Forecasting Nevada Gross Gaming Revenue and Taxable Sales Using Coincident and Leading Employment Indexes*

Abstract

This paper provides out-of-sample forecasts of Nevada gross gaming revenue and taxable sales using a battery of linear and non-linear forecasting models and univariate and multivariate techniques. The linear models include vector autoregressive and vector error-correction models with and without Bayesian priors. The non-linear models include non-parametric and semi-parametric models, smooth transition autoregressive models and artificial neural network autoregressive models. In addition to gross gaming revenue and taxable sales, we employ recently constructed coincident and leading employment indexes for Nevada's economy. We conclude that non-linear models generally outperform linear models in forecasting future movements in gross gaming revenue and taxable sales.

Keywords: Forecasting, Linear and non-linear models, Nevada gross gaming revenue, Nevada taxable sales

JEL classification: C32, R31

* We thank the Associate Editor and an anonymous referee for many helpful comments that substantially improved the quality of the paper. All remaining errors are, however, solely ours.

** *Corresponding author*

1. Introduction

The Great Recession in the US creates significant challenges for state governments as they plan for future budgets. That is, most state governments live under a balanced operating budget from year-to-year. Forecasting state revenues, therefore, becomes a significant task.

Nevada faces the most severe decline in state tax revenue in its history. In prior recessions, Nevada's economy barely noticed the national recession and continued to grow, or at least experienced a much lower decline in its economy than the nation as a whole. The Great Recession that witnessed significant declines in the leisure and hospitality, construction, and finance, insurance, and real estate sectors caused Nevada to experience its worst recession ever.

Gross gaming revenue and taxable sales comprise two major components of Nevada's tax base, generating sales and use taxes, and gaming taxes respectively. In fiscal 2008, sales and use taxes constituted 35.6 percent of total Nevada taxes while gaming taxes constituted 29.0 percent. On total general fund revenue, which adds licenses, fees, fines and other revenues to taxes, sales and use taxes comprise 32.3 percent and gaming taxes comprise 26.3 percent. In sum, well over 50 percent of Nevada revenues come from the taxable sales and gross gaming revenue tax bases.

Thus, this paper provides out-of-sample forecasts of Nevada gross gaming revenue and taxable sales using a battery of linear and non-linear forecasting models and univariate and multivariate techniques. Linear models include vector autoregressive (VAR), Bayesian VAR (BVAR), vector error-correction (VEC), and Bayesian VEC (BVEC) models. Non-linear models include semi-parametric (SP), non-parametric (NP), smooth transition autoregressive (STAR), and artificial neural network (ANN) models. In addition to the two components of Nevada's tax base, this paper also employs recently constructed CBER-DETR Nevada Coincident and Leading Employment indexes to capture the state of the Nevada economy.

We organize the rest of the paper as follows. Section 2 discusses the construction of the Nevada coincident and leading Employment indexes. Section 3 reviews the existing literature on forecasting Nevada taxable sales and gross gaming revenue. Section 4outlines the various methodologies used to forecast Nevada gross gaming revenue and taxable sales – VAR, BVAR, VEC, and BVEC models, semi-parametric and non-parametric models, smooth transition autoregressive models, and artificial neural network models. Section 5 describes the data and reports the results of the various linear and non-linear forecasting methods. Section 6 concludes.

2. Coincident and Leading Employment Indexes for Nevada

Coincident indexes include a number of economic series that collectively represent the current state of the economy. Each series in a coincident index contains some information about the turning points in the business cycle. Since series do not all show the same turning points, a coincident index provides a collective call on the business cycle. This averaging process produces better information about cyclical turning points than any one of the individual series in the index can generate on their own.

Leading indexes provide valuable information about the future path of the economy, combining information from several economic series and collectively forecasting future movements in the economy. As with coincident indexes, each series provides some information but it is unlikely that the individual series will show identical turning points. The combined information in leading indexes produces better predictions about future turning points.

Dua and Miller (1995) construct Connecticut coincident and leading employment indexes following well-developed procedures used by the Department of Commerce and described in U.S. Department of Commerce (1977, 1984) and in Niemira and Klein (1994). These procedures adopt methods developed by National Bureau of Economic Research researchers Geoffrey H. Moore and Julius Shiskin in the 1950s.

Several characteristics of a time series are evaluated to select the components of a composite index. The most important of these is cyclical timing determined by the consistency with which the cyclical turning points in a series coincide with or lead the business cycle turns. Other factors include the periodicity of the data, their reliability, and the promptness with which they are available. The components are standardized to prevent the more volatile series from dominating the index.

Dua and Miller (1995) base their indexes on employment-related time series only due to data availability constraints. The components of the indexes are available monthly with a short time lag and are reliable. Since employment conditions generally mirror overall economic activity, the coincident and leading indexes serve as measures of current and future economic activity, respectively.

The Nevada coincident employment index comprises four individual employment-related series that track current employment activity - the total unemployment rate (inverted), the insured unemployment rate (inverted), nonfarm employment, and total (household) employment. The index, therefore, combines information from different sources. Total (household) employment and the unemployment rate are based on a survey of about 600 Nevada households. The insured unemployment rate, on the other hand, comes from the data on unemployment insurance claims filed with the state. Finally, nonfarm employment is based on a survey of employers by the state.

The Nevada leading employment index includes six components that predict employment activity - the initial claims for unemployment insurance (inverted), the short-duration (less than

15 weeks) unemployment rate (inverted), housing permits, commercial permits, construction employment, and the real Moody's Baa interest rate (inverted). Each variable has some intuitive appeal. The initial claims for unemployment insurance is one of the first steps taken by someone who loses his/her job. So, this variable quickly reflects job market changes. The short-duration unemployment rate measures changes in those unemployed for 15 weeks or less. This variable also quickly reflects changes in the job market. Housing and commercial permits captures the intention to build in the near future, which closely relates to construction employment. Finally, the Baa interest rate, although a national variable, provides useful leading information for the state (Banerji, Dua, and Miller 2006).

