Investigating the effects of ageing population on economic growth and healthcare expenditure for the US: A Bayesian-VAR

Tian Cater[a]

[a] *Advanced Time Series Econometrics 872 Project*

[b] *15 January 2023*

1. Introduction

Healthcare expenditure is positively correlated with economic growth. Specifically, the demand for healthcare services increases in response to economic growth. (Fogel, 2004; Lopreite & Mauro, 2017; Wang, 2011). Over time, improved healthcare provisions increase life expectancy, requiring more healthcare financing to sustain the larger proportion of elderly individuals. In particular, the US has experienced a considerable shift in its demographic, a result of higher life expectancy combined with lower rates of fertility (Linden & Ray, 2017; Wiener & Tilly, 2002; Williams, Cylus, Roubal, Ong, Barber, Organization & others, 2019). As this disparity in demographics continues to widen, the increased healthcare costs will not be able to be offset by economic growth as the labour force starts to diminish relatively (Cheng, Yang, Schwebel, Liu, Li, Cheng, Ning & Hu, 2020).

To this end, this project aims to investigate the relationships between an ageing population, healthcare expenditure, life expectancy, and economic growth in the US empirically by applying Bayesian vector autoregression (BVAR) techniques.

Impulse response analysis shows that the ageing population in the US not only negatively impacts economic growth but it also reverses the generally accepted relationship between health expenditure and economic growth; An increase in health expenditure exerts both short and long-run downward pressure on economic growth as the additional spending fails to generate more productivity.

Moreover, the findings show that, contrary to common belief, an increase in the gap between elderly and young individuals (the ageing index) causes persistent long-run reductions in life expectancy, indicating that the lower fertility rates have a more significant impact than the lower mortality rates amongst the elderly.

Email address: 19025831@sun.ac.za (Tian Cater)

The rest of this paper is structured as follows. Section 2 provides a brief review of the existing literature and underscores the contributions of this project. Section 3 presents the BVAR methodology and the subsequent priors used in estimating the model. Sections 4.1 and 5 discuss the sample data and its stationarity, respectively. Lastly, Section 6 specifies the selected prior distribution's features and justification, provides the estimated results and conducts impulse response analysis and forecasting.

2. Literature Review

Various studies have investigated economic growth's one-way relationship with healthcare expenditure. Murillo, Piatecki & Saez (1993) shows that the variation in healthcare spending is significantly affected by GDP per capita for the nineteen OECD countries. On the other hand, Baltagi & Moscone (2010) also uses data from 1973 through 2006 for the twenty OECD countries and finds that relatively high healthcare expenditure is uncommon, even during abnormally high economic growth. Braendle & Colombier (2016) is another study where healthcare spending and income per capita are shown to be positively correlated.

Furthermore, other literature investigates the bivariate causality between economic growth and expenditure. Wang (2011), for example, finds that economic growth is promoted by healthcare spending between 1991 through 2010 using data on thirty-one countries. Complementarily, Amiri & Ventelou (2012) shows for OECD countries that the two-way relationship between economic growth and healthcare expenditure maintains Granger causality. Also, using data from twenty OECD countries from 1981 through 2015, Amiri & Linden (2016) shows that this bilateral association between healthcare spending and GDP is evident in more than three-quarters of the OECD countries considered. Studies also show that the reciprocal affiliation between these two variables is apparent in low-middle and high-income countries Halıcı-Tülüce, Doğan & Dumrul (2016).

The second widely investigated relationship in the literature is between healthcare spending, population ageing (or life expectancy), and economic growth. For example, Murthy & Okunade (2016) finds, using an ARDL model for data from 1958 to 2008 for the U.S., that advancements in healthcare technology, real income per capita, and the proportion of the population over 65 are the leading causes of higher per capita healthcare spending. Jaba, Balan & Robu (2014) used panel data for 110 nations between 1998 and 2012 and found that life expectancy and health spending are significantly positively correlated. This positive relationship is iterated by Linden & Ray (2017) for twenty-eight OECD countries for a sample period between 1985 and 2015.

Estimating a BVAR model for Italy for 1992 to 2014, Lopreite & Mauro (2017) indicates that healthcare expenditure reacts less to economic growth than changes in the ageing index. Lopreite & Zhu (2020) also estimates a BVAR and compares the effect of an ageing population on health spending and GDP between China and the U.S. Their findings show that China's ageing population is a much more

dire concern for economic growth prospects compared to the U.S. The mutual relationship between an ageing population and economic growth is driven by the age structure of the population, which impacts the speed at which economic growth is elicited, whereas this slower growth prospect, in turn, results in demographic changes (for example, Lee, Mason & Miller (2000), Bloom, Canning & Fink (2010), and Liotta, Canhao, Cenko, Cutini, Vellone, Illario, Kardas, Poscia, Sousa, Palombi & others (2018))

This project contributes to the literature in at least two ways. First, in a similar fashion to Lopreite & Zhu (2020), I investigate the relationship between an ageing index, healthcare spending, economic growth, and life expectancy for the U.S. using a BVAR model. The Bayesian approach is specifically appropriate in this case as the sample is small (consisting of only annual data) and benefits from added predictive accuracy compared to the classic VAR model (Chan, 2017).

Second, I estimate and compare the BVAR model for both Minnesota prior and the conditional normal inverse-Wishart prior, incorporating prior beliefs on parameter distributions. [1] Thus, the estimation enables comparing the short and medium-term impulse response analyses and forecasting performance for both adopted priors.

3. Methodology

I start by considering the following (reduced form) VAR(p) model with $\boldsymbol{y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ represents a vector of dependent variables at time t:

$$\mathbf{y}_t = \boldsymbol{b} + \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + ... + \boldsymbol{A}_p \boldsymbol{y}_{t-p} + \boldsymbol{\varepsilon}_t, \tag{3.1}$$

where $\boldsymbol{\varepsilon}_t \sim N(0, \boldsymbol{\Sigma})$ is the $n \times 1$ white noise error vector with covariance matrix $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Sigma}$, $\boldsymbol{b}$ is a $n \times 1$ vector of intercepts, and $\boldsymbol{A}_1, .., \boldsymbol{A}_p$ is the $n \times n$ coefficient matrices for the p lags, respectively. That is, the VAR(p) model (3.1) is a multiple equation regression with explained variables entering as the lagged explanatory variables. More generally, one can rewrite the VAR(p) as:

$$\boldsymbol{y}_t = \boldsymbol{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t, \tag{3.2}$$

---

[1]Lopreite & Mauro (2017) used a BVAR and compared the performance of the Minnesota and normal inverse-Wishart prior for Italy, whereas I use the conditional normal inverse-Wishart prior.

where $\boldsymbol{X}_t = \boldsymbol{I}_n \otimes (1, \boldsymbol{y}'_{t-1}, ..., \boldsymbol{y}'_{t-p})$ and $\boldsymbol{\beta} = vec([\boldsymbol{b}, \boldsymbol{A}_1, ..., \boldsymbol{A}_p])$. Then stack the observations over $t = 1, ..., T$ to get:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{3.3}$$

where $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{I}_T \otimes \boldsymbol{\Sigma})$ and $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_T)'$ is a $Tn \times nk$ matrix.

