Forecasting State-Level Food Insecurity Rates in the United States

Akash Issar
The Agricultural and Food Policy Center
Department of Agricultural Economics
Texas A&M University
akash.issar@ag.tamu.edu

Pourya Valizadeh
The Agricultural and Food Policy Center
Department of Agricultural Economics
Texas A&M University
pourya.valizadeh@ag.tamu.edu

Henry Bryant
The Agricultural and Food Policy Center
Department of Agricultural Economics
Texas A&M University
henry.bryant@ag.tamu.edu

Bart Fischer
The Agricultural and Food Policy Center
Department of Agricultural Economics
Texas A&M University
bart.fischer@ag.tamu.edu

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25,2023

Forecasting State-Level Food Insecurity Rates in the United States

Akash Issar* Pourya Valizadeh† Henry Bryant ‡
Bart Fischer §

(Updated) July 19, 2023

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

Food insecurity affects the health and well-being of the households in the United States. Less is known about the future trends in the prevalence of household food insecurity rates at both the national and state level. The latter is particularly important given that household food insecurity varies considerably across states. We build a framework that produces national and state level annual projections of household food insecurity rates by exploiting three broad groups of state-level variables (economic, policy, and demographic variables). We estimate a static and dynamic two-way fixed effects (TWFE) models that conducts a specification search over a large set of forecasting models. Coefficient estimates are obtained for the in-sample period (1996-2016) and projected on the post-sample observations (2017-2021) to obtain the point forecast estimates. We evaluate the best model using the five-year average post-sample root mean squared forecast error. Our preliminary results indicate that the household food insecurity rates forecasts is best estimated via a dynamic TWFE log-linear model with determinants being the previous year's household food insecurity rates and unemployment rate, as well as contemporaneous house price index, poverty rate and SNAP policy sub-indices.

*Department of Agricultural Economics, Texas A&M University. akash.issar@ag.tamu.edu
†Department of Agricultural Economics, Texas A&M University. pourya.valizadeh@ag.tamu.edu
‡Department of Agricultural Economics, Texas A&M University. henry.bryant@ag.tamu.edu
§Department of Agricultural Economics, Texas A&M University. bart.fischer@ag.tamu.edu

1 Introduction

Food insecurity - a lack of continuous and assured access to sufficient food by all people at all times to lead an active and healthy life - affects the health and well-being of many economically vulnerable households in the United States (US). Food insecurity saw a gradual decline from its peak in 2011 until 2019 (prior to the onset of the COVID-19 pandemic). With the onset of the pandemic, researchers had predicted a major rise in food insecurity rates in the absence of measures to alleviate the household economic problems brought about by the pandemic (Gundersen et al., 2021). With the assistance of multiple players in combatting food insecurity (stimulus checks, charitable food assistance programs, increase in Supplemental Nutrition Assistance Program (SNAP) benefits and a resilient agricultural supply chain), food insecurity saw no spike in 2020-2021 and remained around the 2019 level (Gundersen, 2023).

Although national trends in household food insecurity rates show a national average of 10.2% (as of 2021), there is considerable variation in household food insecurity rates across the states and less is known about state-level future trends in household food insecurity rates. The objective of this study is to build an econometric model for post-sample forecasting (as defined in Granger and Huang (1997)) of state's household food insecurity rates in the US and evaluate its forecasting performance on national and state-level point estimates. Furthermore, we use state-level panel data rather than national time-series data to model food insecurity rates. Panel data modelling allows us to capture observed and unobserved heterogeneity as well as allows to conduct state-level counterfactual scenarios.

Existing works on modelling food insecurity and its determinants focus, primarily, on inference (Smith and Gregory, 2023; Restrepo et al., 2021; Bartfeld and Men, 2017; Bartfeld and Dunifon, 2006; McKernan et al., 2021; Han, 2016) and those works on (post-sample) forecasts of food insecurity rates employ a national time series approach to global food security issue (Lentz et al., 2019; Wang et al., 2022). Thus, our current work advances these prior works to provide a panel-data forecast modelling of food insecurity for the Unites States at the national and state level (that is currently unavailable in the literature).

The only available study providing limited projections of food insecurity rates in the US was undertaken by Gundersen et al. (2021). Gundersen et al. (2021) provide contemporaneous spatial out-of-sample forecasts of food insecurity using Feeding America's Map the Meal Gap (MMG) model developed in Gundersen et al. (2014). They model annual state-level food insecurity rates and provide contemporaneous annual projections at the sub-state (county) level. Gundersen et al. (2021) conduct out-of-sample forecasts by providing (current) county food insecurity rates using state-level regression estimation. MMG has been widely used by food banks to direct resources as well as provide state and county governments a plan of action to target communities struggling with food insecurity.

Gundersen et al. (2021) used the MMG model to guide policymakers on the likely projections of food insecurity rates as a result of sky-rocketing unemployment rates with the onset of the COVID-19 pandemic. They projected an increase of 17 million food insecure households in 2020. However, food insecurity rates reported by the United States Department of Agriculture (USDA) for the year 2020 and 2021 show that food insecurity rates remained around their 2019 level (Coleman-Jensen et al., 2022). This can be attributed to the central role played by multiple agents: stimulus checks, charitable food assistance programs, increase in SNAP benefits and a resilient agricultural supply chain (Gundersen, 2023). Thus, a good forecasting model of food insecurity can help policymakers with national and state-level trends in prevalence of food insecurity and to direct resources to areas likely to experience high food insecurity.

The MMG model, on the other hand, was built to provide current county-level projections and not, necessarily, post-sample forecasts. Moreover, MMG and other available food insecurity models' performance for forecasting into the future is limited. Thus, we undertake the task of developing an econometric forecasting model of household food insecurity at the state and national level and evaluate its performance using post-sample observations (rather than an in-sample fit, usually done for causal inference).

We develop a forecast model of national and state-level annual household food insecurity rates using the relationship between food insecurity and its determinants well established in the prior literature (Gundersen, 2022; Gundersen et al., 2014; Bartfeld et al., 2006; Bartfeld and Dunifon, 2006; Bartfeld and Men, 2017). We classify our inputs into three broad categories: economic, policy, and demographic. We collect data on states' annual household food insecurity rates from 1996 to 2021 from the Current Population Survey's annual Food Security Module (CPS-FSM) (Smith and Gregory, 2023).[1] The survey provides annual household responses to the Food Security Supplement (FSS), and we aggregate household responses to the state-level using

[1]See, Smith and Gregory (2023) for an overview on food insecurity in the US and its modelling and measurement.

sampling weights. We consider households with and without children in aggregating the state's annual household food insecurity rates. Furthermore, we conduct a specification search across various models that differ in the combination of the set of predictors.

To evaluate forecast modelling performance, we compare the post-sample forecast point estimates to our observed food insecurity rates in the post-sample. To do this, we divide our available historical data (1996-2021) into two sets of observation: in-sample and post-sample. The in-sample observations take the historical data from 1996 to 2016 to estimate the parameters of the predictors that are then projected onto the post-sample observations from 2017 to 2021. We calculate the five-year (2017-2021) average post-sample root mean squared forecast error (RMSFE) for each model specification as a summary measure of forecast performance. The model that generates the lowest post-sample RMSFE is selected as the most accurate model (among the universe of forecasting models considered herein) estimator for forecasting national and state-level annual household food insecurity rates.

Our model specification search estimates the in-sample coefficients using the static & dynamic two-way fixed effects (TWFE) panel data estimator. TWFE controls for time-invariant and time-variant unobserved heterogeneity with state and year fixed effects, respectively. In our forecast modelling framework, we conduct a specification search over different models that vary with the combination of predictors, the adjustment of time fixed effects into the future, log versus level outcomes and static versus dynamic TWFE.

The remainder of the paper proceeds as follows. Section 2 provides the conceptual framework in modelling food insecurity and discusses data sources. Section 3 details the methods used to develop a food insecurity model and describes the sources of variation underlying the set of forecast specifications. We then present our results from the in-sample estimated coefficients and the post-sample point forecast estimates for both national- and state-level annual household food insecurity rates in Section 4. Finally, Section 5 concludes.

2 Conceptual Framework

We model state's annual household food insecurity rates with demographics, economic characteristics and local policies as determinants that are consistent in prior literature in explaining the probability of a household being food insecure. An overview of these determinants are provided in Gundersen and Ziliak (2018) and have been applied in Gundersen et al. (2021, 2014) to construct county level food insecurity rates; as well as in Bartfeld and Men (2017); Bartfeld and Dunifon (2006) explore the potential factors in explaining the probability of being food insecure for households with children.

Table 1 lists the set of predictors we employ to model the relationship between states' household food insecurity rates and their determinants. Our choice of this subset of predictors from the list of predictors available stems from the availability of projected estimates of these inputs into the future. We keep these set of predictors limited to those whose future projections from reputable sources (e.g. Congressional Budget Office (CBO) or the S&P Global's IHS Markit) are available. We do this to extend our forecasts into the next decade where our model would require future values of our predictors.[2]

Table 1 categorizes the set of inputs into three broad categories: economic, local policies, and demographics. State economic conditions include seasonally adjusted average annual state unemployment rate (in percent) measuring labor force activity; House Price Index (HPI) measures the accessibility & affordability of home mortgage loans; annual estimates of state-level poverty rates (in percent); average annual household personal income per-capita (in real 2021 USD) and State GDP per capita (in real 2021 USD).