Figures 1 and 2 report the coincident and leading indexes using data through December 2009. At that time, it looked like the coincident and leading indexes bottomed in October 2009 All data are seasonally adjusted and come from DETR, CBER, and the Federal Reserve Bank of St. Louis FRED® data. The description of the construction method is posted at http://cber.unlv.edu/nvindices.pdf. Data availability restricts our coverage in the two indexes to monthly series beginning in January 1976. The data series for household employment, nonfarm employment, the unemployment rate, initial claims, and the real Moody's Baa bond rate all begin in January 1976. Housing permits and the insured unemployment rate begins in January 1980 and March 1987, respectively. Commercial permits, construction employment, and the short-duration unemployment rate begin in January 1988, January 1990, and January 2001, respectively. Thus, the coincident index uses three series through March 1987, when we add the insured unemployment rate. The leading index begins with two series and adds housing permits in January 1980, commercial permits in January 1988, construction employment in January 1990, and finally, the short-duration unemployment rate in January 2001.

In sum, the Nevada coincident employment index contains four individual series that gauge current economic activity. By construction, it includes more information than that in a single measure of economic activity such as the unemployment rate. Likewise, the leading employment index contains six individual series that predict future economic activity.

3. Literature Review

Several papers model and forecast Nevada gross gaming revenue. Cargill and Eadington (1978) develop simple models of Nevada gross gaming revenue – structural and autoregressive integrated moving average (ARIMA). They actually investigate gross gaming revenue for three different regions in Nevada – Las Vegas, Reno-Sparks, and South Lake Tahoe. The single equation structural models include California personal income and dummy variables for the 1973-74 energy crisis and recessions. Using quarterly data, the ARIMA models employ one regular difference and one seasonal difference with one autoregressive term.[1] They estimate the ARIMA model from 1955:Q1 to 1974:Q4 and forecast 1975:Q1 to 1977Q4, reporting the errors as a percent of the actual values.[2]

Cargill and Morus (1988) develop an eight variable Bayesian vector autoregressive (BVAR) model of the Nevada economy. The model includes four national drivers – real GNP, the GNP deflator, employment, and the 4-6 month commercial paper rate -- and one California driver – employment -- along with three Nevada variables – gross gaming revenue, taxable sales,

---

[1] Eisendrath, Bernhard, Lucas, and Murphy (2008) use intervention analysis to determine the effect of the terrorist attack on 9/11 on Nevada gross gaming revenue. They conclude that the recovery from the attack occurred rather quickly with most recovery occurring in five months and complete recovery within two years.

[2] Traditionally, for forecasting purposes, time-series models generally perform as well as or better than dynamic structural econometric specifications. Zellner and Palm (1974) provide the theoretical rationalization. Any dynamic structural model implicitly generates a series of univariate time-series models for each endogenous variable. The dynamic structural model, however, imposes restrictions on the parameters in the reduced-form time-series specification. Dynamic structural models prove most effective in performing policy analysis, albeit subject to the Lucas critique. Time-series models prove most effective at forecasting. That is, in both cases errors creep in whenever the researcher makes a decision about the specification. Clearly, more researcher decisions relate to a dynamic structural model than a univariate time-series model, suggesting that fewer errors enter the time-series model and allowing the model to produce better forecasts.

and non-farm employment. From variance decompositions, they conclude that Nevada proves more isolated from national events than is California. Moreover, much of the movement in Nevada variables gets explained by Nevada's variables. For example, the variance of Nevada's employment helps to explain the variances of both Nevada's taxable sales and gross gaming revenue. And, the variance of Nevada's taxable sales helps to explain the variance of Nevada's employment. For forecasting purposes, they assume that the national variables are exogenous to the model. The forecasts from their model for Nevada's taxable sales and employment prove better than the naïve model while those for gross gaming revenue do not. They suspect that the adoption of the lottery in the middle of their forecast horizon drove the poor forecasts for gross gaming revenue.[3]

Shonkwiler (1992) develops a state-space model using the Kalman filter, which introduces stochastic parameter estimation, to forecast gross taxable gaming revenue in Nevada. The econometric procedure allows the slope of the trend in the data series to vary over time. Under various assumptions within this stochastic trend model, different ARIMA models emerge. The stochastic trend model out-performed the BVAR model on Cargill and Rafiee (1990) in forecasting quarterly gaming revenue out of sample over 1988 to 1989, although both methods tended to overestimate gross gaming revenue.

4. Methodology:

This section describes the various methodologies used to forecast Nevada gross gaming revenue and taxable sales. We first describe the specification of the linear models -- *VAR, VEC, BVAR,* and *BVEC* models. Then, we discuss the specification of the NP and SP models, which could

[3] Cargill and Raffiee (1990) update the Nevada BVAR model to include Nevada personal income as well as to include a dummy variable for the California lottery.

represent linear or nonlinear specifications. Finally, we conclude with a description of the specification of the parametric nonlinear models – STAR and ANN models.

Several additional considerations determine the class of nonlinear parametric models used, since a large number of nonlinear models exist for forecasting. These include univariate threshold autoregressive (TAR), STAR, autoregressive ANN (AR-ANN), Markov switching autoregressive (MSAR) models, and their multivariate extensions. Our study considers only STAR, ANN, and their multivariate extensions MSTAR and MAR-ANN, respectively.

Two important conceptual differences distinguish TAR, STAR, and MSAR models. First, MSAR models use less prior information than TAR and STAR models. That is, we can interpret regime probabilities in MSAR models as a regime function estimated from the data. Second, regime changes are predetermined in TAR and STAR models. That is, we can predict regime changes using the past data, because the switch variable is a lagged (predetermined) and known value of the time series that we predict. On the other hand, regime change in an MSAR mode is governed by an exogenous Markov process. Exogenous regime changes are, thus, difficult to predict in MSAR models due to presence of additional disturbances.