The goal of Bayesian methods is to obtain the posterior distribution that summaries all the information about the parameter vector given the data. That is, to estimate the parameters of the model ($\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$) one needs to apply Bayes rule:

$$p(\boldsymbol{\beta}, \boldsymbol{\Sigma}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\Sigma})p(\boldsymbol{\beta}, \boldsymbol{\Sigma})}{p(\boldsymbol{y})} \tag{3.4}$$

$$\propto p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\Sigma})p(\boldsymbol{\beta}, \boldsymbol{\Sigma}), \tag{3.5}$$

where $p(\boldsymbol{\beta}, \boldsymbol{\Sigma}|\boldsymbol{y})$ is the joint posterior distribution , $p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\Sigma})$ the likelihood function, and $p(\boldsymbol{\beta}, \boldsymbol{\Sigma})$ the joint prior distribution.

3.1. The Minnesota Prior

The first BVAR estimation imposes the Minnesota or Litterman prior proposed by Litterman (1986) and Doan, Litterman & Sims (1984). The reasons why I adopt the Minnesota prior are fourfold: firstly, the analysis uses three of the four series in levels, and the Minnesota prior's robustness is optimal in this case; secondly, using the levels of the series, it is possible to analyse, through impulse response analyses and forecasts, the short and medium term dynamic relationships among the variables; thirdly, the estimated coefficients of the unrestricted VAR model is shrunk away from its OLS estimated and towards the prior mean by the Minnesota prior, thereby generating estimation gains. Lastly, the Minnesota prior has been shown to perform very well when considering smaller samples, as is the case here.

The Minnesota prior is a shrinkage prior that holds $\boldsymbol{\Sigma}$ fixed to a data-based approximation for posterior sampling of $\boldsymbol{\beta}$, which simplifies prior elicitation and computation. That is, it replaces $\boldsymbol{\Sigma}$ with an estimate $\hat{\boldsymbol{\Sigma}}$, and assumes $\boldsymbol{\Sigma}$ is a diagonal matrix:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} \tag{3.6}$$

, with $\hat{\sigma_{ii}^2} = s_i^2$, the OLS estimate of the error variance in the ith equation.

The Minnesota prior assumes, for $\boldsymbol{\alpha} = vec(\boldsymbol{\beta})$, that:

$$\boldsymbol{\alpha} \sim N(\bar{\boldsymbol{\alpha}}, \boldsymbol{\Phi}_\alpha), \tag{3.7}$$

where the prior mean of the coefficients $\bar{\boldsymbol{\alpha}}$ is a $kn \times 1$ vector of zeros, except for elements that relate to the first order own-lag terms. Let $i, j \in \{1, ..., n\}$ where equations and variables are indexed by i and j, respectively. Then, the construction of the prior variance matrix $\boldsymbol{\Phi}_\alpha$ for $\boldsymbol{\beta}$ can be specified for a general form of the decay function $d(l)$ as in Canova (2007):

$$\boldsymbol{\Phi}_{\alpha(i,j)}(l) = \begin{cases} \frac{V_1}{d(l)} \\ \frac{V_1 V_2 \sigma_j^2}{d(l)\sigma_i^2} \\ V_1 V_3, \end{cases} \tag{3.8}$$

corresponding to own lags, cross variable lags, and exogenous variables respectively.

The hyperparameters V_1, V_2, V_3, and V_4 govern the diagonal variance-covariance matrix. These hyperparameters has the following characteristics:

(i) V_1 dictates the comparative significance of the sample and prior information, indicating variance of the first lag's overall tightness. Therefore, applying a modest value entails the prior dominating the sample information. On the other hand, the prior information becomes uninformative as V_1 tends to infinity and the estimates of the posterior converge towards the VAR coefficients.

(ii) V_2 establishes the comparative significance for the lags of variables on variables other than itself. The comparative significance of the information encompassed in the exogenous variables is controlled by V_3.

(iii) V_4 The decay function $d(l)$ is indexed by the hyperparameter $V_4 > 0$, where:

- if $d(l) = l^{V_4}$, then there is harmonic decay.

- if $d(l) = V_4^{-l+1}$, the geometric decay,

- and if $V_1 = 1$, then linear decay.

In this project, I use the specification by Koop & Korobilis (2010) where $d(l) = l^2$ and the ratio of the variance of each equation has been inverted for the cross-equation coefficients, thereby setting the hyperparameter $V_4 = 2$:[2]

$$\boldsymbol{\Phi}_{\alpha(i,j)}(l) = \begin{cases} \frac{V_1}{l^2} \\ \frac{V_2 \sigma_i^2}{l^2 \sigma_j^2} \\ V_3 \sigma_i^2. \end{cases} \tag{3.9}$$

Finally, with $\boldsymbol{\Sigma}$ fixed, the posterior distribution of $\boldsymbol{\alpha} = vec(\boldsymbol{\beta})$ is given by:

$$p(\boldsymbol{\alpha}|\boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{y}) \propto N[\boldsymbol{\Phi}_\alpha^{-1} + (\boldsymbol{\Sigma^{-1}} \otimes \boldsymbol{X}^T \boldsymbol{X}) , \boldsymbol{\Phi}_\alpha^{-1}\bar{\alpha} + vec(\boldsymbol{X}^T \boldsymbol{y} \boldsymbol{\Sigma}^{-1})] \tag{3.10}$$

3.2. The Conditional Normal Inverse-Wishart Prior

The priors take the following form

$$p(\boldsymbol{\Sigma}) \sim IW[\frac{1}{\boldsymbol{\gamma}}\boldsymbol{\Phi}_\sigma^{-1}, (\gamma)], \tag{3.11}$$

$$p(\boldsymbol{\alpha}|\boldsymbol{\Sigma}) \sim N[vec(\bar{\boldsymbol{\beta}}), \boldsymbol{\Sigma} \otimes (\Phi)_\beta], \tag{3.12}$$

where the location matrix $\boldsymbol{\Phi}_\Sigma$ is formed in an identical way to the Minnesota prior; as the residual variance from equation-by-equation estimation of AR(p) models. The scaling of those parameters are governed by the hyperparameters V_1 and V_3, where:

---

[2]The selection and justification of the remaining hyperparameters is discussed later in section 6 after the data and stationarity concerns have been analysed.

$$\mathbf{\Phi}_{\beta(i,i)}(l) = \begin{cases} \frac{V_1}{l^2 \sigma_i^2} \\ V_1 V_3, \end{cases} \tag{3.13}$$

for lags and exogenous variables respectively. The conditional posterior distributions are given by:

$$p(\mathbf{\Sigma}|\mathbf{X}, \mathbf{y}) \sim IW[\frac{\gamma}{\mathbf{\Phi}_{\Sigma}} + \bar{\boldsymbol{\beta}}^T \mathbf{\Phi}_\beta^{-1} \bar{\boldsymbol{\beta}} + \mathbf{y}^T \mathbf{y} - \tilde{\boldsymbol{\beta}}^T \tilde{\mathbf{\Sigma}}_\beta^{-1} \boldsymbol{\beta}^2 \; , \; n - p + \boldsymbol{\gamma}], \tag{3.14}$$

$$p(\boldsymbol{\alpha}|\mathbf{\Sigma}) \sim N[vec(\bar{\boldsymbol{\beta}}), \mathbf{\Sigma} \otimes \mathbf{\Phi}_\beta], \tag{3.15}$$

where

$$\tilde{\mathbf{\Sigma}}_\beta^{-1} = \mathbf{\Phi}_\beta^{-1} + \mathbf{X}^T \mathbf{X}, \tag{3.16}$$

$$\tilde{\boldsymbol{\beta}} = \tilde{\mathbf{\Sigma}}_\beta (\mathbf{\Phi}_\beta^{-1} \bar{\boldsymbol{\beta}} + \mathbf{X}^T \mathbf{y}). \tag{3.17}$$