State-level policies include real effective state minimum wage (in 2021 USD), SNAP participation per capita, broad and narrow SNAP policy index measuring the ease and accessibility of being on SNAP, and real annual average SNAP benefits per-person (in 2021 USD). We include SNAP variables as SNAP is the primary and largest food assistance program food assistance program in the US to address food insecurity.[3]

State demographics capture household level characteristics. They include age composition of individuals in a state. Age composition is measured as the share of

[2]Information on the Congressional Budget Office is available at: https://www.cbo.gov/data/budget-economic-data. And the information on the S&P Global's IHS Markit is available at: https://www.spglobal.com/en/.

[3]The other food assistance programs are: special supplemental nutrition program for Women, Infant and Children (WIC), National School Lunch Program (NSLP) and School Breakfast Program (SBP). We do not include these program statistics into our set of predictors for two reasons: 1) SNAP is the largest food assistance program that makes up the majority of Food Government food assistance program expenditure; 2) The future state-level projections of these programs are not available. The central goal of SNAP, the largest food assistance program in the US, is the alleviation of food insecurity and has been well documented that SNAP reduces food insecurity (McKernan et al., 2021; Han, 2016; Gundersen et al., 2017; Gundersen, 2022; Borjas, 2004; Corman et al., 2022).

state population that makes up the following age demographics: 0 to 14 years old; 15 to 29 years old; 30 to 44 years old; 45 to 59 years old; 60 to 74 years old; and 75 years old and higher. The racial/ethnic composition measures the share of the state population with Hispanic; white and non-white individuals. Education measures the share of the state population with a high school diploma or less. Finally, home ownership measures the share of the state population that owns a home.[4]

2.1 Data

USDA has measured household food insecurity through an 18-item questionnaire in the Food Security Supplement of the Current Population Survey (CPS-FSS) every December since 1995.[5] CPS-FSS has been shown to be a reliable and a valid instrument for measuring household food insecurity (Marques et al., 2015). CPS-FSS asks 10 out of the 18 questions to adults in the households measuring their experiences in food security. Households with children are asked the additional eight questions to measure the food security experience among children. Households are then classified as food insecure if they respond affirmatively to three or more food insecurity experiences in the past year.[6]

The questions in the CPS-FSS ask households about their experience with food acquisition and purchase owing to financial constraints in acquiring adequate food to feed the family on a regular basis.[7] Figure 1 displays the annual trend in the United States for overall food insecurity rates as well as very low food security rates from 1996-2021.[8] Very low food secure households experience reduced food intake and disruption in their eating pattern or skip meals. As of 2021, 10.2% of households

[4]Disability status of household's members is an important determinant to food insecurity rates (Henly et al., 2022). However, we fail to include a measure capturing share of households with disabled members as reporting of disability status is available in the CPS from 2009 onwards. Thus, for our objective of forecast modelling to learn from a long history of food insecurity starting from 1996 we have to exclude a measure of disability status.

[5]See, Smith and Gregory (2023) for an overview of food insecurity in the US and its measurement and modelling.

[6]USDA categorizes the responses into four categories: high food secure (no reported affirmative responses of food-access problems or limitations.), marginal food secure (one or two reported affirmative responses), low food secure (three to five affirmative responses), and very low food secure (six or more (out of 10 questions) or 8 or more (out of 18 questions) affirmative food insecure responses) (Coleman-Jensen et al., 2022).

[7]See Coleman-Jensen et al. (2022) for the complete 18 item questionnaire.

[8]It is important to note that the screener process of the Food Security Module (FSM) was different prior to 1998 and since 1998 USDA provides a consistent food security status. However, for our forecast estimation rather than dropping data for 1996 and 1997 we use the food insecurity rates for those years to leverage information starting from 1996.

(13.2 million) experienced food insecurity. Both food insecurity and very low food security rates rose greatly during the Great Recession (there was a 4- percentage-point increase in food insecurity rates in 2008 from the 2005-2007 rates) and peaked in 2011 and has been gradually declining (reaching the pre-Great Recession levels in 2019). And, 3.7% of households experienced very low food security in 2021. The very low food security rates experienced a 2-percentage-point jump in 2008 relative to their rates in 2007.

Food insecurity rates vary considerably across states. Not only do household level characteristics (income, age, race, education) impact a household's food insecurity experience, the health of the state economy (e.g., cost-of-living, rent) also impacts food insecurity rates. Figure 2 provides a snapshot of the variation in household food insecurity rates for the years 2001, 2008, 2016, and 2021, across 50 states and the District of Columbia. As of 2021, New Hampshire had the lowest food insecurity rate of around 6% and Arkansas had the highest food insecurity rate of around 18%, compared to a national average of 10.2%.[9]

Historical state-level food insecurity rates are constructed from the annual household survey using the CPS-FSS from 1996 to 2021. We aggregate the household level responses to the state level using the survey sampling weights. We gather data on state-level macroeconomic variables (e.g., unemployment rate, poverty rate, GDP per capita, etc.) from various sources such as the Bureau of Labor Statistics, the Bureau of Economic Analysis, and the US Census. State-level policies capturing SNAP measures (participation rate, policy parameters, and average benefits per-person) are taken from USDA's Food & Nutrition Services (FNS) SNAP Data Tables and the SNAP Policy Database.[10] Lastly, information on a state's demographic composition is obtained from the US Census and the CPS' Annual Social and Economic Supplement (ASEC) (Flood et al., 2022).

2.2 Summary Statistics

Table 2 provides summary statistics of the household food insecurity rates and its determinants. The summary is provided for 50 states and Washington D.C. from 1996 to 2021 (a total of 1,326 observations). Table 2 also provides the sources of data. We see that, for our observations, household food insecurity has averaged

[9]In the Appendix, figure A2 and A3 display household food insecurity rates as a three-year average starting from 1996 to 2021.

[10]SNAP Policy Database, currently available, includes SNAP policy parameters only upto 2016. So, we make a strong assumption of using the 2016 values post-2016.

around 12% with some states experiencing food insecurity as low as 4% to high as 23%.

SNAP policy parameters measure the state-level SNAP policies that States have the autonomy to implement beyond the Federal regulations that govern the SNAP program. We obtain the monthly State policy options from the USDA SNAP Policy Database from 1996 to 2016.[11] The SNAP policy parameters available in the FNS are summarized into two overall indices: broad and narrow SNAP policy index. Broad SNAP policy index is a summary measure of 28 SNAP policy parameters where we take the average of the policies active in any given year. The detailed list of the 28 SNAP policy parameters are shown in column 1 of the appendix table A1. The narrow SNAP policy index is constructed from a smaller set of 9 policies that are essential in capturing SNAP accessibility and used in prior works (See, for e.g., Stacy et al. (2018)). Column 2 of appendix table A1 lists the nine polices we use to capture the narrow index.

Furthermore, these SNAP policies can be grouped into three broad categories: eligibility, transaction costs and stigma (Stacy et al., 2018). *Eligibility* measures SNAP policies that expand or restrict SNAP participation. For example, States that implement Broad Based Categorical Eligibility (BBCE) and its components help in increasing SNAP participation as BBCE couples other food stamp participation with other welfare programs. *Transaction costs* are SNAP policies that reduce the administrative burden and participants' tangible cost in participating in SNAP. For example, States that implement longer re-certification periods help in reducing the administrative costs of SNAP case-workers repeatedly reviewing households' information and reduces SNAP households time and money in getting re-certified. *Stigma*, the final group, includes the policies that reduce the stigma and apprehension among potential SNAP participants in participating in SNAP. For example, States that waive face-to-face interviews help SNAP households in overcoming the stigma of being on SNAP.

The SNAP policy indices (overall or the three sub-indices) are constructed as the simple average of the number of active expansionary policies aimed at expanding SNAP participation. We measure each policy as a positive binary outcome that takes on the value of one if the policy is thought to increase SNAP participation and zero otherwise. We count the number of positive SNAP policies in a given month in a

[11]SNAP Policy Database and documentation can be found at: `https://www.ers.usda.gov/data-products/snap-policy-data-sets/snap-policy-database-documentation/`

particular State and take the simple average of the number of positive policies. Thus, our SNAP policy indices range between 0 and 1 where values closer to 1 indicate a more expansionary SNAP program that increases accessibility to SNAP program.

3 Methods

Equation 1 describes the relationship between household food insecurity rates at the state level and its determinants:

$$FI_{st} = \alpha + \boldsymbol{\beta_1}\textbf{Economic}_{st} + \boldsymbol{\beta_2}\textbf{Policies}_{st} + \boldsymbol{\beta_3}\textbf{Demographics}_{st} + \mu_s + \delta_t + \varepsilon_{st} \tag{1}$$

where s indicates a state and t indicates time (here, year). **Economic**, **Policies** and **Demographics** are a vector of state economic conditions, state-level policies and state demographics , respectively, as listed in table 1. μ_s and δ_t capture state and time unobserved heterogeneity through state and year fixed effects, respectively.[12]

Our approach draws on state-level panel data rather than national time-series data to model food insecurity rates. Panel data modelling serves two primary advantages. First, we can conduct *counterfactual* policy analysis of important state-level policy changes. For example, SNAP plays a pivotal role when the state of the economy is in poor health by providing supplemental income to households to purchase food and policymakers frequently re-evaluate and re-adjust the SNAP policy parameters depending on the state of the economy (Anderson et al., 2022; Gundersen, 2022; Schanzenbach, 2023). Second, econometrically, a panel data method is superior to a time-series model because it incorporates important state-level observed and unobserved heterogeneity in the data which could improve (post-sample) forecast performance. The latter is particularly important considering that US states are heterogeneous entities and differ from each other in several observed and unobserved dimensions, and subsequently in their responses to local and macroeconomic shocks as well as national policy changes.