The forecasting performance comparison in Marcellino (2002, 2004), Terasvirta (2006), and Deschamps (2008) generally favor STAR models over the MSAR models. In this study, we do not examine TAR models since they nest within STAR models and their implied instantaneous regime shifts lead to poor forecasting performance around turning points. We choose ANN models since they generally rank on top in many forecasting comparison studies. They also serve as a benchmark to compare the nonlinear models. Moreover, ANN models provide superior ability to approximate arbitrary close to any nonlinear function. ANN model prove successful in forecasting economic time series, such as stock prices, exchange rates and in option prices (Hornik et al., 199; Swanson and White, 1995, 1997).

4.1. VAR, VEC, BVAR, and BVEC Specifications

Following Sims (1980), we write an unrestricted VAR model as follows:

(1) $$y_t = A_0 + A(L)y_t + \varepsilon_t ,^4$$

where *y* equals a ($n \times 1$) vector of variables, which in our case, includes four variables -- gross gaming revenue (GGR), taxable retail sales (TRS), the leading employment index (LI), and the coincident employment index (CI); *A(L)* equals an ($n \times n$) polynomial matrix in the backshift operator *L* with lag length *p,*[5] and ε equals an ($n \times 1$) vector of error terms. In our case, we assume that $\varepsilon \sim N(0, \sigma^2 I_n)$, where I_n equals an ($n \times n$) identity matrix. Note that, we find that all four variables are non-stationary in levels based on Augmented–Dickey–Fuller (ADF), the Dickey-Fuller with GLS detrending (DF-GLS), the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS), and the Phillips-Perron (PP) tests of unit roots. Hence, we estimate the (V)AR models in natural-logarithmic differences of the variables.[6]

With cointegrated (non-stationary) series, we can transform the standard VAR model into a VEC model. The VEC model builds into the specification of the cointegration relations so that they restrict the long-run behavior of the endogenous variables to converge to their long-run, cointegrating relationships, while at the same time describing the short-run dynamic adjustment

---

[4] *A(L)* = $A_1 L + A_2 L^2 + ... + A_p L^p$; and A_0 equals an ($n \times 1$) vector of constant terms.

[5] We use three lags as confirmed by the sequential modified LR test statistic, Akaike information criterion (AIC), and the final prediction error (FPE) criterion, obtained from an estimation of a stable VAR model. Stability, as usual, implies that no roots lie outside the unit circle. These results are available upon request from the authors.

[6] In an earlier version, we estimated the VAR model using the variables in level form. The Associate Editor suggested possible spurious correlation and recommended using the growth rates of the variables. We appreciate this comment, which we implemented in this revision . We recover the forecast in levels using the actual value of the preceding period.

of the system. The cointegration terms, known as the error correction terms, gradually correct through a series of partial short-run adjustments.

More explicitly, for our four variable system, if each series y_t is integrated[7] of order one, (i.e., *I(1)*),[8] then the error-correction counterpart of the VAR model in equation (1) converts into a VEC model as follows.[9]

(2) $$\Delta y_t = \pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-1} + \varepsilon_t \text{, where } \pi = -[I - \sum_{i=1}^{p} A_i] \text{and} \Gamma_i = -\sum_{j=i+1}^{p} A_j .$$

We impose economic priors on the single cointegrating vector for these four variables based on the trace and maximum–Eigen-value test statistics (Johansen, 1995).[10] The estimated cointegrating vector nearly implies a one-to-one relationship between Nevada gross gaming revenue and taxable sales, which we impose in a constrained VEC model. Additionally, we restrict all insignificant coefficients in the long-run vector and the short-run adjustment coefficients to zero to ensure not only a theoretically consistent, but also a parsimonious, model. Note that we could not reject the restrictions at the one-percent level of significance. Thus, we report the results of a restricted VEC (RVEC) as well, besides the unrestricted VEC.

VAR and VEC models typically use equal lag lengths for all variables in the model, which implies that the researcher must estimate many parameters, including many that prove statistically insignificant. This over-parameterization problem can create multicollinearity and a loss of degrees of freedom, leading to inefficient estimates, and possibly large out-of-sample

---

[7] A series is integrated of order *q,* if it requires *q* differences to transform it into a zero-mean, purely non-deterministic stationary process.

[8] See LeSage (1999) and references cited therein for further details regarding the non-stationarity of most macroeconomic time series.

[9] See Dickey *et al.* (1991) and Johansen (1995) for further technical details.

[10] The Associate Editor also suggested the analysis of this paragraph. These results are available upon request from the authors.

forecasting errors. Some researchers exclude lags with statistically insignificant coefficients. Alternatively, researchers use near VAR models, which specify unequal lag lengths for the variables and equations.

Litterman (1981), Doan *et al.,* (1984), Todd (1984), Litterman (1986), and Spencer (1993), use a Bayesian VAR (BVAR) model to overcome the over-parameterization problem. Rather than eliminating lags, the Bayesian method imposes restrictions on the coefficients across different lag lengths, assuming that the coefficients of longer lags may approach more closely to zero than the coefficients on shorter lags. If, however, stronger effects come from longer lags, the data can override this initial restriction. Researchers impose the constraints by specifying normal prior distributions with zero means and small standard deviations for most coefficients, where the standard deviation decreases as the lag length increases. The first own-lag coefficient in each equation is the exception with a unitary mean. Finally, Litterman (1981) imposes a diffuse prior for the constant. We employ this "Minnesota prior" in our analysis, where we implement Bayesian variants of the classical VAR and VEC models.

Formally, the means and variances of the Minnesota prior take the following form:

(3) $$\beta_i \sim N(1, \sigma^2_{\beta_i}) \text{ and } \beta_j \sim N(0, \sigma^2_{\beta_j})$$

where β_i equals the coefficients associated with the lagged dependent variables in each equation of the VAR model (i.e., the first own-lag coefficient), while β_j equals any other coefficient. In sum, the prior specification reduces to a random-walk with drift model for each variable, if we set all variances to zero. The prior variances, $\sigma^2_{\beta_i}$ and $\sigma^2_{\beta_j}$, specify uncertainty about the prior means $\overline{\beta}_i = 1$, and $\overline{\beta}_j = 0$, respectively.