3.3. The Gibbs Sampler

I then apply an Gibbs sampling algoritm for the VAR(p) to derive the posterior densities. The alorithm is as follows: Pick some initial values $\boldsymbol{\alpha}^{(0)} = \boldsymbol{c}_0$ and $\mathbf{\Sigma}^{(0)} = \boldsymbol{C}_0 > 0$, then repeat the following steps from $r = 1$ to R :

- Draw $\boldsymbol{\alpha}^{(r)} \sim p(\boldsymbol{\alpha}|\mathbf{X}, \mathbf{y}, \mathbf{\Sigma}^{(r-1)})$ (Multivariate Normal)

- Draw $\mathbf{\Sigma}^{(r)} \sim p(\mathbf{\Sigma}|\mathbf{X}, \mathbf{y}, \boldsymbol{\alpha}^{(r)})$ (Inverse-Wishart)

4. The Data

The data consists of annual U.S. data sampled over the period stretching from the start of 2016 to the end of 2020, gathered from the Federal Reserve Bank of St. Louis' FRED database. The time series are plotted in Figure 4.1 and consist of nominal Gross Domestic Product (GDP) per capita, healthcare expenditure per capita (HE), life expectancy at birth (LE), and a self-constructed ageing index (AI). [3]

---

[3]As will be evident in section 5, nominal GDP growth per capita will be log difference to render it stationary, therby representing its annual growth.

Healthcare expenditure is the total government's current expenditures on health (in Dollars), transformed to per capita terms using the age 16 and over civilian non-institutional population (HE). The ageing index (AI) is constructed similarly to Lopreite & Mauro (2017), as the logged ratio between the number of civilians aged 65 and over and the number of civilians aged 14 and less. The AI shows a persistent rise over time as the elderly population increases (numerator) while the young population decreases (denominator). Lastly, the life expectancy at birth (LE) series is defined as the average number of years a newborn is anticipated to live if, at the time of birth, the mortality rate remains constant throughout the newborn's lifetime.
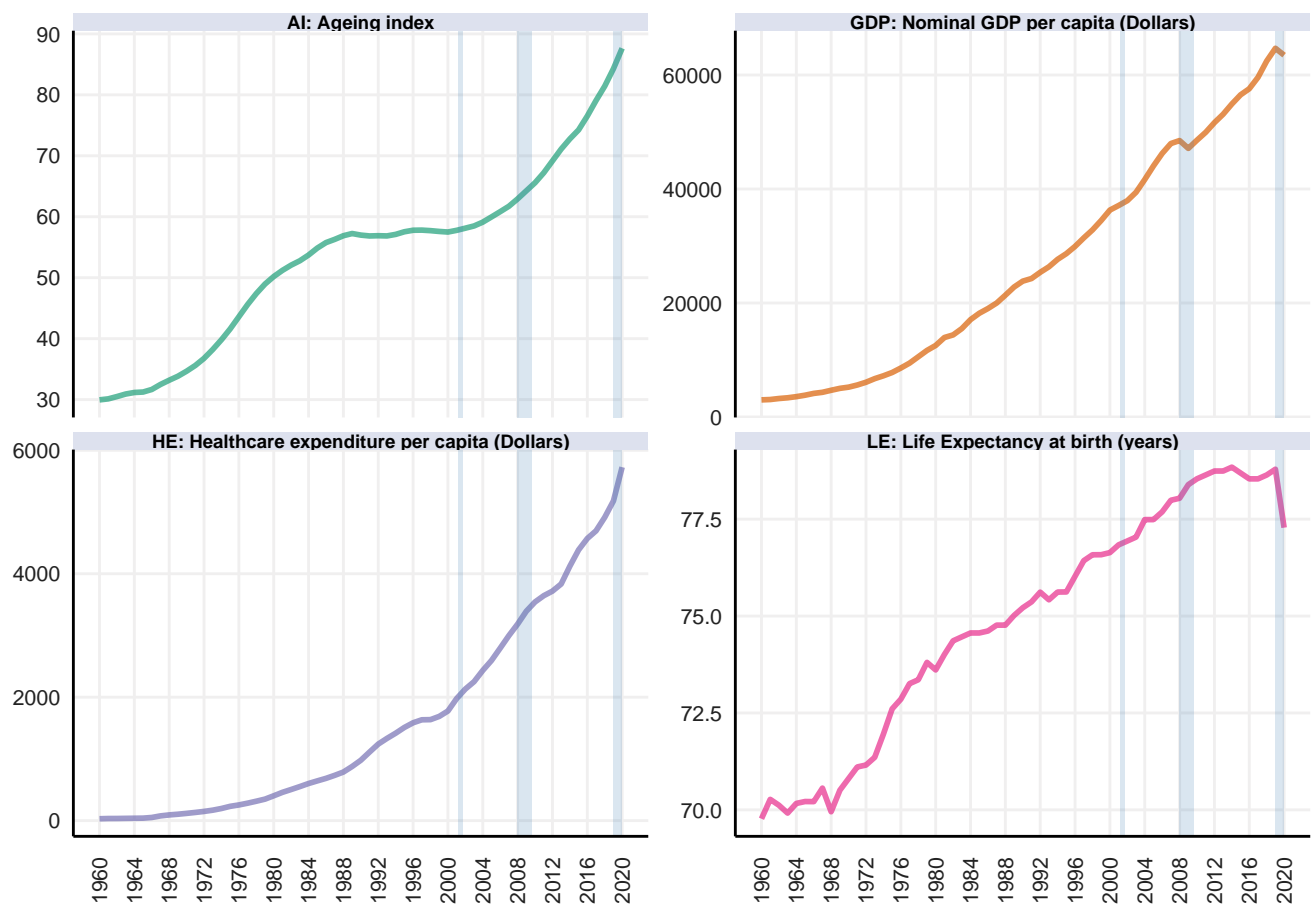
Notes: The shaded areas reflect economic recessions as defined by the NBER and the COVID–19 pandemic from 2019.
Source: Federal Reserve Bank of St. Louis' FRED database; Bureau of Economic Analysis.

Figure 4.1: Time series data sampled.