The regression parameters (β's) of equation (1) are estimated using a fixed estimation forecast scheme (see, West (2006) for detailed overview on different forecast

[12]State fixed effects (μ_s) enter as 50 state dummy variables (one state dummy variable dropped as reference category) in our regression. And, time fixed effects (δ_t) enter either as year dummies or polynomials of linear time trend. We discuss more on modelling of time fixed effects below.

estimation schemes used for forecast evaluation).[13] In a fixed estimation forecast scheme, we divide our historical observations of 1,326 (51 states from 1996 to 2021) into two sets where the first set of observations of 1,071 (51 states from 1996 to 2016) are used as in-sample observations to estimate our regression parameters that are projected onto the second set of observations of 255 (51 states from 2017 to 2021) that are used as post-sample to evaluate forecast performance of the model (See figure 3).

We evaluate model performance by calculating the post-sample RMSFE and choose the best model that produces the lowest post-sample RMSFE. Our post-sample periods in the fixed estimation scheme are the 5-year period following 2016 (i.e., 2017 to 2021). We calculate our 5-year average post-sample RMSFE based on the post-sample period forecast estimates, as in equation 2:

$$\text{RMSFE}_m = \frac{1}{51}\sum_{s=1}^{51}\sqrt{\frac{1}{\texttt{H}}\sum_{t=2017}^{2021}(FI_{st} - \hat{FI}_{s,t})^2} \tag{2}$$

where, H is the number of periods in the post-sample window (in our application, this is 5 years (2017-2021)); FI_{st} is the observed household food insecurity rates for state s for post-sample time period t; $\hat{FI}_{st}$ is the point forecast estimate. RMSFE_m is the post-sample RMSFE generated for each model m included in our specification. We run numerous model specification search of static and dynamic TWFE that vary with the set of determinants and compare each model's post-sample point forecast performance through post-sample RMSFE statistics given in equation 2.

3.1 Time Fixed Effects Adjustment

Time fixed effects in panel/longitudinal studies capture macroeconomic shocks common across units. To model time fixed effects (δ_t in equation 1) in our static & dynamic TWFE panel data estimator we include the task of adjusting for time fixed effects alongwith our task of searching across specifications by adding to that specification search variation in how time fixed effects are estimated in-sample to provide post-sample future (unobserved) time fixed effects estimates. Traditionally, time fixed effects, such as year fixed effects, are incorporated using time (year) dummies. However, using time dummies raises the complication about post-sample future estimates

[13] In the appendix section A.5 we provide a quick overview of other forecast estimation schemes to obtain regression parameters: recursive, rolling and fixed scheme.

of time effects as time does not repeat in future. Thus, we use our in-sample estimates of time fixed effects and project it post-sample using the following ways:

1. Last time fixed effects: use the last period time fixed effects estimate and use the that estimate as our post-sample time fixed effects estimate.[14]

2. Predict time fixed effects by modelling functional form of time fixed effects: we predict post-sample time fixed effects estimate by estimating a linear, quadratic, cubic polynomial of time trend.

In addition to modelling time fixed effects using year dummies, time fixed effects can also be modelled using linear time trend and/or its polynomial functional form. Gösser and Moshgbar (2020) provide a deep exploration on when a linear time trend and its polynomial approximates time fixed effects when time dummies are hard to model or are inapplicable. Gösser and Moshgbar (2020) show that trend and its polynomial form (they call this "*smoothing time fixed effects*"), under some regularity conditions, "*yield consistent estimates by controlling for time fixed effects, also in cases time-dummy variables are inapplicable*" (Gösser and Moshgbar, 2020).

Thus, in our forecast modelling task we add, in addition to the time dummies approach above, to our econometric method adjustment of time fixed effects via a linear trend as well as quadratic linear time trend functional form.

3.2 Variation in Model Specification

Based on our discussion above, we can summarize our forecast modelling approach through the following ways:

1. Variations from permutations of explanatory variables:

    a) State economic conditions

    b) State local policies

    c) State demographic characteristics

2. Logarithmic versus level outcome of state food insecurity rates.[15]

[14]This approach of using last time period dummy into the future has been applied in Gundersen et al. (2021). They use the estimated coefficient of the 2018 year dummy to project county level household food insecurity rates for the pandemic year of 2020.

[15]For logarithmic outcomes while transforming from log to levels we use the Goldberger (1968) transformation.

3. Static versus dynamic TWFE estimator.

4. Time fixed effects adjustment in TWFE:[16]

   a) Linear time trend

   b) Polynomials of linear time trend (quadratic)

   c) Last in-sample period year fixed effects estimate projected into the future

   d) Estimates of future projections of year fixed effects through a linear, quadratic, or cubic polynomial time trend.

4 Results

Table 3 presents the estimated coefficients of our top five models with the lowest five-year average post-sample RMSFE for the period of 2017 to 2021. Each of the models in table 3 are estimated via a dynamic TWFE specification (with one-year lag of household food insecurity rates) and estimated with a quadratic polynomial of linear time trend. Each of the top five models show that one-year lag of household food insecurity rates, one-year lag of unemployment rate and contemporaneous house price index and poverty rate are strong predictors of state's annual household food insecurity rates. And, the models vary by additional predictors: broad or narrow snap policy sub-indices, snap participation per-capita. Thus, we see that within our top five models SNAP prorgam statistics (participation or policy indices) play an important role in prediction.[17]

Bottom rows of table 3 provide the five-year average of post-sample RMSFE. The lowest post-sample RMSFE is model 1 in table 3 with the broad SNAP policy sub-indices (eligibility, transaction costs, and stigma) as additional explanatory predictors. And, the broad SNAP eligibility sub-index statistically significant coefficient tells us that as more SNAP eligibility parameters are in place, it helps reduce household food insecurity.

Figure 4 overlays the estimated point-forecasts of our top twenty models for the post-sample period (2017-2021). All of the top twenty models are dynamic TWFE models estimated using with a quadratic polynomial of linear time trend. In figure 4 we also highlight the estimated point forecasts from our best model (model 1 in

[16]The adjustment of time fixed effects in forecast modelling using TWFE estimator on an out-of-sample data is briefly discussed and applied in Auffhammer and Steinhauser (2012).

[17]Additionally, we obtain the expected sign and relationship between SNAP program parameters and household food insecurity rates.

table 3) as shown through a red dashed line. The best model (as well as the other 19 models) does relatively well in tracking the post-sample observed food insecurity rates. Moreover, evident from table 3 and figure 4 the post-sample RMSFEs are in the neighborhood of each other.

Figure 5 plots the post-sample (2017-2021) RMSFEs for the 1 year up to 5 years ahead for the top 20 models. Figure 5 shows how the post-sample RMSFE (estimated from a fixed estimation scheme) across 20 models increases as we move into one period ahead in our forecast horizon.

With the point forecast estimates of our best model available, we are further interested in the uncertainty around these estimates. Thus, we measure the uncertainty in our forecast estimates coming from the error uncertainty using a multi-step ahead forecast interval procedure. We conduct 5,000 simulation and re-estimate the forecast estimates 5,000 times and plot 95%, 90%, 60% and 30% forecast interval. In figure 6, we plot the observed (actual) national annual household food insecurity rates (black solid line), point forecast estimates (red solid line) and the uncertainty surrounding our best model (gray shaded regions). The wide forecast interval at the start of the post-sample period is reflective of the historical volatilities observed in food insecurity rates (for example, the spike in food insecurity rates observed at the onset of the Great Recession (about 4 percentage point increase in 2008's food insecurity rates relative to 2007)).

4.1 State-Level Forecast Performance

As stated in the introduction section 1, the objective of our current research is to provide both national and state-level forecasts. In the previous section 4, we showed our best model specification and its performance on national annual household food insecurity rates. In this section, we show how the best forecasting model performs by providing annual state-wise forecast estimates. In figure 7 , we compare actual state-wise household food insecurity rates to predicted estimates for the post-sample period using our best model described in section 4 to summarize the performance of state-wise household food insecurity rates performance, we take three-year average of household food insecurity rates and categorize the food insecurity rates into one percentage points interval. The darker shaded regions in figure 7 indicate higher level food insecurity rates.

The left panels of figure 7 display the observed states' household food insecurity rates and the right panels are the corresponding states projected estimates. We observe in figure 7 that our best performing model is able to capture the variation in household food insecurity across states for both 2016-2018 and 2019-2021 three year average. However, for the three-year period of 2019-2021, our predicted estimates under-estimate, marginally, for states like Texas, New Mexico and Louisiana, for example.

5 Conclusion

Food insecurity, a condition in which households lack access to adequate and nutritious food because of limited financial and other resources, is a major public health issue in the United States. In 2021, 13.5 million US households (or, 33.8 million individuals), including 2.3 million households with children, were food insecure (Coleman-Jensen et al., 2022). Economic shocks and uncertainty, like the Great Recession and the COVID-19 pandemic, exacerbated household and state economic stressors, especially faced by economically vulnerable households, highlighting the need for accurately forecasting food insecurity rates and understanding changes in food insecurity rates, and how they relate to an array of programs in the US safety net.

We measure state household food insecurity rates as the share of households that report to be food insecure (Coleman-Jensen et al., 2022; Tiehen et al., 2020). We use sampling weights to aggregate household-level data to the state level. We model food insecurity as a function of three broad groups of variables (economic, policy, and demographic). We divide the observational periods (1996-2021) into in-sample (1996-2016) and out-of-sample periods (2017-2021). We employ static and dynamic two-way fixed effects (TWFE) panel data estimators that estimate the parameters for the in-sample observations and project the estimated coefficients for the out-of-sample observations to obtain the 5-year point forecasts for each state and year. Our forecast modelling includes a specification search over combination of determinants of household food insecurity established in prior literature (Gundersen, 2022; Bartfeld and Men, 2017).

