces the largest number of most accurate forecasts for 1-month-ahead forecast. For 6-month-ahead forecast, UVAR model produces the largest number of most accurate forecasts. For 12-month-ahead forecast, IO2_VAR perform the best in terms of number of most accurate forecasts. For 18-month-ahead forecast, MVAR produces the largest number of most accurate forecasts. For 24-month-ahead forecasts, IO2_VAR produces the largest number of most accurate forecasts.

Results indicate that in the short run (within one year), the models that do not have any I-O information (AR, UVAR, and MVAR) perform better than the models with I-O information (IO1_VAR, IO2_VAR) in terms of number of most accurate forecasts. In particular, IO2_VAR performs worst for 1-month-ahead and 6-month-ahead forecasts. On the other hand, results show that in the long run (after one year), MVAR and IO2_VAR perform much better than the other model versions in terms of number of most accurate forecasts. Results also indicate that overall IO1_VAR performs the worst in terms of the number of the most accurate forecasts.

Turning our attention to the number of most accurate forecasts by industry (Table 3), the AR model generates the most accurate forecasts for five industries (SEA, ONR, TRN, WHO, and RET). The IO2_VAR model generates the most accurate forecasts for four industries (SEA, GOV, MAN, and WHO). UVAR and MVAR produce the most accurate forecast for three industries, respectively, while IO1_VAR produces the most accurate forecast for only two

Forecast Length	AR	UVAR	MVAR	IO1_VAR	IO2_VAR
1 month	6*	3	3	3	0
6 months	2	8*	3	1	1
12 months	2	3	3	3	4*
18 months	1	1	7*	2	4
24 months	4*	2	2	3	4*
Total	15	17	18	12	13

TABLE 2. Number of Most Accurate Forecasts by Length of Forecast

* Denotes the model that produces the most number of accurate forecasts for the period

paper, sensitivity analysis was conducted for some of these sub-periods (i.e., lags of 1 and 3, 1 and 4, and 1 and 6). Also, we omitted insignificant lags between lag 1 and the maximum lag in order to avoid over-parameterization and degree of freedom problem.

Industry	AR	UVAR	MVAR	IO1_VAR	IO2_VAR
OIL	0	0	5*	0	0
SEA	2*	1	0	0	2*
GOV	0	1	0	0	4*
ONR	2*	1	1	1	0
TRN	3*	1	1	0	0
CON	0	2*	2*	1	0
MAN	0	0	0	2	3*
WHO	2*	0	1	0	2*
RET	3*	1	1	0	0
INF	0	0	2	3*	0
FIN	2	0	0	3*	0
PRO	1	1	1	1	1
EDU	0	5*	0	0	0
LEI	0	3*	1	0	1
OSV	0	1	3*	1	0
Total	15	17	18	12	13

 TABLE 3. Number of Most Accurate Forecasts by Industry

* Denotes the model that produces the most number of accurate forecasts for the industry

industries. As Table 3 shows, IO2_VAR, which has economic base information as priors, forecasts most accurately the employment of two of the three basic sectors (SEA and GOV). For nonbasic sectors, the model produces the most accurate forecasts for MAN and WHO. It is expected that models that incorporates interindustry relationships (UVAR, IO1_VAR, and IO2_VAR) perform comparatively better for industries which depend heavily on local economy (nonbasic industries). Indeed, as Table 3 shows, 34 out of a total of 42 most accurate forecasts by these three models are for nonbasic industries. On the other hand, models without I-O information (AR and MVAR) are expected to perform comparatively better for industries that are least dependent on the local economy. Oddly though, 26 out of a total of 33 most accurate forecasts by these two models (AR and MVAR) are nonbasic industries. In particular, 13 out of a total of 15 most accurate forecasts by AR are in nonbasic industries which include two trade industries (Wholesale Trade and Retail Trade, Table 3).