5. Stationarity

To test the four variables for stationarity, I conduct the general-to-specific Augmented Dickey-Fuller (ADF) test approach (Dickey & Fuller, 1979) . The approach entails testing hypothesis on three specifications of ADF test that have the following regressions with the corresponding test statistic and hypothesis:

i.   Variable with trend and drift:

$$\Delta y_t = b_0 + \alpha y_{t-1} + b_2 t + \sum_{j=2}^{p} \beta_j \Delta y_{t-j+1} + \varepsilon_t \tag{5.1}$$

$$H_0 : \alpha = 0 \ \& \ b_0 = 0 \ \& \ b_2 = 0 \ (test \ statistic = \phi_2)$$
$$H_0 : \alpha = 0 \ \& \ b_0 = 0 \ (test \ statistic = \phi_3)$$
$$H_0 : \alpha = 0 \ (test \ statistic = \tau_3)$$

ii.   Variable with drift and no trend:

$$\Delta y_t = b_0 + \alpha y_{t-1} + \sum_{j=2}^{p} \beta_j \Delta y_{t-j+1} + \varepsilon_t \tag{5.2}$$

$$H_0 : \alpha = 0 \ \& \ b_0 = 0 \ (test \ statistic = \phi_1)$$
$$H_0 : \alpha = 0 \ (test \ statistic = \tau_2)$$

iii.   Variable with no drift or trend:

$$\Delta y_t = \alpha y_{t-1} + \sum_{j=2}^{p} \beta_j \Delta y_{t-j+1} + \varepsilon_t \tag{5.3}$$

$$H_0 : \alpha = 0 \ (test \ statistic = \tau_1)$$

One or more test statistics are attached to each regression above, testing restrictions on the parameters of interest. The corresponding test statistics differ depending on the different specifications of each equation. In short, $\tau = 0$ suggests the existence of a unit root, $\phi_1 = 0$ or $\phi_3 = 0$ indicates no intercept and a unit root, and $\phi_2 = 0$ no intercept nor trend and a unit root for each of the four variables. The test statistics and their corresponding critical values are provided in Table 5.1 for each variable.[4]

---

[4]The null hypothesis under each specification is rejected at the corresponding significance level if the absolute value of the test statistic is larger than the absolute value of the critical value. I.e. Reject the null i.f.f. $|test \ stat.| > |critical value|$.

Table 5.1: General-to-specific ADF tests for $p = 3$.

Test Statistic	CV: 1%	5%	10%	Test Statistic	CV: 1%	5%	10%
GDP (*growth per capita*)				**HE**			
Level variable with trend and drift							
$\tau_3 = -5.23$	-4.04	-3.45	-3.15	$\tau_3 = 2.04$	-4.04	-3.45	-3.15
$\phi_2 = 6.2$	6.5	4.88	4.16	$\phi_2 = 3.17$	6.5	4.88	4.16
$\phi_3 = 14.33$	8.73	6.49	5.74	$\phi_3 = 4.6$	8.73	6.49	5.74
Level variable with drift and no trend							
$\tau_2 = -3.41$	-3.51	-2.89	-2.58	$\tau_2 = 3.05$	-3.51	-2.89	-2.58
$\phi_1 = 8.7$	6.7	4.71	3.86	$\phi_1 = 4.8$	6.7	4.71	3.86
Level variable with no drift and no trend							
$\tau_1 = -2.195$	-2.60	-1.95	-1.61	$\tau_1 = 3.1$	-2.60	-1.95	-1.61
LE				**AI** (*logged*)			
Level variable with trend and drift							
$\tau_3 = 0.97$	-4.04	-3.45	-3.15	$\tau_3 = -6.77$	-4.04	-3.45	-3.15
$\phi_2 = 3.93$	6.5	4.88	4.16	$\phi_2 = 16.36$	6.5	4.88	4.16
$\phi_3 = 2.15$	8.73	6.49	5.74	$\phi_3 = 22.94$	8.73	6.49	5.74
Level variable with drift and no trend							
$\tau_2 = -1.65$	-3.51	-2.89	-2.58	$\tau_2 = -1.6$	-3.51	-2.89	-2.58
$\phi_1 = 5.07$	6.7	4.71	3.86	$\phi_1 = 2.2$	6.7	4.71	3.86
Level variable with no drift and no trend							
$\tau_1 = 2.59$	-2.60	-1.95	-1.61	$\tau_1 = 1.15$	-2.60	-1.95	-1.61

The number of lags for each variable set to 3 according to the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Upon initially running the ADF tests on nominal GDP per capita, the presence of a unit root for all three specifications (5.1), (5.2), and (5.3). However, nominal GDP growth per capita (by log differencing) is stationary when including a drift term. This can be seen in Table 5.1; The null hypothesis of no trend, no drift, and the presence of a unit root cannot be rejected at the 1% significance level ($|\phi_2| < |CV_{1\%}|$). However, the null hypothesis of no drift and the presence of a unit root can be rejected at the 1% significance level ($|\phi_3| > |CV_{1\%}|$). This is verified by the same reasoning when hypothesising using ϕ_1 and τ_2.

The results of the ADF tests is the same for both healthcare expenditure (HE) and life expectancy at birth (LE); The null hypothesis of no trend, no drift, and the presence of a unit root cannot be rejected at all significance levels. However, when considering the specification in equation (5.2), then the hypothesis of no drift and a unit root is rejected at the 5% significance level. Therefore, an AR(3) with drift for both HE and LE are stationary. Under the same reasoning, the (logged) ageing index (AI) is stationary when including drift and and trend.

6. Results: Estimation, Impulse Responses, and Forecast Analysis

The estimated BVAR's number of lags are set to 3 according to the AIC and BIC and a constant term is included in the model following the stationarity tests conducted in previous section 5. I choose the four covariance matrix hyperparameters of the Minnesota prior (shown in equations 3.8 and 3.9) in the following fashion: V_1, reflecting the importance of prior beliefs, is set to a comparatively minor value as the prior information affect the estimation regarding the information contributed by the data; V_2 (lags across variables) and V_3 (exogenous variables) is set to larger than zero as, for this analysis, the estimates of the cross variable lags and the exogenous variables is central. The 'lag decay' hyperparameter V_4 is set to two similar to Koop & Korobilis (2010), thereby adopting a harmonic decay structure.

The posterior distributions of the coefficients for the Minnesota and conditional normal inverse-Wishart priors are shown in Figures A.1 and A.2 in Appendix A, respectively. This section proceeds by analysing and comparing the impulse response functions and forecasts under both priors.

6.1. Impulse Response Functions

The percentage-point response to a one standard deviation shock for each variable on each other for the Minnesota and conditional inverse-Wishart prior is depicted in Figures 6.1 and 6.2, respectively. In the case of responses from GDP (growth) and HE, the responses are given as the percentage point times 100.

In the case of the Minnesota prior (Figure 6.1), a HE shock decreases the GDP growth rate by approximately 1.3% over 3 years, increasing back after 5 years and then marginally increasing it past pre-shock levels. Importantly, a shock from AI reduces GDP growth by approximately 0.8% over 2 years, returning to its pre-shock levels a year later. The shock from HE on AI shows that the additional expenditure on healthcare only starts to increase AI after 4 years. In addition, a shock on HE has a persistent and negative effect on LE, showing that healthcare expenditure is implemented in a more reactive than proactive way. For example, during the COVID-19 pandemic, the increase in HE was in response to LE persistently declining. The shock from LE to HE, in turn, shows that as the LE starts to increase, there is a gradual and persistent reduction in HE.

Another interesting observation is that a shock from AI has a negative and persistent long-run effect on LE. This shows that the increase in the ageing index (AI) is more dependent on higher mortality for individuals aged 14 or less than lower mortality for ages 65 and over, causing the persistent reduction in LE. The corresponding shock on LE positively impacts GDP growth; however, the wide error bands show that this observation is not robust.