Our preliminary results uses the fixed estimation scheme to estimate the parameter of both static and dynamic TWFE panel data estimator.[18] Our best performing forecasting model is chosen using the model specification that provides the best (lowest) 5-year average post-sample RMSFE. The lowest post-sample RMSFE we obtain is about 0.016 (a 1.6 percentage point difference between the observed and the

[18]In the next development of our forecast model, we supplement our results with the rolling window scheme and evaluate the difference between the two schemes arising from parameter instability.

predicted household food insecurity rates). The best forecast model comes from a dynamic TWFE model (with a quadratic linear time trend smoothing of time fixed effects). The predictors that generate the best forecast model are one-year lag of household food insecurity rates, one-year lag of unemployment rate and contemporaneous house price index, poverty rate and broad SNAP policy sub-indices. Our best model generates point-forecast estimates that performs relatively well in producing both the national and state-level household food insecurity rates.

As we improve upon on this current version of the forecast model, we aim to strengthen the forecast modelling framework by incorporating the rolling window scheme that would account for parameter instability. Finally, with the best model selected, we aim to construct counterfactual scenarios as well as estimate household food insecurity rates for the next decade. The future post-out-of-sample projections for national and state-level household food insecurity rates will use the estimated projections of the predictors sourced from well-established reputable sources like CBO and the S&P.

References

Anderson, Theresa, Elaine Waxman, and Craig Gundersen, "The impact of the Meals-to-You program on food insecurity," *Applied Economic Perspectives and Policy*, 2022, *44* (3), 1499–1512.

Auffhammer, Maximilian and Ralf Steinhauser, "Forecasting The Path of U.S. CO2 Emissions Using State-Level Information," *The Review of Economics and Statistics*, 02 2012, *94* (1), 172–185.

Bartfeld, Judi and Rachel Dunifon, "State-level predictors of food insecurity among households with children," *Journal of Policy Analysis and Management*, 2006, *25* (4), 921–942.

_ , _ , Mark Nord, and Steven Carlson, "What Factors Account for State-to-State Differences in Food Security?," Economic Information Bulletin 7086, United States Department of Agriculture, Economic Research Service November 2006.

Bartfeld, Judith and Fei Men, "Food Insecurity among Households with Children: The Role of the State Economic and Policy Context," *Social Service Review*, 2017, *91* (4), 691–732.

Borjas, George J, "Food insecurity and public assistance," *Journal of Public Economics*, 2004, *88* (7), 1421–1443.

Coleman-Jensen, Alisha, MP Rabbitt, and CA Gregory, "Household Food Security in the United States in 2021, ERR-309, US Department of Agriculture," *Economic Research Service*, September 2022, pp. 1–40.

Corman, Hope, Dhaval M. Dave, Ofira Schwartz-Soicher, and Nancy E. Reichman, "Effects of welfare reform on household food insecurity across generations," *Economics & Human Biology*, 2022, *45*, 101101.

Ditzen, Jan, "XTCSE2: Stata module to estimate the exponent of cross-sectional dependence in large panels," Statistical Software Components, Boston College Department of Economics July 2019.

Eberhardt, Markus, "MULTIPURT: Stata module to run 1st and 2nd generation panel unit root tests for multiple variables and lags," Statistical Software Components, Boston College Department of Economics February 2011.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J Robert Warren, and Michael Westberry, "Integrated public use microdata series, current population survey: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2021," *URL https://doi. org/10.18128 D*, 2022, *30.*

Goldberger, Arthur S., "The Interpretation and Estimation of Cobb-Douglas Functions," *Econometrica*, 1968, *36* (3/4), 464–472.

Gösser, Niklas and Nima Moshgbar, "Smoothing time fixed effects," DICE Discussion Paper 343, Düsseldorf 2020.

Granger, Clive WJ and Ling-Ling Huang, "Evaluation of Panel Data Models: Some suggestions from time series. Discussion Paper 97-10," April 1997. Available at SSRN: https://ssrn.com/abstract=52400 or http://dx.doi.org/10.2139/ssrn.52400.

Gundersen, Craig, "SNAP: The Most Important Component of Our Efforts to End Food Insecurity in the United States," *American Journal of Public Health*, 2022, *112* (10), 1370–1371. PMID: 35981279.

_ , "Lessons Learned from the Lack of Increase in Food Insecurity during COVID-19.," *J Nutr*, Apr 2023, *153* (4), 920–921.

_ **and James P Ziliak**, "Food Insecurity Research in the United States: Where We Have Been and Where We Need to Go," *Applied Economic Perspectives and Policy*, 2018, *40* (1), 119–135.

_ **, Brent Kreider, John Pepper, and Valerie Tarasuk**, "Food assistance programs and food insecurity: implications for Canada in light of the mixing problem," *Empirical Economics*, 2017, *52* (3), 1065–1087.

_ **, Emily Engelhard, and Elaine Waxman**, "Map the Meal Gap: Exploring Food Insecurity at the Local Level," *Applied Economic Perspectives and Policy*, 2014, *36* (3), 373–386.

_ **, Monica Hake, Adam Dewey, and Emily Engelhard**, "Food Insecurity during COVID-19.," *Appl Econ Perspect Policy*, Mar 2021, *43* (1), 153–161.

Han, Jeehoon, "The Impact of SNAP on Material Hardships: Evidence From Broad-Based Categorical Eligibility Expansions," *Southern Economic Journal*, 2016, *83* (2), 464–486.

Henly, Megan, Debra L. Brucker, and Alisha Coleman-Jensen, "Food insecurity among those with disability: Cross-survey comparison of estimates and implications for future research," *Applied Economic Perspectives and Policy*, 2022, *n/a* (n/a).

Im, Kyung So, M.Hashem Pesaran, and Yongcheol Shin, "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 2003, *115* (1), 53–74.

Lentz, E.C., H. Michelson, K. Baylis, and Y. Zhou, "A data-driven approach improves food insecurity crisis prediction," *World Development*, 2019, *122*, 399–409.

Marques, Emanuele S, Michael E Reichenheim, Claudia L de Moraes, Marina ML Antunes, and Rosana Salles-Costa, "Household food insecurity: a systematic review of the measuring instruments used in epidemiological studies," *Public Health Nutrition*, 2015, *18* (5), 877–892.

McKernan, Signe-Mary, Caroline Ratcliffe, and Breno Braga, "The Effect of the US Safety Net on Material Hardship over Two Decades.," *J Public Econ*, May 2021, *197.*

Pesaran, M. Hashem, "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence," *Journal of Applied Econometrics*, 2007, *22* (2), 265–312.

Restrepo, Brandon J., Matthew P. Rabbitt, and Christian A. Gregory, "The Effect of Unemployment on Food Spending and Adequacy: Evidence from Coronavirus-Induced Firm Closures," *Applied Economic Perspectives and Policy*, 2021, *43* (1), 185–204.

Schanzenbach, Diane Whitmore, "Understanding SNAP: An overview of recent research," *Food Policy*, 2023, *114*, 102397.

Smith, Travis A. and Christian A. Gregory, "Food Insecurity in the United States: Measurement, Economic Modeling, and Food Assistance Effectiveness," *Annual Review of Resource Economics*, 2023, *15* (1), null.

Stacy, Brian, Laura Tiehen, and David Marquardt, "Using a policy index to capture trends and differences in state administration of USDA's Supplemental Nutrition Assistance Program," Technical Report, U.S. Department of Agriculture, Economic Research Service 2018.

Tiehen, Laura, Cody N Vaughn, and James P Ziliak, "Food insecurity in the PSID: A comparison with the levels, trends, and determinants in the CPS, 1999–2017," *Journal of Economic and Social Measurement*, 2020, *45* (2), 103–138.

Wang, Dieter, Bo Pieter Johannes Andrée, Andres Fernando Chamorro, and Phoebe Girouard Spencer, "Transitions into and out of food insecurity: A probabilistic approach with panel data evidence from 15 countries," *World Development*, 2022, *159*, 106035.

West, Kenneth D., "Chapter 3 Forecast Evaluation," in G. Elliott, C.W.J. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 1, Elsevier, 2006, pp. 99–134.

6 Figures

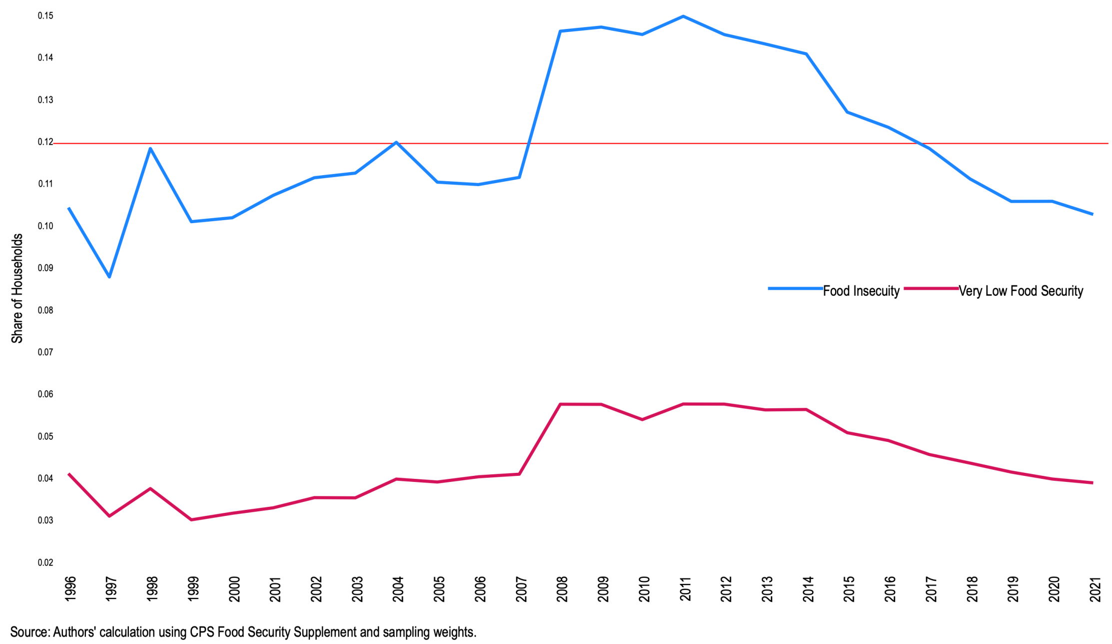


Figure 1. Prevalence of food insecurity over the years. A U.S. national average.