Results from Tables 2 and 3 show that, in terms of the number of most accurate forecasts, Minnesota prior information (MVAR) seems to slightly improve the forecasting performance compared with the first two models (AR and UVAR) that do not have any prior information. In fact, in terms of number of most accurate forecasts, MVAR performs the best; MVAR produces 18 most accurate forecasts while AR and UVAR produce 15 and 17 such forecasts, respectively. These results are somewhat consistent with those from previous studies (e.g., Partridge and Rickman, 1998).

Comparing models that use prior I-O information (IO1_VAR) with the first three models, it is seen that the I-O information in IO1_VAR does not improve the forecasting performance in terms of the number of the most accurate forecasts (Table 2). In fact, IO1_VAR produces the smallest number of the most accurate forecasts (12 out of 75 forecasts, Tables 2 and 3). Also, the results in Tables 2 and 3 do not provide any strong evidence that IO1_VAR perform better either in the short run (within 1 year) or in the long run (after 1 year) than the first three model

variants, which do not have any I-O information, in terms of the number of the most accurate forecasts. These results are in contrast with those in some previous studies (e.g., Magura, 1990; LeSage and Magura, 1991; and Partridge and Rickman, 1998) which report that models with I-O information specified as prior perform better in the long run than those with no I-O information.

Results also show that IO2_VAR produces smaller number of most accurate forecasts than the first three models; IO2_VAR produces 13 most accurate forecasts out of a total of 75 forecasts. However, IO2_VAR performs significantly better in the long run in terms of the number of the most accurate forecasts; the model produces 4 most accurate forecasts for the last three forecast periods (12-month-ahead, 18-month-ahead, and 24-month-ahead). This means that the economic base information incorporated in the f(i, j) matrix used in IO2_VAR improves the forecasting capability in the long run in terms of the number of most accurate forecasts.

Table 4 presents average RMSE across industries for each model by length of forecast. The results in Table 4 indicate that IO2_VAR produces on average the most accurate forecasts in three of all five forecast periods. The average RMSEs in IO2_VAR are 0.509 (12-month-ahead), 0.698 (18-month-ahead), and 0.513 (24-month-ahead). For 1-month-ahead forecasts, the AR model produces on average the most accurate forecast. For other forecast periods, the AR model performs worse than IO2_VAR.

Table 4 demonstrates that MVAR model performs the worst in terms of average RMSE; in all five forecast periods, the average RMSEs are the largest among the five model variants. This result is contrasted with the results in Table 2 which shows that the model (MVAR) produces the largest number of most accurate forecasts. Table 4 also shows that the reducedform employment information in IO1_VAR improves to some extent the average forecasting performance compared with UVAR and MVAR, but IO1_VAR performs worse on average than AR. Table 4 provides strong evidence that the economic base information incorporated in f(i, j)matrix in IO2_VAR indeed improves the average forecasting accuracy, especially in the long run.

Although not reported in this paper, mean absolute errors (MAEs) were also calculated as an alternative measure of forecasting accuracy. However, this study found that use of MAEs does not change the major findings and conclusions in the above. One notable difference between the results from using the two alternative measures of forecasting accuracy is that, if MAE is used, UVAR produces the largest number of most accurate forecasts; the total number of most accurate forecasts by the model (UVAR) is 19 while the number is 17 (Table 2), if RMSE is used.

TABLE 4. Average RMSE by Length of Forecast								
Length of	AR	UVAR	MVAR	IO1_VAR	IO2_VAR			
Forecast								
1 month	0.287*	0.299	0.307	0.305	0.302			
6 months	0.548	0.501*	0.591	0.589	0.540			
12 months	0.564	0.620	0.667	0.563	0.509*			
18 months	0.954	1.035	1.065	0.908	0.698*			
24 months	0.907	1.043	1.570	1.026	0.513*			
Average	0.652	0.699	0.840	0.678	0.513*			

* Denotes the model that produces on average the most accurate forecast for the step (period)