Compared to the impulse responses when the conditional inverse-Wishart prior was adopted (Figure

6.2), the observations are generally the same with a few differences. Specifically, a shock from HE has an insignificant effect on AI. However, the negative impact of HE on LE is amplified compared to the responses under the Minnesota prior, further exacerbating the reactiveness, as opposed to the proactiveness, of HE to LE.
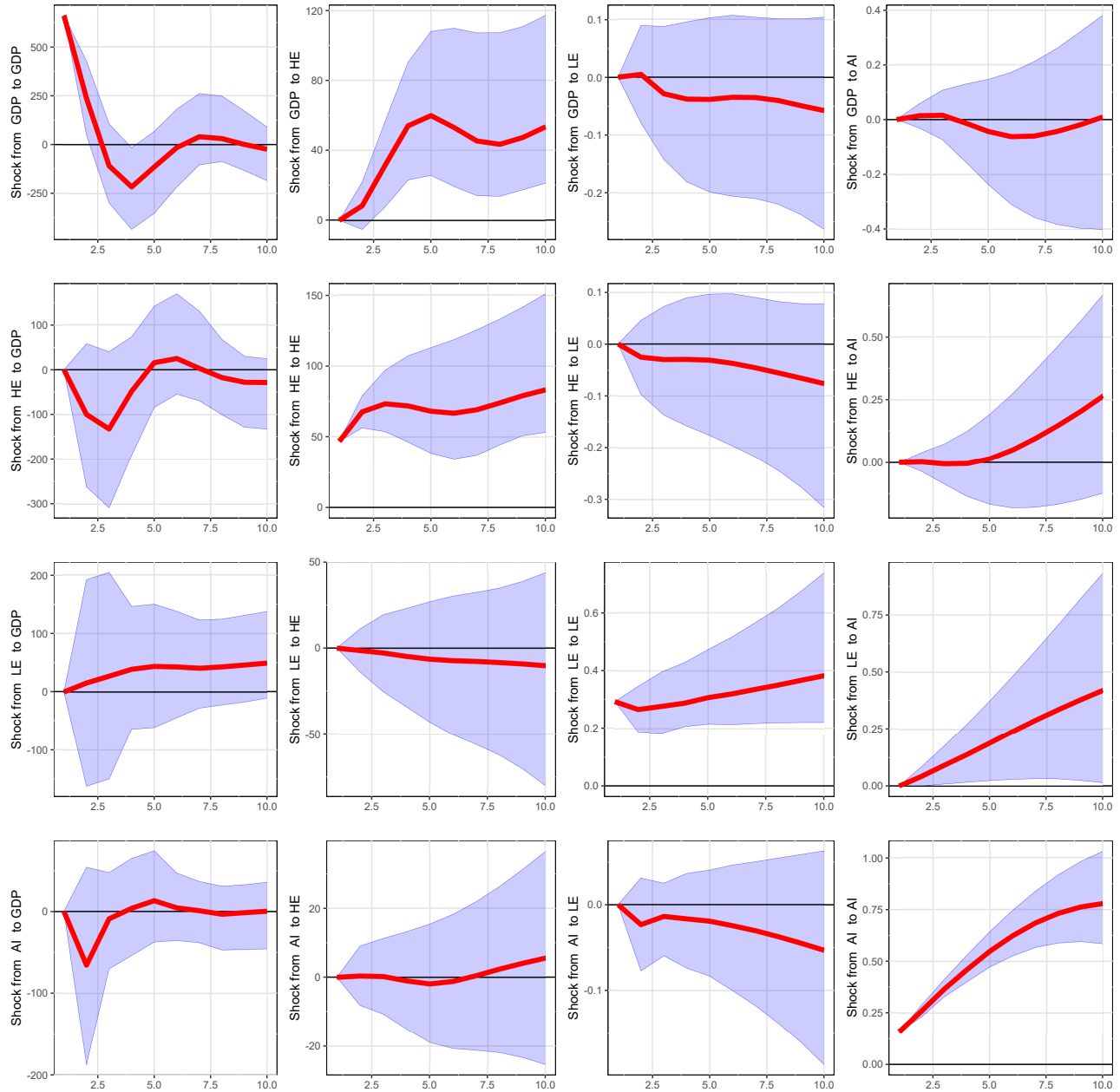

Figure 6.1: Minnesota prior: Impulse responses. Each plot shows the percentage-point response of the indicated variable from its steady state to a one-standard-deviation indicated shock when the Minnesota prior was adopted. Periods are in years. In the case of responses from GDP (growth) and HE, the responses are given as the percentage point times 100.
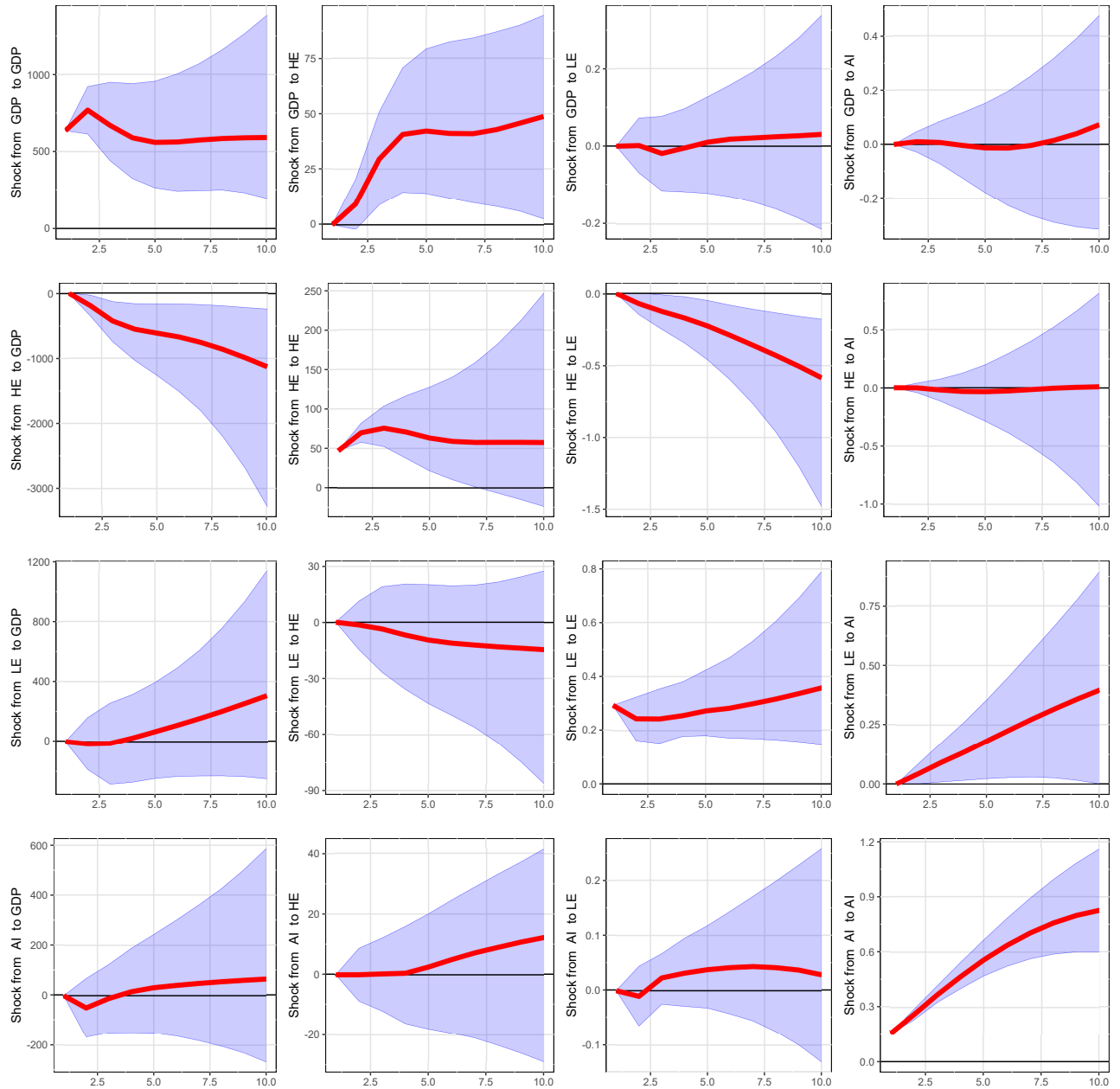
Figure 6.2: Conditional normal inverse-Wishart prior: Impulse responses. Each plot shows the percentage-point response of the indicated variable from its steady state to a one-standard-deviation indicated shock when the Conditional normal inverse-Wishart prior was adopted. Periods are in years. In the case of responses from GDP (growth) and HE, the responses are given as the percentage point times 100.