Doan *et al.,* (1984) propose a formula to generate standard deviations that depend on a small numbers of hyper-parameters: *w, d*, and a weighting matrix *f*(*i, j*) to reduce the over-parameterization in the VAR and VEC models. This approach specifies individual prior variances for a large number of coefficients, using only a few hyper-parameters. The specification of the standard deviation of the distribution of the prior imposed on variable *j* in equation *i* at lag *m*, for all *i, j* and *m*, equals $S_1(i, j, m)$, defined as follows:

(4) $$S_1(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j},$$

where $f(i, j) = 1$, if $i = j$ and k_{ij} otherwise, with ($0 \le k_{ij} \le 1$), and $g(m) = m^{-d}$, with $d > 0$. The estimated standard error of the univariate autoregression for variable *i* equals $\hat{\sigma}_i$. The ratio $\hat{\sigma}_i / \hat{\sigma}_j$ scales the variables to account for differences in the units of measurement and, hence, causes specification of the prior without consideration of the magnitudes of the variables. The term *w* indicates the overall tightness and equals the standard deviation on the first own lag, with the prior getting tighter as the value falls. The parameter *g(m)* measures the tightness on lag *m* with respect to lag 1, and equals a harmonic shape with decay factor *d*, which tightens the prior at longer lags. The parameter *f*(*i, j*) equals the tightness of variable *j* in equation *i* relative to variable *i*, and by increasing the interaction (i.e., the value of k_{ij}), we loosen the prior.[11] [12] Sims et al. (1990) indicate that with the Bayesian approach entirely based on the likelihood function, the associated inference does not require nonstationarity, since the likelihood function exhibits the same Gaussian shape regardless of the presence of nonstationarity. Thus, we specify the

[11] For an illustration, see Dua and Ray (1995).

[12] We estimate the (B)VAR and (B)VEC models, as well as the random-walk model, using the Econometrics Toolbox in MATLAB.

variables in the B(V)AR models in the natural-logarithmic-level form, using four lags, since we find that 3 lags prove optimal based on lag-length tests conducted on the first-differenced classical VAR.

4.2. Nonparametric and Semi-Parametric Models

We now proceed with a nonparametric and semi-parametric regression approach for forecasting both gross gaming revenue and taxable retail sales. To ensure stationarity of the variables, we work with the growth rates, and not the actual levels, when fitting the models, and making the forecasts.[13] Eventually after the forecasts of the growth rates are made, the forecasts related to the actual levels are recovered once more.

First, we abbreviate all the variables that are used in these models. The variables used for modeling and initial forecasting, correspond to the growth rates of: GGR, TRS, LI, and CI, but for convenience of understanding, we use the same terminologies to address the growth rates as we do the levels. Thus, we abbreviate the growth rate of GGR as GGR too. We also use GGR1, GGR2, and GGR3 to denote the first, second, and third lags of the growth rate of GGR, respectively.[14] The rest of the cases follow accordingly.

We examine three competing models, and examine their forecasting abilities. These specifications occur as follows:

Model 1: Nonparametric full regression model (NPFR model)

(5) $$GGR = f(GGR_1, GGR_2, GGR_3, TRS_1, TRS_2, TRS_3, LI_1, LI_2, LI_3, CI_1, CI_2, CI_3) + \varepsilon_{GGR}$$

(6) $$TRS = g(GGR_1, GGR_2, GGR_3, TRS_1, TRS_2, TRS_3, LI_1, LI_2, LI_3, CI_1, CI_2, CI_3) + \varepsilon_{TRS}$$

[13] Note that non-parametric and semi-parametric estimation, as well as smooth-transition-autoregressive and artificial-neural-network models, require that the variables are stationary to avoid spurious estimates. Hence, we convert all the variables to their monthly growth rates.

[14] Note that with the semi-parametric, non-parametric, and smooth transition autoregressive models, the artificial–neural-network models and the VECMs based on the growth rates of the variables, use three lags.

Model 2: Nonparametric partial regression model (NPPR model)

(7) $$GGR = f(GGR_1, GGR_2, GGR_3) + \varepsilon_{GGR}$$

(8) $$TRS = g(TRS_1, TRS_2, TRS_3) + \varepsilon_{TRS}$$

Model 3: Semi-parametric full regression model (SPFR model)

(9) $$\begin{aligned} GGR &= \alpha_0 + \alpha_1 GGR_1 + \alpha_2 GGR_2 + \alpha_3 GGR_3 \\ &\quad + f(TRS_1, TRS_2, TRS_3, LI_1, LI_2, LI_3, CI_1, CI_2, CI_3) + \varepsilon_{GGR} \end{aligned}$$

(10) $$\begin{aligned} TRS &= \beta_0 + \beta_1 TRS_1 + \beta_2 TRS_2 + \beta_3 TRS_3 \\ &\quad + g(TRS_1, TRS_2, TRS_3, LI_1, LI_2, LI_3, CI_1, CI_2, CI_3) + \varepsilon_{TRS} \end{aligned}$$

Here, *f(.)* and *g(.)* denote unknown functions that are estimated from the data. The ε_{GGR} and ε_{TRSR} are mean-zero errors, with unchanged variance over the entire data set. The parameters $\alpha_i; \beta_i; i = 1; 2; 3$ are constants estimated from the data. Therefore, the semi-parametric model can also be described as a partially linear nonparametric model.[15]

In the time series context, nonparametric regression could lead to some issues with correlated errors (see, for example, Opsomer *et. al.* 2001). For instance, the data-driven band-width selection techniques in the kernel smoothing methodology could break down in this context. In such cases, we could use a correlation-corrected method called CDPI to yield stable results. In our case, for Models 1, 2, and 3, using 3 lags guarantees the absence of autocorrelation. As a result, the responses in equations (5) to (10) exhibit uncorrelated errors. Also, stationarity checks ensure constant variances in each model. For use of stationary data in nonparametric analysis, also see: Milas and Naraidoo (forthcoming) and Naraidoo and Paya (forthcoming). Finally, we compare such models based on their prediction errors or forecast performances.