	AR	UVAR	MVAR	IO1_VAR	IO2_VAR	Total
AR	NA	4	5	3	1	13
UVAR	1	NA	5	2	2	10
MVAR	0	0	NA	0	0	0
IO1_VAR	2	3	5	NA	0	10
IO2_VAR	4	3	5	5	NA	17

TABLE 5. Pairwise Comparisons of Forecast Models

We also derived the reduced-form employment I-O relationships using only IMPLAN data, and ran IO1_VAR model with these I-O relationships as priors. Although the detailed results are not shown in a table, this study found out that the average RMSE for each forecast length with priors based on only IMPLAN data is larger than that generated with the priors based on the hybrid data. While the average RMSE across all the forecast periods from the ready-made model (which uses only IMPLAN data) is 0.715, the RMSE from the hybrid model (which uses both IMPLAN and local data) is 0.678 (last row, Table 4). Results indicate that it is necessary to use local data in forecasting employment and economic impact analysis.

As in Rickman, Miller, and McKenzie (2009), we made pairwise comparisons of the forecasting performance of the five models (Table 5). The results in the table are based on the average RMSEs in Table 4. In Table 5, number in each cell represents the number of wins that the model in the row has over the corresponding model in the column. The pairwise comparison measure is a measure in which the influence of outliers on the average errors is eliminated. For example, the number four in AR row and UVAR column in the table means that in four of five forecast periods, AR performs better than UVAR. Thus, the number in row *i* and column *j* plus the number in row *j* and column *i* equals the total number of forecast periods. So there are a total of 50 pairwise comparisons (10 for each forecast period). It is seen from Table 5 that the IO2_VAR, which has the smallest average RMSE (0.513, Table 4), produces the largest number (17) of head-to-head wins against the other models. This, again, reinforces the earlier conclusion that overall IO2_VAR performs the best. Also, AR produces the next largest number of head-to-head wins, which is consistent with the results in Table 4. It is notable that MVAR produces no head-to-head wins against any of the other models.

We also made pairwise comparisons of the forecasting performance of the five models for two forecast time horizons (12-month-ahead and 24-month-ahead forecasts, Table 6). The results in the table are based on the average RMSEs for 12-month-ahead and 24-month-ahead forecasts. In Table 6, number in each cell represents how many industries the model in the row most accurately predicts versus each of the other models in the column. For example, the number nine in AR row and UVAR column in the top pane in the table means that for nine of fifteen industries, AR performs better than UVAR in the 12-month-ahead forecast. It is seen from Table 6 that, in case of 12-month-ahead forecasts, AR model produces the largest number (34) of head-to-head wins against the other models in terms of the number of industries with the most accurate forecasts. The next largest numbers (33 and 32) of head-to-head wins are generated by IO1_VAR and IO2_VAR, respectively. Looking at the results for 24- month-ahead forecasts, however, the largest number of head-to-head wins are produced by IO2_VAR (43), followed by AR (37) and IO1_VAR (34). This implies that, for a long-run forecast (24-month-ahead forecast), IO2_VAR performs significantly better than any other models, further supporting the earlier conclusion that overall IO2_VAR performs the best. A sensitivity analysis was conducted

-	12-month-ahead forecast						
	AR	UVAR	MVAR	IO1_VAR	IO2_VAR	Total	
AR	NA	9	9	7	9	34	
UVAR	6	NA	7	8	6	27	
MVAR	6	8	NA	4	6	24	
IO1_VAR	8	7	11	NA	7	33	
IO2_VAR	6	9	9	8	NA	32	
-		24	l-month-ahead f	orecast			
	AR	UVAR	MVAR	IO1_VAR	IO2_VAR	Total	
AR	NA	9	12	9	7	37	
UVAR	6	NA	9	4	3	22	
MVAR	3	6	NA	2	3	14	
IO1_VAR	6	11	13	NA	4	34	
IO2_VAR	8	12	12	11	NA	43	