6.2. Forecasts

The 20-year forecasts for the four variables under the adopted Minnesota and conditional inverse-Wishart priors are shown in Figures 6.3 and 6.4, respectively. Due to the similarities across Figures,

I only analyse the case for the Minnesota prior.

The projections show a persistent increase in the ageing index (AI) and a continued decrease in life expectancy at birth (LE). This matches my recent experience in the US and again highlights the source of the reduced AI; the increased mortality of individuals aged 14 and younger is larger than the lowered mortality for individuals aged 65 and over, thereby reducing LE whilst increasing AI. As indicated in the impulse response analysis, healthcare expenditure persistently increases due to the lowered life expectancy.
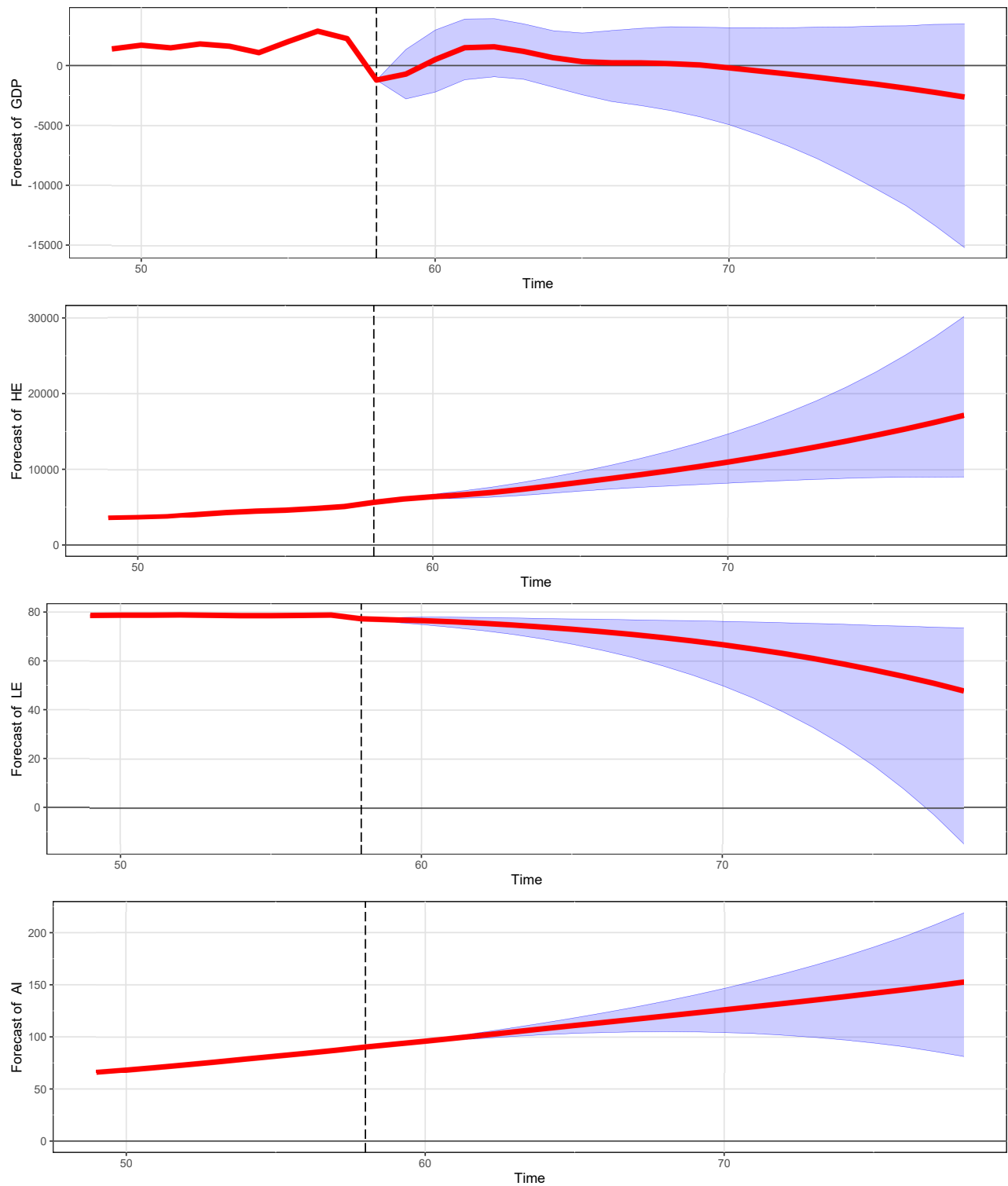
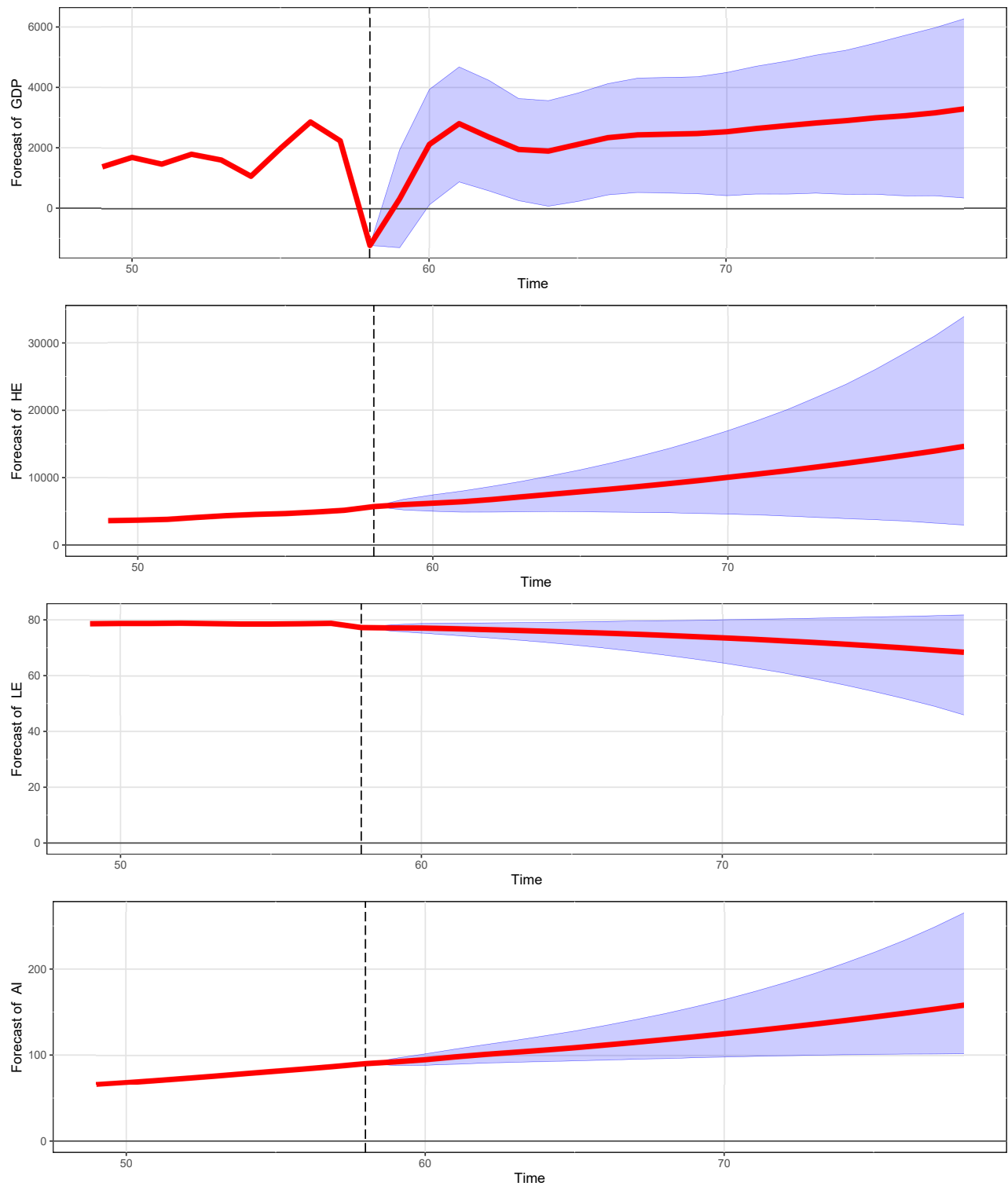Figure 6.3: Minnesota prior: Estimated forecasts. Periods are in years.