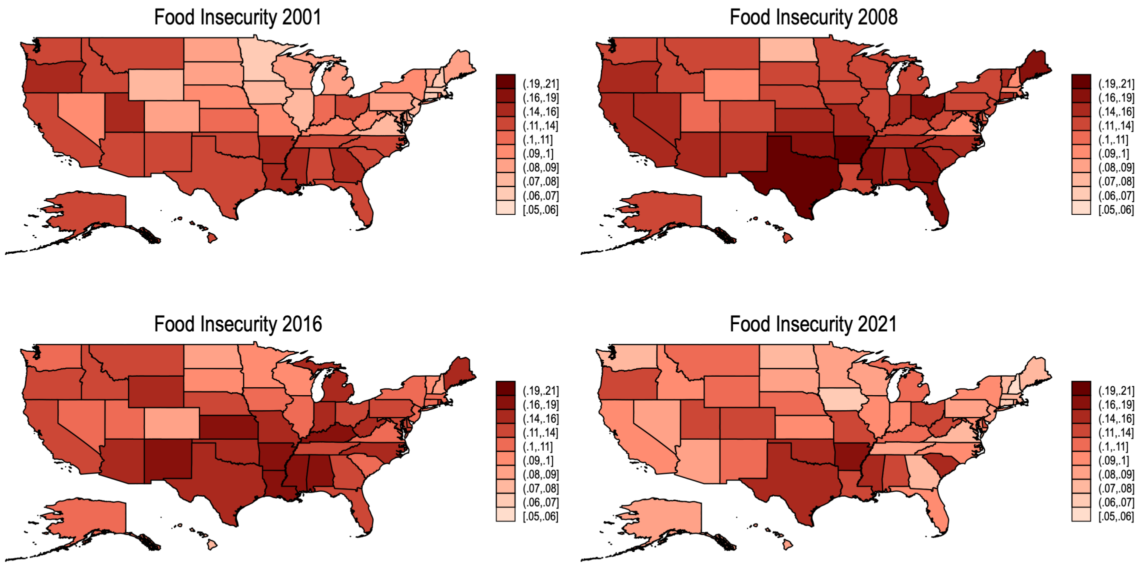


Figure 2. State-level variation in food insecurity rates over the years.

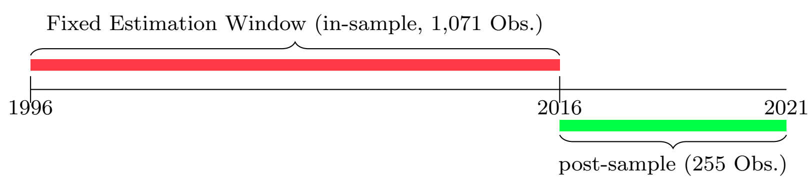


Figure 3. Figurative representation of a fixed estimation forecast scheme for our modelling.

Figure 4. Top 20 models hair graph for out-of-sample projections with the best model's point-forecast estimates overlayed.

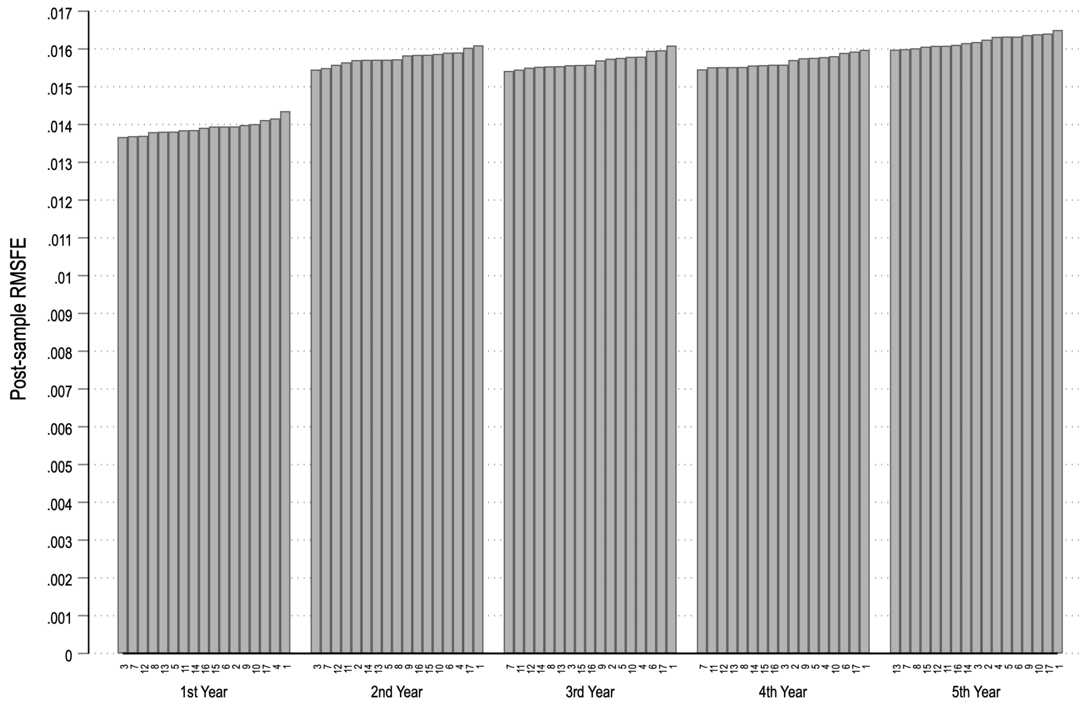


Figure 5. Comparison of post-sample (2017-2021) RMSFE of forecasts across the top 20 models - Household food insecurity rates in the US.

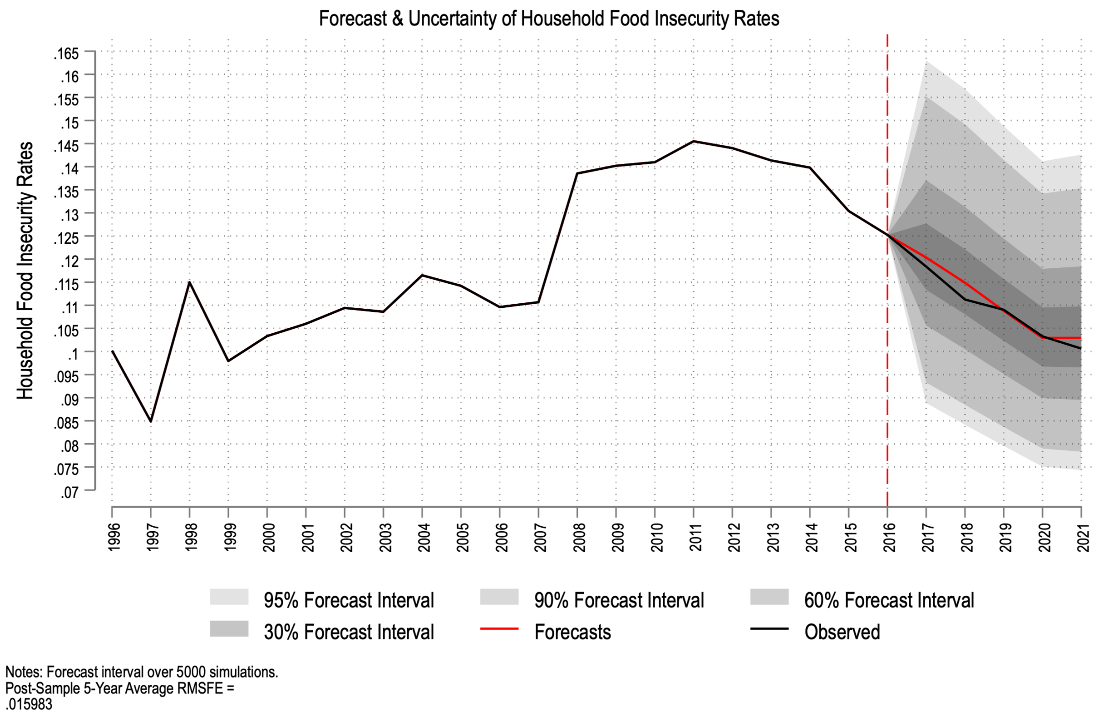


Figure 6. Forecast Interval: Forecast uncertainty (percentiles) around point forecasts from our best model. Comparison between simulated and estimated point forecasts.