---

[15] We use the *np* package in *R* to carry out the regressions outlined above.

We checked the goodness of model fit using Bootstrap testing and found p-values close to 1 for the models used models. When estimating the unknown functions $f(\cdot)$ and $g(\cdot)$ in case of the nonparametric models, a local linear regression using AIC_c bandwidth selection criterion was used. In this case, we also examined all the options for the choice of kernels, and found that the Gaussian kernel of order 2 worked the best yielding highest R-squared values and smallest MSE. The optimum bandwidth chosen by the software was used. In case of the semi-parametric modeling, we first computed data-driven bandwidths of the kernels to be used in the $f(\cdot)$ and $g(\cdot)$ parts of the model since selection of bandwidth for lower levels of tolerance takes a very large proportion of time. We overrode the default tolerances, and set the tolerance levels at 0.1 for the search method as the objective function is well-behaved. The regression type was local constant, and not local linear, as local linear seems to yield smaller R-squared values. Again, for the $f(\cdot)$ and $g(\cdot)$ parts of the model, we used Gaussian kernels of order 2, because they yielded highest R-squared values and lowest MSE.

4.3. Smooth Transition Autoregressive Model Identification

Recent empirical studies show that smooth–transition-autoregressive (STAR) models can successfully model economic time series that move smoothly between two or more regimes, e.g., recession to expansion. When considering the joint dynamic properties of gross gaming revenue, taxable retail sales, the leading index, and the coincident index, it is natural to consider multivariate STAR (MSTAR) models. van Dijk et al. (2002), among many others, discussed MSTAR models. Montgomery, et al. (1998) and Marcellino (2002) report much more favorable forecasting performance for LSTAR forecasts, while the results obtained in Stock and Watson (1999) show that linear models generally dominate nonlinear models in terms of forecasting performance. In spite of specification difficulties, such as the appropriate transition variable, number of regimes, type of transition function, and so on, they prove useful for state dependent multivariate relationships. Recent applications (e.g., Rothman *et al.*, 2001; Psaradakis *et al.*, 2005; Tsay, 1998; De Gooijer and Vidiella-i-Anguera, 2004) find that MSTAR models successfully model nonlinear economic time-series data.

Here, we discuss the specification of MSTAR models, which also follows for the (univariate) STAR models. Define $y_t = (y_{1t,} y_{2t}, \ldots, y_{nt})'$ as $(k \times 1)$ vector time series. In our case, $y = (TRS, GGR, LI, CI)'$, where all variables are in logarithms. We specify the *k*-dimensional MSTAR model as follows:

(11) $$\Delta y_t = (\Theta_{1,0} + \sum_{j=1}^{p} \Theta_{1,j} \Delta y_{t-j}) + (\Theta_{2,0} + \sum_{j=1}^{p} \Theta_{2,j} \Delta y_{t-j}) G(s_t; \gamma, c) + \varepsilon_t,$$

where Δ denotes the first difference operator such that $\Delta x_t = x_t - x_{t-1}$, $\Theta_{i,0}$, $i = 1,2$, are $(k \times 1)$ vectors, $\Theta_{i,j}$, $i = 1,2$, $j = 1,2,\ldots,p$, are $(k \times k)$ matrices, and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})$ is a *k*-dimensional vector of white noise processes with zero mean and nonsingular covariance matrix Σ, $G(\cdot)$ is the transition function that controls smooth moves between the two regimes, and s_t is the transition variable. In both univariate and multivariate cases, we allow the transition variable s_t to equal any lagged component of y_t.

The (M)STAR model in equation (11) defines for two regimes, one associated with $G(s_t; \gamma, c) = 0$ and another associated with $G(s_t; \gamma, c) = 1$. The transition from one regime to the other is smooth and determined by the shape of the $G(\cdot)$ function. In this paper, we consider a logistic transition function

(12) $$G(s_t; \gamma, c) = \frac{1}{1 + \exp\{-\gamma (s_t - c) / \hat{\sigma}_s\}}, \qquad \gamma > 0,$$

where $\hat{\sigma}_s$ is the estimate of the standard deviation of transition variable s_t. The parameter c is the threshold determining the midpoint between two regimes at $G(c;\gamma,c)=0.5$. The speed of transition between the regimes is determined by the parameter γ, with higher values corresponding to faster transition.

To specify both STAR and MSTAR models, we follow the procedure presented in Terasvirta (1998) (see also van Dijk, et al., 2002; Lundbergh and Terasvirta, 2002). The first step specifies the lag order of p =3. We maintain this order in the univariate case as well.

The next step tests linearity against the MSTAR alternative. Since the MSTAR model contains parameters not identified under the alternative, we follow the approach of Luukkonen et al. (1988) and replace the transition function $G(\cdot)$ with a suitable Taylor approximation to overcome the nuisance parameter problem. The testing procedure selects a logistic MSTAR model with a single threshold, which we maintain for the univariate case as well.