TABLE 6. Pairwise Comparisons of Models for Most Accurate Forecasts

TABLE 7. Results from Sensitivity Analysis for Lags(average RMSE by length of forecast)

Forecast Length	AR	UVAR	MVAR	IO1_VAR	IO2_VAR		
Lags 1 and 3							
1 month	0.306	0.292	0.291*	0.300	0.296		
6 months	0.616	0.510*	0.568	0.597	0.549		
12 months	0.617	0.631	0.689	0.620	0.580*		
18 months	0.878	1.087	1.090	0.979	0.758*		
24 months	0.795	1.536	1.536	1.168	0.655*		
Average	0.642	0.811	0.835	0.733	0.568*		
		Lags 1	and 4				
1 month	0.311	0.309	0.296*	0.303	0.299		
6 months	0.621	0.535*	0.568	0.596	0.545		
12 months	0.617	0.659	0.687	0.618	0.560*		
18 months	0.885	1.246	1.096	0.975	0.755*		
24 months	0.764	1.643	1.536	1.140	0.657*		
Average	0.639	0.878	0.837	0.726	0.563*		
		Lags 1	and 6				
1 month	0.312	0.313	0.311	0.311	0.305*		
6 months	0.615	0.472*	0.608	0.601	0.557		
12 months	0.556	0.591	0.683	0.579	0.522*		
18 months	0.868	0.863	1.080	0.915	0.700*		
24 months	0.675	1.510	1.533	1.020	0.506*		
Average	0.605	0.750	0.843	0.685	0.518*		
		Lags 1 a	and 12				
1 month	0.287*	0.299	0.307	0.305	0.302		
6 months	0.548	0.501*	0.591	0.589	0.540		
12 months	0.564	0.620	0.667	0.563	0.509*		
18 months	0.954	1.035	1.065	0.908	0.698*		
24 months	0.907	1.043	1.570	1.026	0.513*		
Average	0.652	0.699	0.840	0.678	0.513*		

* indicates the model that produces on average the most accurate forecast for the step (period).

to examine how the forecasting accuracy changes when different lags are specified.¹⁰ We estimated the five different forecasting models with each of three different specifications of the lags—(1) lags 1 and 3, (2) lags 1 and 4, and (3) 1 and 6. Results are shown in Table 7. Table 7 shows that with (1) lags 1 and 3 and (2) lags 1 and 4, MVAR generates on average the most accurate forecasts for 1-month-ahead forecast and UVAR performs the best for 6-month-ahead forecast. For all the other forecast periods (i.e., 12 month, 18 month, and 24 month), IO2_VAR with these two lag specifications (i.e., lags 1 and 3 and lags 1 and 4) produces on average the most accurate forecasts. With lags 1 and 6, IO2_VAR produces on average the most accurate forecast periods except for 6-month-ahead forecast. Results in Table 7 strongly support the major conclusion of this study that the economic base information in IO2_VAR indeed improves the average forecasting accuracy in the long run.

5. CONCLUSIONS

This study applied a BVAR methodology to forecasting Alaska industry employment. This study found that, in terms of the number of most accurate forecasts, Minnesota prior information (MVAR) seems to improve slightly the forecasting performance compared with the first two models (AR and UVAR) that do not include any prior information. These results are somewhat consistent with those in previous studies (e.g., Partridge and Rickman, 1998). In the short run, the models that do not have any I-O information (AR, UVAR, and MVAR) perform better than the models with I-O information (IO1_VAR, IO2_VAR) in terms of number of most accurate forecasts. However, in the long run, MVAR and IO2_VAR perform significantly better than the other model versions in terms of number of most accurate forecasts.