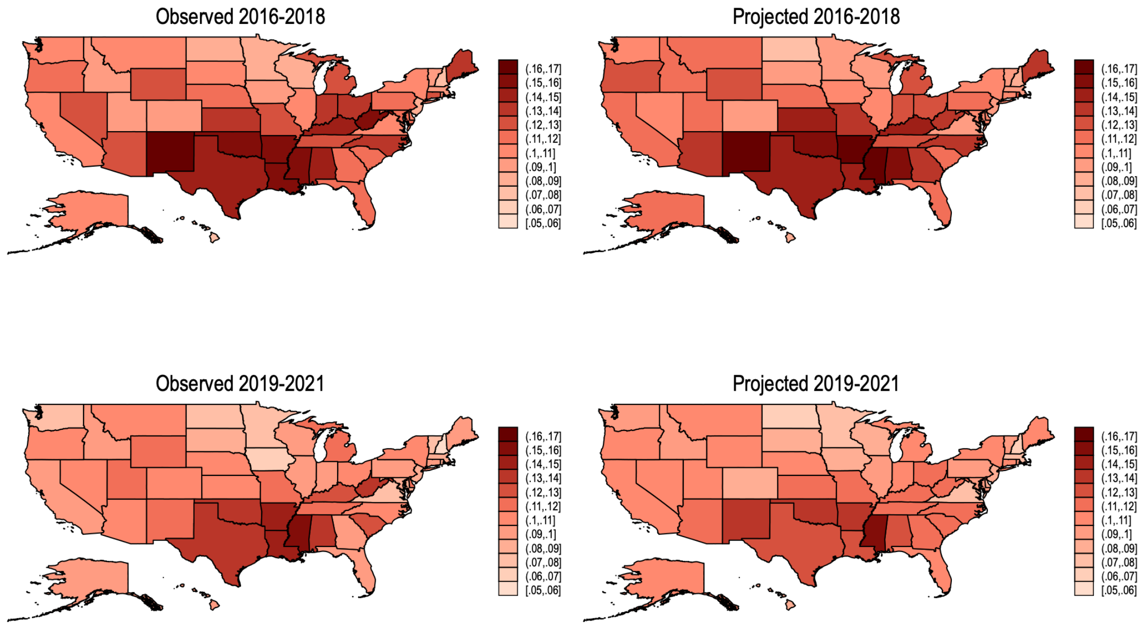


Figure 7. State-level map showing three-year average of observed and projected household food insecurity intervals.

7 Tables

Table 1. Conceptual framework describing the relationship between food insecurity outcome and the modelling input.

State-Level Determinants of Household Food Insecurity		
Economic	Local Policies	Demographics
Unemployment rate (%)	Real State minimum wage	Age composition
House Price Index (HPI)	SNAP participation per-capita	Race/Ethnicity
Poverty rate (%)	Broad SNAP policy index	Education
Real Personal income per-capita	Narrow SNAP policy index	Home ownership
Real State GDP per-capita	Real Average SNAP benefits	

Table 2. Summary Statistics of Household Food Insecurity Rates and its Determinants

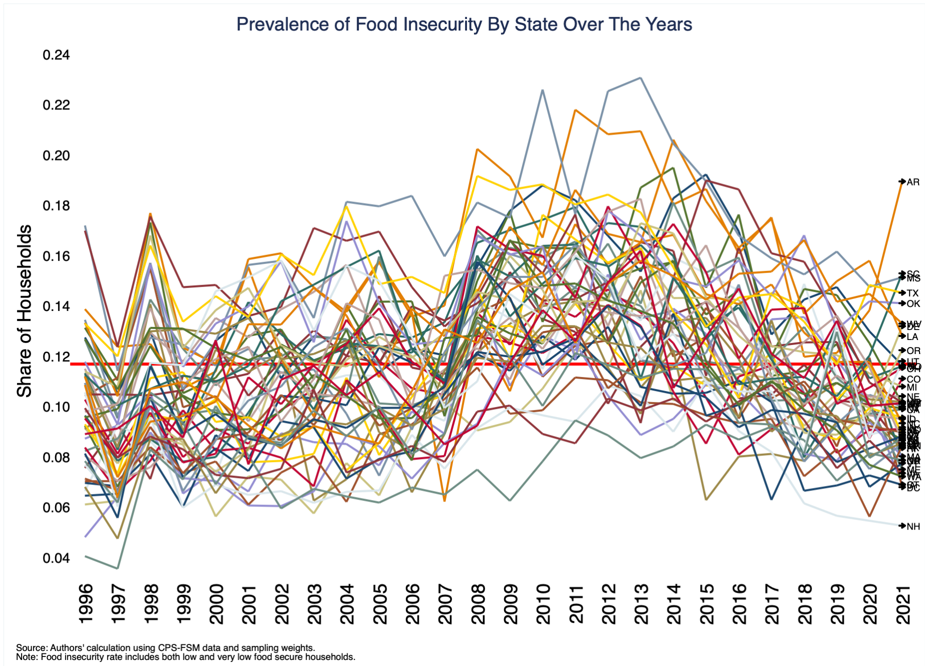
Variable	Mean	SD	Min	Max	Source
Outcome:					
Household Food Insecurity Rate	0.12	0.03	0.04	0.23	CPS-FSM
State Economic Condition					
Unemployment Rate (%)	5.42	1.95	2.11	13.78	BLS
House Price Index (Log)	5.72	0.37	4.83	6.91	US FHFA
Poverty Rate (%)	13.04	3.24	5.60	23.90	Census SAIPE
Household Personal Income (Real USD, log per-capita)	3.90	0.18	3.46	4.57	BEA
State Policies					
Effective Minimum Wage (Real USD, log)	2.14	0.14	1.90	2.74	BLS
SNAP Participation (per-capita, log)	-2.33	0.44	-3.54	-1.38	FNS
Average SNAP Benefits (per-person)	110.04	38.84	55.77	420.15	FNS
Narrow SNAP Policy Indices:					
Overall SNAP Index	0.52	0.21	0	0.89	FNS
Eligibility Sub-Index	0.32	0.20	0	0.75	FNS
Transaction Sub-Index	0.52	0.36	0	1	FNS
Stigma Sub-Index	0.92	0.18	0	1	FNS
Broad SNAP Policy Indices:					
Overall SNAP Index	0.48	0.18	0.11	0.87	FNS
Eligibility Sub-Index	0.30	0.20	0	0.70	FNS
Transaction Sub-Index	0.57	0.20	0.17	1	FNS
Stigma Sub-Index	0.58	0.27	0	1	FNS
State Demographics					
Share of Hispanic Population	0.07	0.08	0.00	0.43	CPS-ASEC
Share of Non-White Population	0.17	0.13	0.00	0.75	CPS-ASEC
Share of home owners	0.68	0.06	0.38	0.82	CPS-ASEC
Share of High-School Or Less Education	0.42	0.08	0.20	0.68	CPS-ASEC
Share of Age 0 to 14	0.20	0.02	0.14	0.28	US Census
Share of Age 15 to 29	0.21	0.01	0.17	0.29	US Census
Share of Age 30 to 44	0.21	0.02	0.17	0.28	US Census
Share of Age 45 to 59	0.20	0.02	0.13	0.24	US Census
Share of Age 60 to 74	0.13	0.03	0.06	0.20	US Census

Note: CPS-FSM: Current Population Survey (CPS) - Food Security Module (FSM); BLS: Bureau of Labor Statistics; US FHFA: U.S. Federal Housing Finance Agency; Census SAIPE: US Census Bureau's Small Area Income and Poverty Estimates; BEA: Bureau of Economic Analysis; FNS: USDA's Food & Nutrition Services; CPS-ASEC: CPS' Annual Social and Economic Supplement.

Table 3. In-Sample Coefficient Estimates of the Top 4 Best Performing Models

	Model 1	Model 2	Model 3	Model 4	Model 5
One-Year Lag of Log HH Food Insecurity	0.165***	0.166***	0.164***	0.158***	0.165***
One-Year Lag Unemp. Rate	0.002**	0.002**	0.002**	0.002***	0.002**
Log House Price Index	-0.032***	-0.033***	-0.032***	-0.034***	-0.032***
Poverty Rate	0.002**	0.002**	0.002**	0.003***	0.002**
Broad SNAP Eligibility Sub-Index	-0.026***		-0.025***	-0.022***	
Broad SNAP Transaction Sub-Index	0.003			0.005	
Broad SNAP Stigma Sub-Index	0.002			0.005	
Log of SNAP Participation				-0.015**	
Narrow SNAP Eligibility Sub-Index		-0.021***			-0.021***
Narrow SNAP Transaction Sub-Index					-0.003
Narrow SNAP Stigma Sub-Index					-0.001
MFE	-0.001466	-0.001211	-0.001821	-0.002192	-0.001543
RMSFE	0.015983	0.015996	0.016018	0.016061	0.016084