The third step in our MSTAR model identification selects the transition variable s_t. In order to identify the appropriate transition variable, we run the linearity tests for several candidates, $s_{1t}, s_{2t}, \ldots, s_{mt}$, and select the one that gives the smallest p-value for the test statistic. Here, we consider lagged monthly changes of all four variables as the candidate transition variable. Let $s_t = \Delta x_{t-d}$, where x is any of the four variables $\{TRS, GGR, LI, CI\}$. We test linearity with theses four variables for delays $d = 1, 2, \ldots, 8$. We obtain the smallest p-value with $x_t = TRS_t$ and $d = 3$. For the univariate case, we follow the same procedure to select s_t. We select TRS_{t-3} for taxable retail sales as the transition variable, and GGR_{t-2} for gross gaming revenue. Analytical point forecasts are not available for non-linear (V)AR models when the disturbance term is Gaussian even when $h \geq 2$, as $E[f(x)] \neq f[E(x)]$, where h is the number of steps-ahead for the forecasts.[16]

4.4. Artificial Neural Network Model Identification

Artificial–neural-network (ANN) models perform well in forecasting nonlinear and chaotic time series (Lachtermacher and Fuller, 1995). As analogues to the STAR models, we consider both multivariate autoregressive ANN (MAR-ANN) and univariate autoregressive ANN (AR-ANN) models. We estimate the univariate models only for forecasting taxable retail sales and gross gaming revenues. Lisi and Schiavo (1999) use an ANN models for predicting European exchange rates, finding that they performed as well as the best model, a chaos model. Using statistical tests, Lisi and Schiavo (1999) discover no significant difference between the ANN and chaos models. Stern (1996) applies ANN models to several simulated data from autoregressive models of order 2, AR(2), with various signal to noise ratios. The results showed that ANN models do not generate good predictions with a small signal to noise ratio. ANN models seem most suitable for forecasting time series with small signal to noise ratios, given sufficient data and appropriate data transformations. Success of ANNs in forecasting nonlinear time series reflects their universal function approximation capability (White, 1988, 1989). This includes any linear or nonlinear function (Cybenko, 1989; Funahashi, 1989; Hornik, et al., 1989; Wasserman, 1989). Because of this approximation capacity, neural networks offer several potential advantages over alternative methods for non-normal and non-linear data (Hansen et al., 1999).

Researchers use a variety of neural-net architectures for time-series prediction. The most widely used architecture for time series prediction is the multilayer perceptron (MLP) (also

---

[16] Details of the bootstrapping procedure are available upon request from the authors. We implement all computations of the STAR models with the RSTAR package (Version 0.1-1) in R developed by the one of the authors of this paper.

known as a feed-forward neural network) (Sarle, 2002). The MLP is capable of resolving a wide variety of problems (Bishop, 1995; Kaastra, Boyd, 1996). In this paper, we also prefer the MLP network for (M)AR-ANN based forecasting. In an MLP network, the units are partitioned into layers. Usually, the MLP network contains an input and an output layer, and one or more hidden layers of neurons between the input and output layers. In the MLP architecture, data are always transmitted from the input layer to the output layer. In our case, each input neuron represents one of the lagged values, while the output neuron(s) represent dependent variable(s) or MLP network forecasts. The MLP is a network with links from each unit in the *k*th layer directed only to units in the (*k* + 1)st layer. In the (M)AR-ANN models, the lags of variables enter as inputs to the first layer, and outputs from the network appear in the last layer. A weight ("connection strength") is associated with each link, and a network is trained ("learned") by modifying these weights, thereby modifying the network function that maps inputs to outputs. We use the (M)AR-ANN model with *q*-hidden layers, which we write as follows:

(13) $$\Delta y_t = \sum_{i=1}^{q} \beta_i' G(\Theta_{i,0} + \sum_{j=1}^{p} \Theta_{i,j} \Delta y_{t-j}) + \varepsilon_t,$$

where Δ and y_t are as before, β_i, $i = 1, 2, \ldots, q$, are parameters called weights or connection strengths, $\Theta_{i,0}$, $i = 1, 2, \ldots, q$, are $(k \times 1)$ vectors, $\Theta_{i,j}$, $i = 1, 2, \ldots, q$, $j = 1, 2, \ldots, p$, are $(k \times k)$ matrices, $G(\cdot)$ is the "squashing (activation) function" called the "hidden unit", and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})$ is a *k*-dimensional vector of white noise processes with zero mean and nonsingular covariance matrix Σ.

In building ANNs for forecasting time series, researchers frequently subdivide the sample into three sets (Bishop, 1995; Ripley, 1996). These sets are called training, validation, and test sets. The training set is used to construct the network, the validation set is used to obtain forecast performance measures, and the test set is used to check for generalization capacity of the network. This method can usefully construct networks with good generalization capability that performs well with new cases. During the network's training stage, the weights iteratively adjust, using an algorithm such as the back propagation of Rumelhart, et al. (1986), on the basis of the training set's values, in order to minimize the error between the network's predicted output and the actual (desired) output. We use sum–of-squared errors (SSE) as a criterion to determine the optimal weights based on the training set. Nevertheless, ANN training based on the training set may lead to overfitting. In order to avoid overfitting, the validation set controls the learning process. We evaluate an ANN's performance by changing the number of hidden layers and type of activation function at hidden and output layers, on the basis of the mean squared error (MSE) obtained when the trained ANN forecasts the period in the validation set. Finally, the test set, which is an independent set of data, provides an unbiased estimate of the generalization error or forecasting performance. No optimal rules exist to select the size of each set of data, although by general agreement, the training set should be the largest. In this paper, we use data from 1982:M2 to 2002:M12 as the training set (251 observations, 74.93%), data from 2003:M1 to 2006:M5 as the validation set (41 observations, 12.24%), and data from 2006:M6 to 2009:M12 as the test set (43 observations, 12.84%). We evaluate a network's performance based on the 1 to 24 step-ahead forecasts in the validation period and we select the best performing network based on the minimum MSE. Then, we select the network that based on the validation set is used to forecast the test period.[17]

Creating an MLP network involves five sets of parameters: the learning rule, network architecture, learning rate and momentum factor, activation function of the hidden and output

[17] See Section 5 for further details.

layers, and number of iterations. Over the years, researchers develop many methods to train an ANN. (see Fine, 1999). MacKay (1992) proposed a Bayesian framework, called the Bayesian regularization, to overcome the problems in interpolation of noisy data. Bayesian regularization facilitates the selection of parsimonious models as well as maximum likelihood estimation. Bayesian regularization advantageously expands the cost function to search not only for the minimal error, but also for the minimal error using the minimal weights. In the Bayesian regularization approach, one determines a set of smaller models nested within a larger model and the algorithm chooses one of these smaller models, providing a method to select parsimonious models. The procedure first assigns prior probabilities to each of the smaller models and then determines the model that posts the highest posterior probability. Following the recommendation in Foresee and Hagan (1997), we fit the models using the Levenberg–Marquardt algorithm.