The I-O information in IO1_VAR does not improve the forecasting performance in terms of the number of the most accurate forecasts compared to the first three model variants (AR, UVAR, and MVAR) either in the short run or in the long run. These results are in contrast with those in some previous studies (e.g., Magura, 1990; LeSage and Magura, 1991; and Partridge and Rickman, 1998) which report that models with I-O information specified as prior perform better in the long run than those with no I-O information. Most importantly, the IO2_VAR performs the best in the long run in terms of the number of the most accurate forecasts and in terms of average forecasting performance. This implies that the economic base information in the model version indeed improves the forecasting capability in the long run. These findings indicate that, if the economy in question (such as Alaska economy) depends to a large extent on natural resource as an economic base, standard I-O-based BVAR models, which may be suitable for manufacturing-based economies, do not perform as well as a model that incorporates as prior the economic base relationships between natural resource economic base sectors and nonbasic sectors.

A natural extension of the approach used in this study is to consider the linkages between the state of Alaska and the RUS. Many of the primary and intermediate inputs used in Alaska

¹⁰ We also examined how the results change in response to change in the overall tightness parameter (θ) in Equation (3) above, for MVAR, IO1_VAR, and IO2_VAR. In the MVAR model, the decrease in θ from 0.2 to 0.1 results in increase in overall forecast accuracy for 1-month-ahead, 12-month-ahead, and 24-month-ahead forecasts. On the other hand, when θ increases from 0.2 to 0.4, the overall forecast accuracy increases in all forecast periods. In IO1_VAR model, the reduction in θ results in an increase in overall forecast accuracy in all forecast periods except for 1-month-ahead forecast as the IO properties in IO1_VAR model are weighted more heavily. In IO2_VAR model, the forecast accuracy decreases in all forecast periods with the decrease in θ . When θ increases in IO2_VAR, the accuracy increases in 1-month-ahead, 6-month-ahead, and 12-month-ahead forecasts, whereas the accuracy decreases in the other forecast steps. Overall, change in θ in the Bayesian models does not lead to a significant change in the average forecast accuracies except for 24-month-ahead forecasts in IO1_VAR.

industries are imported from the RUS, especially from the state of Washington. Therefore, a future study would be to develop a multiregional model which considers the linkages between industries in Alaska and those in other states as well as the linkages between industries within Alaska.

REFERENCES

- Ahn, Sung K. (2007) Report on the Development of Dynamic Economic Base Models for Alaska Fisheries. Project Report to Alaska Fisheries Science Center.
- Ahn, Sung K. and Gregory C. Reinsel. (1990) "Estimation for Partially Nonstationary Multivariate Autoregressive Models," *Journal of the American Statistical Society* Association, 85, 813–823.
- Amirizadeh, Hossain and Richard M. Todd. (1984) "More Growth Ahead for Ninth District States," *Federal Reserve Bank of Minneapolis Quarterly Review*, 8 (Fall), 8–17.
- Bewley, Ronald. (2002) "Forecast Accuracy, Coefficient Bias and Bayesian Vector Autoregressions," *Mathematics and Computers in Simulation*, 59, 163–169.
- Bischoff, Charles W., Haleform Belay, and In-Bong Kang. (2000) "Bayesian VAR Forecasts Fail to Live Up to Their Promise," *Business Economics*, 35(3), 19–29.
- Carlino, Gerald A., Robert H. DeFina, and Keith Sill. (2001) "Sectoral Shocks and Metropolitan Employment Growth," *Journal of Urban Economics*, 50, 396–417.
- Chang, Sheng-Wen and N. Edward Coulson. (2001) "Sources of Sectoral Employment Fluctuations in Central Cities and Suburbs: Evidence from Four Eastern U.S. Cities," *Journal of Urban Economics*, 49, 199–218.
- Coulson, N. Edward. (1993) "The Sources of Sectoral Fluctuations in Metropolitan Areas," *Journal of Urban Economics*, 33, 76–94.
- Doan, Thomas, Robert B. Litterman, and Christopher Sims. (1984) "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.
- Dua, Pami and Subhash C. Ray. (1995) "A BVAR Model for the Connecticut Economy." *Journal of Forecasting*, 14, 167–180.
- Findley, David F., Brian C. Monsell, William R. Bell, Mark C. Otto, and Bor-Chung Chen. (1998) "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program," *Journal of Business Economics and Statistics*, 16, 127–152.
- Fullerton, Thomas M. Jr. (2001) "Specification of a Borderplex Econometric Forecasting Model," *International Regional Science Review*, 24, 245–260.
- Gupta, Rangan and Sonali Das. (2008) "Spatial Bayesian Methods of Forecasting House Prices in Six Metropolitan Areas of South Africa," *South African Journal of Economics*, 76, 298–313.
- Hylleberg, Svend, Robert F. Engle, Clive W. J. Granger, and B. S. Yoo. (1990) "Seasonal Integration and Cointegration," *Journal of Econometrics*, 44, 215–238.