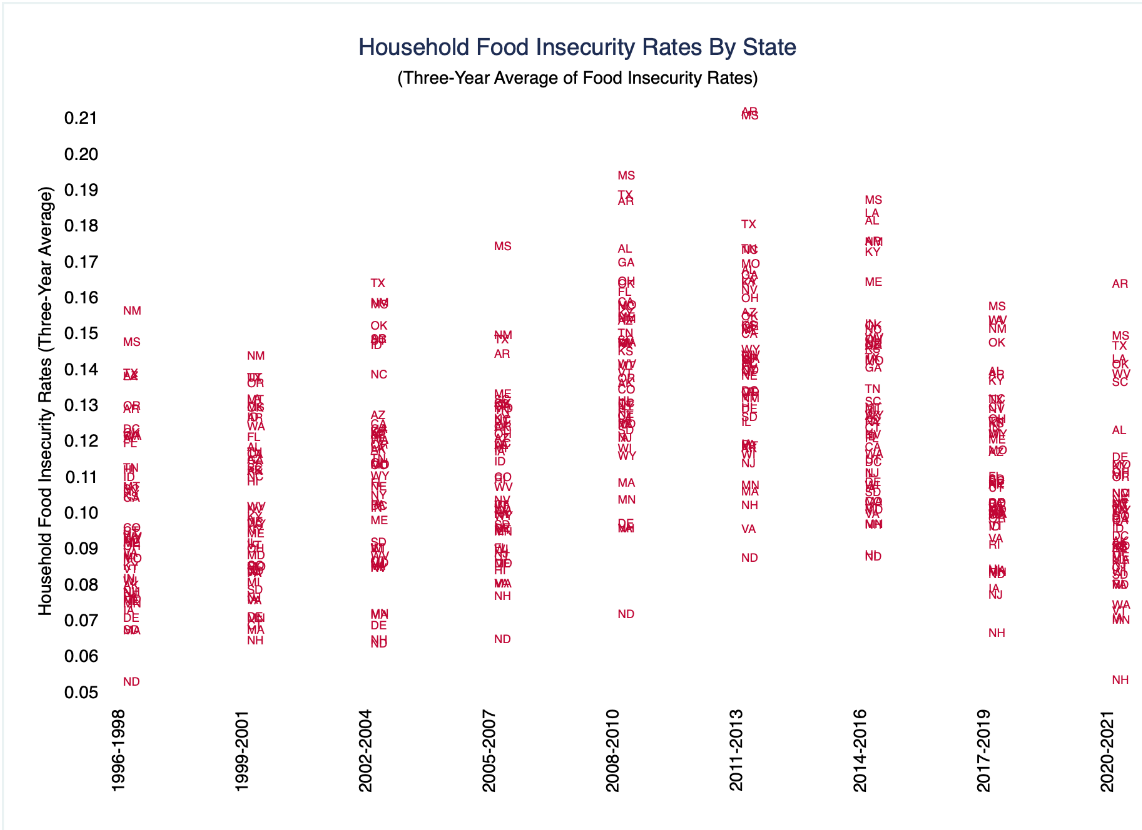
A Appendix Tables and Figures



Appendix Figure A1. State-level variation in food insecurity rates over the years.

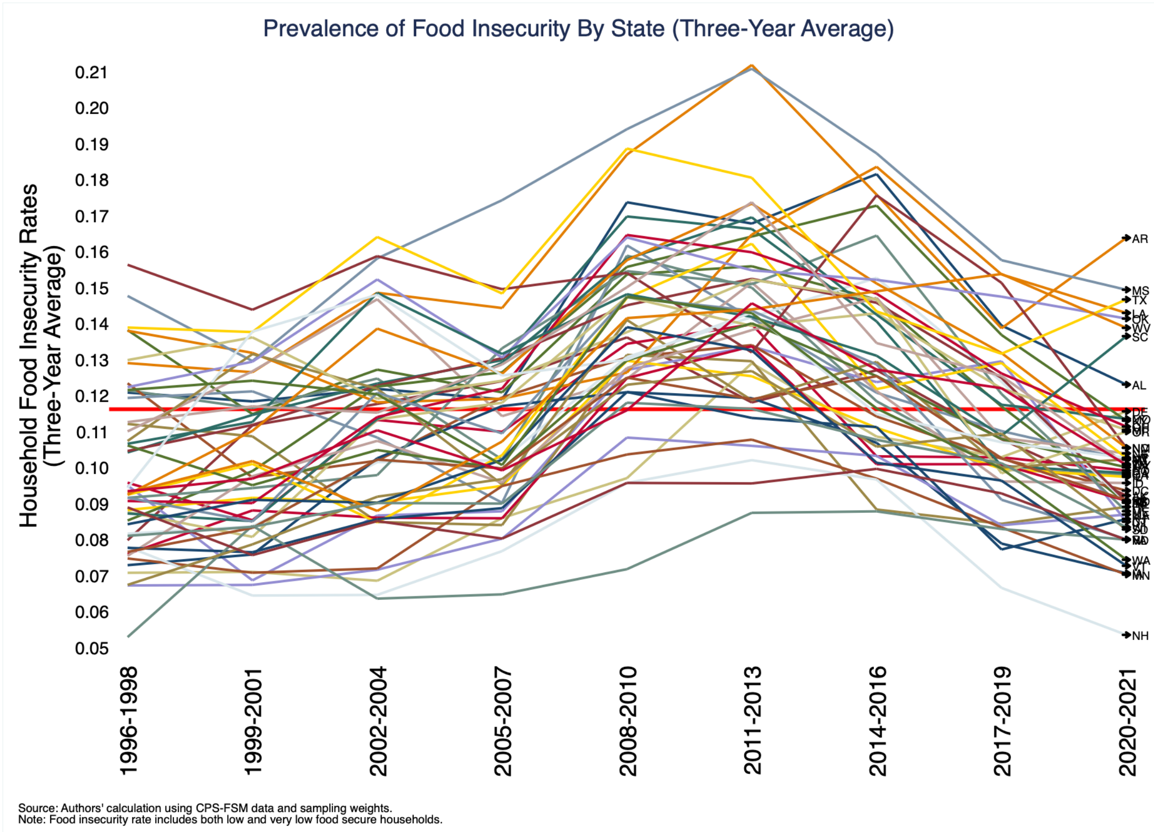
A.1 Three-Year Average State-Level Variation

In figure A2 we display a scatter plot showing how each state varied from low food insecurity rates to high food insecurity rates over the three-year periods. We can see that in the first three-year period (1996-1998) New Mexico (NM) had the highest annual household food insecurity rates of around 15.5% and North Dakota (ND) had the lowest food insecurity rate of around 5.5%. However, by the last and current three-year period (2020-2021), Arkansas (AR) has experienced high food insecurity rate of more than 16% and New Hampshire (NH) has the lowest household food insecurity rates of around 5%. The figure A2 also shows how during the Great Recession all states experienced a level shift upwards in their food insecurity rates.



Appendix Figure A2. Three-year average of State-level variation in household food insecurity rates.

Figure A3 is an alternative visualization showing how the trend in food insecurity rates varied over the three year periods.



Appendix Figure A3. Three-year average trend in State's household food insecurity rates.

A.2 SNAP Policy Index Components

Appendix Table A1. State-Level SNAP Policies Used in the Construction of SNAP Policy Indices

Sub-Index Category	Broad Policy Index (28 Policies)	Narrow Policy Index (9 Policies)
A) Eligibility	1) BBCE implemented 2) BBCE income limit more than 130 3) BBCE removes Asset Test 4) BBCE has Some Asset Limit 5) BBCE excludes at lest one vehicle 6) The State excludes all vehicles in the household from the SNAP asset test. 7) The State excludes at least one, but not all, vehicles in the household from the SNAP asset test. 8) All non-citizen adults allowed 9) All non-citizen children allowed 10) All non-citizen elderly allowed	1) Exempts at least one but not all vehicles from SNAP asset test 2) Exempts all vehicles from SNAP asset test 3) Broad-based categorical eligibility (BBCE) 4) Eligibility restrictions for adult non-citizens
B) Transaction Cost	11) The State operates a Combined Application Project 12) Simplified reporting 13) No SNAP units with earnings with 1-3 month recertification periods. 14) No SNAP units with earnings with 4-6 month recertification periods. 15) No elderly SNAP units with earnings with 3-6 month recertification periods. 16) No elderly SNAP units with earnings with 4-6 month recertification periods. 17) SNAP units with earnings with 7-12 month recertification periods. 18) SNAP units with earnings with 13 plus month recertification periods. 19) Elderly SNAP units with earnings with 7-12 month recertification periods. 20) Elderly SNAP units with earnings with 13 plus month recertification periods. 21) The median certification period greater than 6 months for SNAP units. 22) The median certification period greater than 12 months for elderly SNAP units.	1) Proportion of working households with short recertification periods (1-3 months) 2) Simplified reporting 3) Online application availability
C) Stigma	23) SNAP benefits issued by EBT 24) Face waiver For initial and or recertification 25) No fingerprinting of SNAP applicants Statewide. 26) Online application 27) Transitional benefits 28) The State operates call centers	1) State benefits issued via electronic benefits transfer (EBT) 2) Fingerprinting required during application

A.3 Panel Unit-Root Test

Table A2 provide the test results (p-values) of panel unit root testing of household food insecurity rates and its determinants. The null hypothesis of the test being that the variable has a unit root. We conduct the second-generation panel unit root test that builds on the first-generation panel unit root test by accounting for heterogeneous panels and cross-section dependence (Eberhardt, 2011; Pesaran, 2007; Im et al., 2003). Overall, we find that in table A2 our outcome variable of interest, annual household food insecurity rates is stationary (as we reject the null of unit root presence at 1% level of significance). Furthermore, most of the variables are stationary in table A2, apart from a few. The variables where we fail to reject unit root are owing to the annual average of monthly frequency observations (unemployment rate , snap participation, avaerage SNAP benefits). We could take first-difference of these variable to address non-stationarity , however, that would result in losing variation across states and time. And, furthermore, since our outcome of interest is stationary, we proceed with a level regression specification.