In this paper, the MLP architecture uses three lags of each variable as inputs for MAR-ANN and three lags of own for AR-ANN models of taxable retail sales and gross gaming revenues. An MLP network's capacity to learn depends on the number of hidden neurons. Despite its significant role, no statistical criteria exist to select the optimum number of hidden neurons. We select the best ANN with Bayesian regularization, bearing in mind the overfitting issue, based on its MSE in the validation set, using the least possible number of hidden neurons (Masters, 1993; Smith, 1993; Rzempoluck, 1998). For both MAR-ANN and AR-ANN models, we try ANNs with maximum q set to 9. We obtain the best performing MAR-ANN with $q = 3$, and the best performing AR-ANN with $q = 2$ for taxable retail sales and with $q = 1$ for gross gaming revenue.

In our study, the input layer neurons use a linear activation function, while the hidden and output layer neurons use a sigmoid activation function, $G(\cdot)$. Two sigmoid functions widely used in MLP are the logistic (providing continuous values between 0 and 1) and hyperbolic tangent sigmoid, called tansig, functions (providing continuous values between -1 and 1). In this study, we use the tansig function in the hidden and output layers of the MLP networks, since it allows much faster learning in comparison to the logistic function (Fahlman, 1988; Fausett, 1994). We scale our data onto -1 and 1, which is the range covered by the tansig function.

The learning rate parameter plays a crucial role in the training process of MLP networks. The learning rate controls the change in the weights in each step of the iteration of training. In order to obtain optimum weights, researchers should avoid both too-small and too-large size changes in weights. We use a learning rate of 0.25, which provides good results in most practical cases (Rumelhart et al., 1986). We can increase the speed of learning by filtering, based on the past changes, the oscillations caused by the learning rate. The momentum factor parameter controls the effect of past changes, which should be a number close to 1. In this study, we use a momentum factor equal to 0.85.[18]

5. Data and Results:

This section reports our data sources and econometric findings. In addition to the monthly Nevada coincident and leading employment indexes, we use monthly data from January 1982 through December 2009 on seasonally-adjusted Nevada gross gaming revenue and taxable retail sales. We use January 1982 through December 2002 as the in-sample period and January 2003 through May 2006 as the out-of-sample period, which is updated recursively to generate one- to twenty four-months-ahead forecasts. Our forecasting analysis compares the various models against a benchmark random-walk model. Finally we use the period from June 2006 to December 2009 for carrying out ex ante forecasts in an attempt to predict the downturn in gross

[18] We implement all computations of the ANN models with the Neural Network Toolbox (Version 6.0) in MATLAB.

gaming revenue (which peaked in November 2006) and taxable retail sales (which peaked in February 2007).

5.1. VAR, VEC, BVAR, and BVEC Forecast Results

We begin with a series of linear forecasting models. Tables 1 and 2 report the results. The best performing models bifurcate across taxable sales and gross gaming revenue. For taxable sales, the VEC models generally provide the best forecasting performance on average across all forecast horizons. Specifically, the VEC model outperforms all other models over steps two to twelve, while, in the medium-run, the RVEC model stands out. The VAR dominates at the longer run horizon. Though the BVEC models do not perform better relative to the VEC, RVEC, or VAR models in any specific steps, they generally rank, on average, second in forecast performance, usually by a small margin. In sum, the VEC and BVEC models provide the best forecasting performance for taxable sales with average performance about 44 percent better than the random walk model benchmark.

For gross gaming revenue, the VAR models generally provide the best forecasting performance in the medium- to longer-run horizons. Averaged across all forecast horizons, however, the RVEC model performs the best, as well as at shorter-run horizons, interchanging often with the AR model. On average, the BVAR model with the most loose prior (w = 0.3, d = 0.5) performs exceptionally well coming second to the RVEC model, producing a RMSE of 0.6814 relative to the random-walk model, which, in turn, is only marginally bigger than 0.6791 – the average relative RMSE from the RVEC model.[19] As the prior gets tighter, however, the

[19] Note in Table 2, where the relative RMSEs from the RVEC model and the BVAR model with w = 0.3 and d = 0.5 equal each other at two decimal places only (i.e., 0.68). Going to four decimal places, the RVEC model produces a lower relative RMSE. We choose this specification as the optimal model for the gross gaming revenue and use it to perform the ex ante forecasts, the discussion of which follows in the next paragraph. Not surprisingly, the BVAR model with w = 0.3 and d = 0.5, produces a similar ex ante forecast path for the gross gaming revenue. The results are available upon request from the authors.

classical VAR model beats all the BVAR models, on average, and outperforms the VAR models at the first three months forecast horizon. Moreover, occasionally, the VEC or BVEC model outperforms the VAR and RVEC models. The VEC and BVEC models, however, show an erratic forecasting performance with excellent performance at some horizons and awful performance at other horizons. The VAR, RVEC, and BVAR models show a consistent pattern across all forecast horizons and the RMSEs differ only marginally for these cases. In sum, the VAR, RVEC, and BVAR models provide the best forecasting performance for taxable sales, where average performance exceeds the random-walk benchmark model by about 32 percent.