- LeSage, James P. and Michael Magura. (1991) "Using Interindustry Input-Output Relations as a Bayesian Prior in Employment Forecasting Models," *International Journal of Forecasting*, 7, 231–238.
- LeSage, James P. and J. David Reed. (1989) "The Dynamic Relationship between Export, Local and Total Area Employment," *Regional Science and Urban Economics*, 19, 615–636
- Litterman, Robert B. (1980) "A Bayesian Procedure for Forecasting with Vector Autoregressions," Working Paper, Department of Economics, Massachusetts Institute of Technology.
- _____. (1986) "Forecasting With Bayesian Vector Autoregressions–Five Years of Experience," Journal of Business & Economic Statistics, 4, 25–38.
- Magura, Michael. (1990) "Using Input-Output Data in a Bayesian Autoregressive Forecasting Model," in Luc Anselin and Moss Madden, eds, *New Directions in Regional Analysis: Multiregional Approaches*. Bellhaven Press: London, pp. 133–145.
- _____. (1998) "IO and Spatial Information as Bayesian Priors in an Employment Forecasting Model," *Annals of Regional Science*, 32, 495–503.
- Minnesota IMPLAN Group, Inc. (2001) IMPLAN Pro User's Guide.
- Partridge, Mark D. and Dan S. Rickman. (1998) "Generalizing the Bayesian Vector Autoregression Approach for Regional Interindustry Employment Forecasting," *Journal* of Business and Economic Statistics, 16, 62–72.
- Pierce, David A. (1978) "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality Are Present," in Arnold Zellner (ed.), Seasonal Analysis of Economic Time Series. U.S. Census Bureau, pp. 242–280.
- Puri, Anil and Gokce Soydemir. (2000) "Forecasting Industrial Employment Figures in Southern California: A Bayesian Vector Autoregressive Model," Annals of Regional Science, 34, 503–514.
- Rickman, Dan S. (2001) "Using Input-output Information for Bayesian Forecasting of Industry Employment in a Regional Econometric Model," *International Regional Science Review*, 24, 226–244.
- _____. (2002) "A Bayesian Forecasting Approach to Constructing Regional Input-Output Based Employment Multipliers," *Papers in Regional Science*, 81, 483–498.
- Rickman, Dan S., Steven R. Miller, and Russell McKenzie. (2009) "Spatial and Sectoral Linkages in Regional Models: A Bayesian Vector Autoregression Forecast Evaluation," *Papers in Regional Science*, 88, 29–41.
- Theil, Henri. (1963) "On the Use of Incomplete Prior Information in Regression Analysis," *Journal of the American Statistical Association*, 58, 401–414.
- The Research Group. (2007) *Estimating Economic Impacts of Alaska Fisheries* Using a Computable General Equilibrium Model – Data Acquisition and Reduction Task Documentation. Report Prepared for Alaska Fisheries Science Center. Corvallis, Oregon.
- Todd, Richard M. (1984) "Improving Economic Forecasting With Bayesian Vector Autoregression," *Federal Reserve Bank of Minneapolis Quarterly Review*, 8, 18–29.