Appendix Table A2. Second-Generation Panel Unit Root Test: With and Without Trend

	Constant		Trend	
	Lag 0	Lag 1	Lag 0	Lag 1
	(p-values)	(p-values)	(p-values)	(p-values)
Outcome:				
Household Food Insecurity Rate	0.000	0.000	0.000	0.000
Log of Household Food Insecurity Rate	0.000	0.000	0.000	0.000
State Economic Condition				
Unemployment Rate (% Annual, Seasonally Adj.)	0.593	0.736	0.989	0.992
Log of House Price Index (Base 1980)	1.000	0.003	1.000	0.997
Poverty Rate	0.000	0.002	0.000	0.019
Personal Income Per-Capita	0.870	0.028	1.000	0.999
State Policies				
Log of Eff. Min. Wage	1.000	0.994	1.000	1.000
Log of SNAP per-capita	0.957	0.470	0.999	0.506
Avg. SNAP Benefits	0.000	0.124	0.610	0.963
Narrow SNAP Policy Indices:				
Overall SNAP Index	0.000	0.000	0.000	0.018
SNAP Eligibility Sub-Index	0.000	0.000	0.052	0.526
SNAP Transaction Sub-Index	0.000	0.000	0.000	0.000
SNAP Stigma Sub-Index	0.000	1.000	0.015	1.000
Broad SNAP Policy Indices:				
Overall SNAP Index	0.000	0.000	0.007	0.004
SNAP Eligibility Sub-Index	0.000	0.000	0.720	0.047
SNAP Transaction Sub-Index	0.000	0.000	0.000	0.005
SNAP Stigma Sub-Index	0.000	0.000	0.049	0.000

A.4 Cross-Section Dependence Test (CD Test)

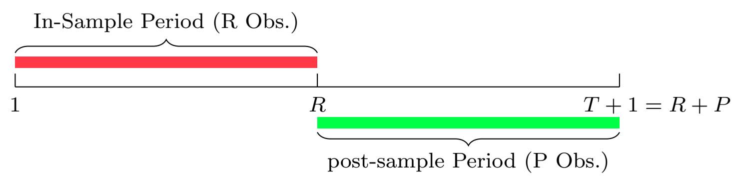
Table A3 provides the test results of weak/strong cross-section dependence test across panels (Ditzen, 2019). The CD test tests the presence of weak and/or strong cross-section dependence across panels. The exponent of CD being greater than 0.5 is a strong indicator of strong cross-section dependence. In table A3 , we see that all variables exhibit a strong cross-section dependence. This CD test result supports the use of second-generation panel unit root in table A2.

Appendix Table A3. Cross-Section Dependence Test Results and Estimated Exponent of Cross-Sectional Dependence

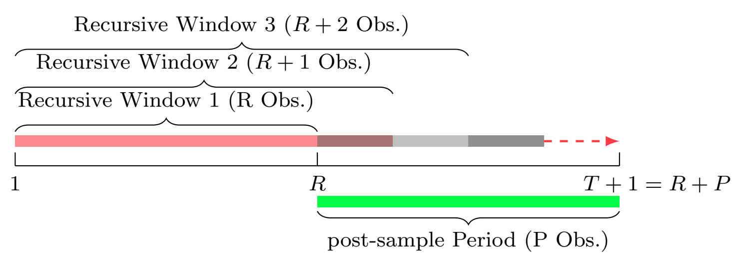
	CD Test			Exponent of CD		
	CD	CDw	CDwplus	$\hat{\alpha}^*_{0.05}$	$\hat{\alpha}$	$\hat{\alpha}^*_{0.95}$
Log of HHFI	86.833	-0.921	3079.376	0.921	1.005	1.089
HHFI	89.269	1.137	3168.634	0.927	1.005	1.083
Unemp. Rate	149.745	-2.402	5344.546	0.914	1.005	1.096
Log of HPI	169.482	-2.778	6048.940	0.882	1.005	1.128
Poverty Rate	129.628	-2.283	4620.551	0.933	1.005	1.077
Personal Inc. per-capita	163.191	-0.317	5826.238	0.897	1.005	1.113
Log of Eff. Min. Wage	80.201	-2.669	3161.188	0.952	1.005	1.058
Log of SNAP per-capita	158.486	-0.228	5658.853	0.961	1.005	1.049
Narrow SNAP Index	161.422	-2.210	5761.720	0.955	1.005	1.055
Narrow SNAP Eligibility	132.897	12.996	4749.737	0.949	1.005	1.061
Narrow SNAP Transaction	149.048	-2.050	5319.428	0.959	1.005	1.051
Narrow SNAP Stigma	73.996	-1.817	2654.309	.	.	.
Broad SNAP Index	162.626	-3.432	5803.468	0.953	1.005	1.057
Broad SNAP Transaction	96.174	-1.080	3872.278	0.884	0.994	1.104
Broad SNAP Stigma	160.610	4.172	5739.093	0.927	1.005	1.083
Broad SNAP Eligibility	136.237	-3.326	4848.952	0.965	1.005	1.045
Avg. SNAP Benefits	178.643	-3.534	6375.301	0.903	1.005	1.107
Num. of States	51	51	51	51	51	51
Time Periods	26	26	26	26	26	26

A.5 Forecast Estimation Scheme

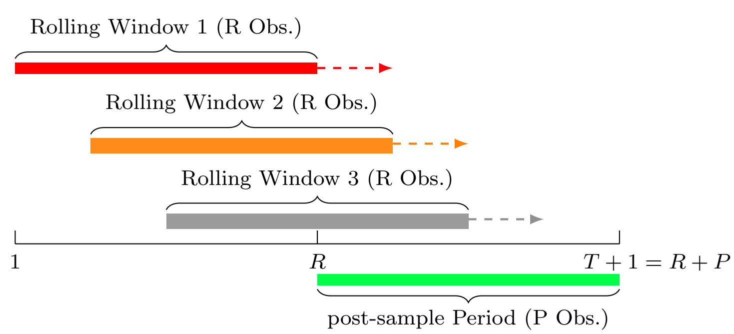
In modelling point-forecast estimates, our preferred model in consideration relies on the estimated regression parameters. We are interested in one-step ahead prediction errors (in our application, one-step ahead would be one-year ahead.). We divide our observed data (say, T+1 observations as total sample size) into in-sample periods (say, using R observations) and post-sample periods (say, using the last P observations). Here, $R + P = T + 1$. We will use the first R observations to estimate our regression parameters to provide our first prediction and use the last P observations, for forecast evaluation. Schematically:



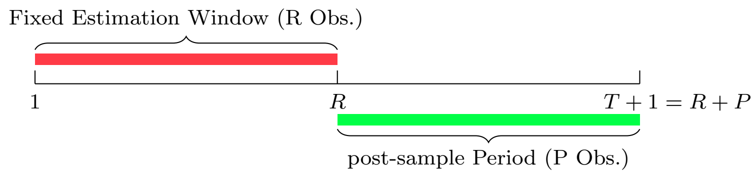
In the forecasting literature, the regression parameters (β's) are estimated using three distinct schemes: recursive, rolling and fixed estimation. In recursive scheme, the starting period of estimation is fixed and the end period is sequentially increased as one makes predictions for successive observations. Schematically:



In a rolling estimation scheme, the starting and end period are movable with the constraint that the sample size between the starting and end periods remain the same. Thus, in rolling the regression parameters are always estimated on a sample size of R. Schematically:



Finally, in the fixed scheme, the first R observations are used to estimate β's just once. And, with those estimated β's, we obtain the point forecast for the post-sample period observations. Schematically:



The preference of the type of forecast estimation scheme relies on the instability of the estimated parameters. West (2006) suggests if it is computationally difficult to update the parameter estimates the preferred scheme is the *fixed scheme*. Otherwise, a *rolling scheme* would be the go-to process to "*to guard against moment or parameter drift that is difficult to model explicitly*".[19]

In each of the forecast estimation scheme, the one-step ahead prediction error for a simple dynamic AR(1) model, for e.g., $y_t = \beta y_{t-1} + e_t$ is:

$$\hat{e}_{t+1} \equiv y_{t+1} - y_t\hat{\beta}_R \qquad \ldots\texttt{[Fixed Scheme]}$$
$$\hat{e}_{t+1} \equiv y_{t+1} - y_t\hat{\beta}_t \qquad \ldots\texttt{[Recursive/Rolling Scheme]}$$

where, $\hat{\beta}_t = \hat{\beta}_R$ for a fixed scheme and for a recursive/rolling scheme $\hat{\beta}_t$ varies with the window size.

To evaluate forecast performance across models, our object of interest is the (Root) Mean Squared Prediction Error (MSPE). To illustrate the steps in calculating the MSPE, let's say we have two linear models:

$$y_t = X_{1t}^{'}\beta_1^* + e_{1t}$$
$$y_t = X_{2t}^{'}\beta_2^* + e_{2t}$$

The population MSPEs are:

$$\sigma_1^2 \equiv Ee_{1t}^2$$
$$\sigma_2^2 \equiv Ee_{2t}^2$$

We can calculate the sample one-step ahead (one-year ahead) forecast errors and

[19]See page 107 of West (2006) for elaborate discussion on the choice of scheme

sample MSPEs as:

$$
\begin{aligned}
\hat{e}_{1t+1} &\equiv y_{t+1} - X'_{1t+1}\hat{\beta}_{1t} \\
\hat{e}_{2t+1} &\equiv y_{t+1} - X'_{2t+1}\hat{\beta}_{2t} \\
\hat{\sigma}_1^2 &= P^{-1}\sum_{t=R}^{T}\hat{e}_{1t+1}^2 \qquad [RMSPE : \sqrt{\hat{\sigma}_1^2}] \\
\hat{\sigma}_2^2 &= P^{-1}\sum_{t=R}^{T}\hat{e}_{2t+1}^2 \qquad [RMSPE : \sqrt{\hat{\sigma}_2^2}]
\end{aligned}
$$

In our forecast modelling framework, here, based on the recommended procedures in West (2006), we also conduct our forecast estimates using both fixed estimation scheme and rolling window estimation scheme and compare the forecast performance. Our post-sample periods in both schemes are the five year period following 2016 (i.e., 2017-2021). We will calculate our five-year post-sample root mean squared forecast error (OOS-RMSFE) based on the post-sample periods' forecast estimates, as in equation below:

$$
\text{post-sample RMSFE} = \frac{1}{\#\text{ of States}}\sqrt{P^{-1}\sum_{t=R}^{T}\hat{e}_{t+1}^2} \qquad (3)
$$