Figure 6.4: Conditional normal inverse-Wishart prior: Estimated forecasts. Periods are in years.

References

10 Amiri, A. & Linden, M. 2016. Income and total expenditure on health in OECD countries: Evidence from panel data and hsiao's version of granger non-causality tests. *Economics and Business Letters.* 5(1):1–9.

Amiri, A. & Ventelou, B. 2012. Granger causality between total expenditure on health and GDP in OECD: Evidence from the toda–yamamoto approach. *Economics letters.* 116(3):541–544.

Baltagi, B.H. & Moscone, F. 2010. Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic modelling.* 27(4):804–811.

Bloom, D.E., Canning, D. & Fink, G. 2010. Implications of population ageing for economic growth. *Oxford review of economic policy.* 26(4):583–612.

Braendle, T. & Colombier, C. 2016. What drives public health care expenditure growth? Evidence from swiss cantons, 1970–2012. *Health Policy.* 120(9):1051–1060.

Canova, F. 2007. *Methods for applied macroeconomic research.* Vol. 13. Princeton university press.

Chaabouni, S., Zghidi, N. & Mbarek, M.B. 2016. On the causal dynamics between CO2 emissions, health expenditures and economic growth. *Sustainable cities and society.* 22:184–191.

Chan, J.C. 2017. Notes on bayesian macroeconometrics. *Manuscript available at http://joshuachan.org.*

Cheng, X., Yang, Y., Schwebel, D.C., Liu, Z., Li, L., Cheng, P., Ning, P. & Hu, G. 2020. Population ageing and mortality during 1990–2017: A global decomposition analysis. *PLoS medicine.* 17(6):e1003138.

Dickey, D.A. & Fuller, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association.* 74(366a):427–431.

Doan, T., Litterman, R. & Sims, C. 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric reviews.* 3(1):1–100.

Fogel, R.W. 2004. Health, nutrition, and economic growth. *Economic development and cultural change.* 52(3):643–658.

Halıcı-Tülüce, N.S., Doğan, İ. & Dumrul, C. 2016. Is income relevant for health expenditure and economic growth nexus? *International journal of health economics and management.* 16:23–49.

Jaba, E., Balan, C.B. & Robu, I.-B. 2014. The relationship between life expectancy at birth and health expenditures estimated by a cross-country and time-series analysis. *Procedia Economics and Finance.* 15:108–114.

Koop, G. & Korobilis, D. 2010. Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends® in Econometrics.* 3(4):267–358.

Lee, R., Mason, A. & Miller, T. 2000. Life cycle saving and the demographic transition: The case of taiwan. *Population and Development Review.* 26:194–219.

Linden, M. & Ray, D. 2017. Life expectancy effects of public and private health expenditures in OECD countries 1970–2012: Panel time series approach. *Economic Analysis and Policy.* 56:101–113.

Liotta, G., Canhao, H., Cenko, F., Cutini, R., Vellone, E., Illario, M., Kardas, P., Poscia, A., et al. 2018. Active ageing in europe: Adding healthy life to years. *Frontiers in medicine.* 5:123.

Litterman, R.B. 1986. Forecasting with bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics.* 4(1):25–38.

Lopreite, M. & Mauro, M. 2017. The effects of population ageing on health care expenditure: A bayesian VAR analysis using data from italy. *Health policy.* 121(6):663–674.

Lopreite, M. & Zhu, Z. 2020. The effects of ageing population on health expenditure and economic growth in china: A bayesian-VAR approach. *Social science & medicine.* 265:113513.

Murillo, C., Piatecki, C. & Saez, M. 1993. Health care expenditure and income in europe. *Health Economics.* 2(2):127–138.

Murthy, V.N. & Okunade, A.A. 2016. Determinants of US health expenditure: Evidence from autoregressive distributed lag (ARDL) approach to cointegration. *Economic Modelling.* 59:67–73.

Wang, K.-M. 2011. Health care expenditure and economic growth: Quantile panel-type analysis. *Economic modelling.* 28(4):1536–1549.

Wiener, J.M. & Tilly, J. 2002. Population ageing in the united states of america: Implications for public programmes. *International journal of epidemiology.* 31(4):776–781.

Williams, G., Cylus, J., Roubal, T., Ong, P., Barber, S., Organization, W.H., et al. 2019. Sustainable health financing with an ageing population: Will population ageing lead to uncontrolled health expenditure growth?
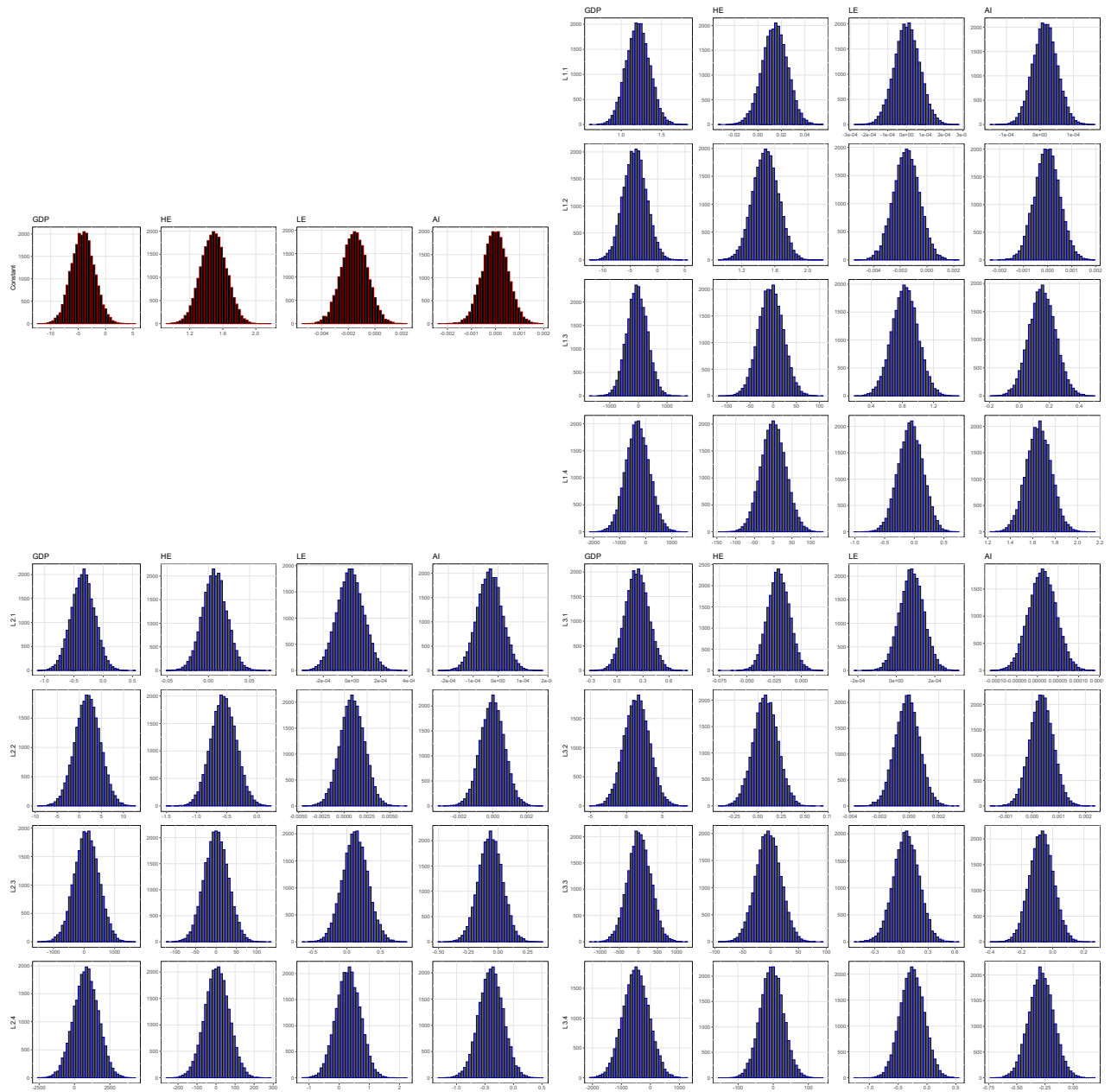
Appendix A.

Figure A.1: Minnesota prior: Estimated BVAR(3) coefficients' posterior distributions using Gibbs sampling.
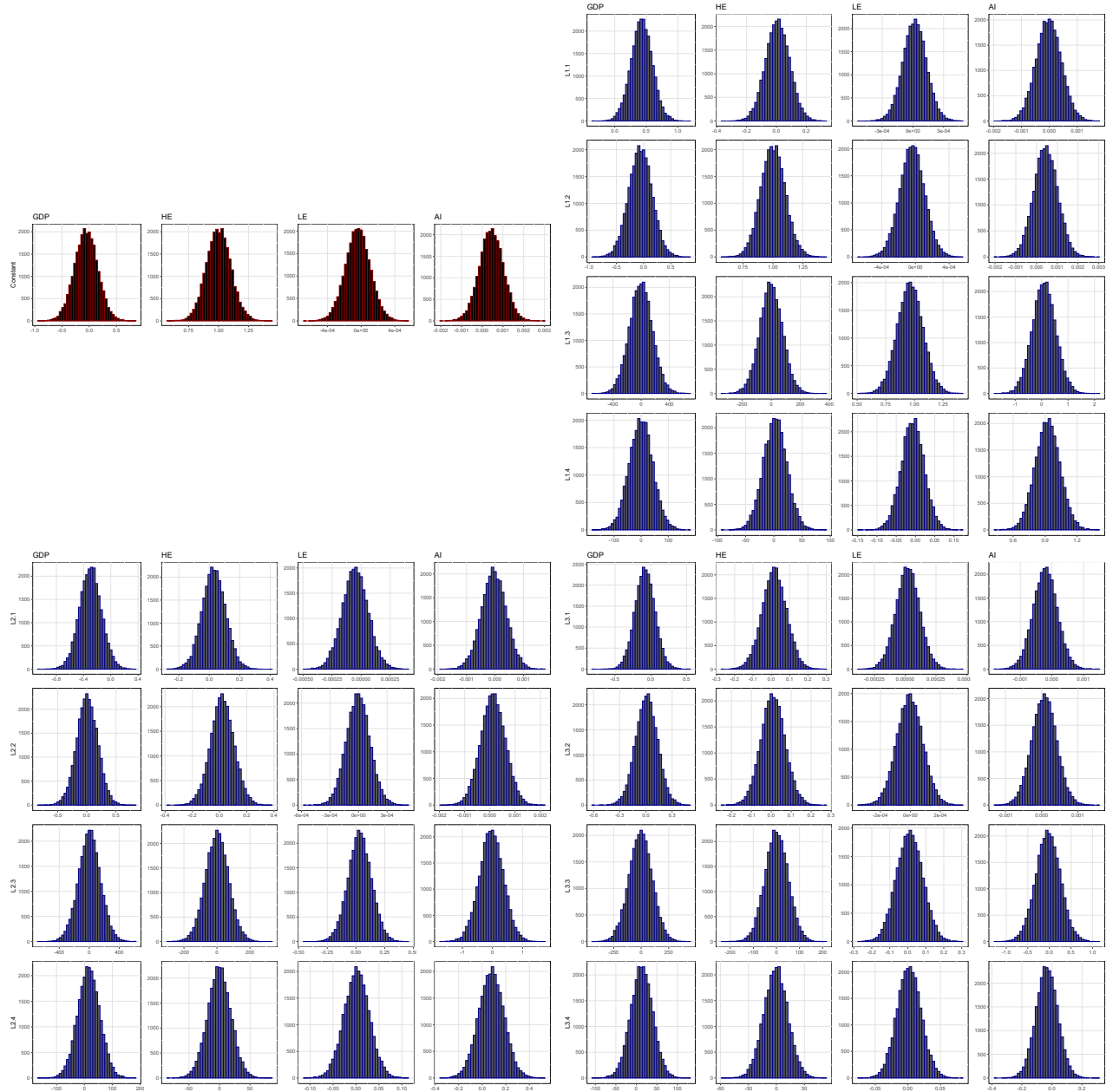
Figure A.2: Conditional normal Inverse-Wishart prior: Estimated BVAR(3) coefficients' posterior distributions using Gibbs sampling.