Figures 3 and 4 provide out-of-sample ex ante forecasts of taxable sales and gross gaming revenue from June 2006 through December 2009. The linear nature of the optimal VEC and RVEC models used to forecast taxable sales and gross gaming revenue make it difficult to move with the twists and turns in either series over the sample period, which sees significant ups and downs in these components to Nevada's tax base. Figures 5 and 6 present the ex ante forecasts for the random-walk model for comparison purposes.[20]

5.2. Nonparametric and Semi-Parametric Forecast Results

The nonparametric and semi-parametric forecasting models in our application generally perform poorly when compared to other models. On average, all three models perform worse than the random walk model. For example, Model 1 produced an especially bad forecast performance for gross gaming revenue with an average of almost 2000 percent worse than the random-walk model. The performance of these three models improves somewhat for longer forecast horizons,

[20] The Associate Editor suggested that we perform out-of-sample ex ante forecasts for a samples longer and shorter than those used in the current text. We do so for out-of-sample ex ante samples from December 2004 to December 2009 and from March 2008 to December 2009. The Figures for the longer and shorter sample periods of the optimal VEC and VAR models as well as the random walk models show exactly the same pattern of one-direction movement in the forecasts with no signs of trying to match any of the twists and turns in the actual series. All Figures appear in a longer version of this paper on line at: http://ideas.repec.org/p/pre/wpaper/201018.html.

especially for the semi-parametric model. The semi-parametric model outperforms the random-walk model from forecast horizon 11 and 13 through 24 for taxable sales and gross gaming revenue, respectively.

Figures 7 and 8 provide out-of-sample ex ante forecasts of taxable sales and gross gaming revenue from June 2006 through December 2009. The optimal semi-parametric model does a much better job than the optimal VEC and VAR models in the prior subsection in keeping the forecast values near the actual values. This outcome occurs in spite of the fact that the semi-parametric model did not outperform the random walk model.[21]

5.3. Smooth Transition Autoregressive Forecast Results

The smooth transition autoregressive forecasts prove much better at forecasting taxable sales and gross gaming revenue than the prior techniques. Generally, the STAR models outperform the MSTAR models, although the differences in performance are not large. From forecast horizon 4 to 24, the STAR model dominates. For horizons 1 to 3, we flip flop between the STAR and MSTAR model as the best performing model. On average, the STAR model outperforms the random-walk model by about 85 percent.

Figures 9 and 10 provide out-of-sample ex ante forecasts of taxable sales and gross gaming revenue from June 2006 through December 2009. For both variables, the ex ante forecasts quickly evolve into a no-change forecast over most of the out-of-sample forecast period.[22]

[21] For the longer and shorter samples, we also see that the optimal semi-parametric model does mirror movements in the actual series generally with a slight lag. One exception occurs. The forecasts of gross gaming revenue in the longer sample tend to drift apart from the actual series beginning with the Great Recession in 2008.

[22] The forecasts for the longer and shorter samples of the STAR models also show the same pattern of a constant forecast shortly after the initial forecasts.

5.4. Artificial Neural Network Forecast Results

The autoregressive artificial neural network model outperforms (or equals in one case) the multivariate autoregressive neural network model at all forecast horizons. Moreover, as we noted for the smooth transition autoregressive models, the performance of the AR-ANN model beats the random-walk model by a large amount. Comparing the two models, the autoregressive artificial neural network model outperforms the smooth transition autoregressive model by a small amount in 20 and 21 out of 24 forecast horizons for taxable sales and gross gaming revenue, respectively. On average, the AR-ANN model beats the random-walk model by about 85 to 86 percent.

Figures 11 and 12 provide out-of-sample ex ante forecasts of taxable sales and gross gaming revenue from June 2006 through December 2009. These graphs tell a story similar to the ex ante forecasts using the smooth transition autoregressive models reported in Figures 9 and 10.[23]

6. Conclusions:

Most state governments face some constitutional or legislative requirement to balance their current services budget. Nevada is no exception. Thus, states necessarily need to forecast revenue in order to determine the level of government spending that the forecast revenue can support. In Nevada, the Economic Forum, a group of five laypersons, makes revenue forecasts that bind the government to spending limits. The Economic Forum hears testimony from various constituencies, including the legislative and executive branches of government and attempts to craft a consensus forecast. Taxable sales and gross gaming revenue comprises a significant portion of Nevada's tax base.

[23] Once again, the forecasts for the longer and shorter samples of the AR-ANN models show the same pattern of a constant forecast shortly after the initial forecasts.

Table 6 summarizes much of our findings. This Table reports the forecast performance of the autoregressive artificial–neural-network models versus the random-walk model and then between the autoregressive artificial-neural-network models and, in turn, the semi-parametric models, the unrestricted and restricted vector-error-correction models for taxable sales and gaming revenue, respectively, and the smooth-transition autoregressive models. That is, the autoregressive artificial neural network and smooth transition autoregressive models provide the best performance of all of the models by a significant margin, but we do not see statistical differences between the forecast errors of artificial-neural-network and smooth-transition autoregressive models, based on the *ENC-T* test of Clark and McCracken (2001).[24]

References

Banerji, A., Dua, P., and Miller, S. M., 2006. Performance evaluation of the new Connecticut leading employment index using lead profiles and BVAR models. *Journal of Forecasting* 25, 415-437.

Bhardwaj, G., and Swanson, N. R., 2006. An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. *Journal of Econometrics* 131(1-2), 539-578.

Bishop, C. M., 1995. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Cargill, T. F., and Eadington, W. R., 1978. Nevada's gaming revenue: Time characteristics and forecasting. *Management Science* 24, 1221-1230.

Cargill, T. F., and Morus, S. A., 1988. A vector autoregressive model of the Nevada economy. Federal Reserve Bank of San Francisco *Economic Review* Winter, 21-32

Cargill, T. F., and Raffie, K., 1990. The Nevada VAR model: Update and general observations. *Nevada Review of Business and Economics* 14, 2-9.

Clark, T. E., and McCracken, M. W., 2001. Tests of Equal Forecast Accuracy and Encompassing for Nested Models. *Journal of Econometrics* 105(1), 85–110.

[24] See Appendix for more explanation.

Cybenko, G., 1989. Approximation by superposition of sigmoidal functions. *Mathematics of Control, Signals and Systems* 2, 303-314.

De Gooijer, J., G., and Vidiella-i-Anguera, A., 2004. Forecasting threshold cointegrated systems. *International Journal of Forecasting* 20, 237-253.

Deschamps, P. J., 2008. Comparing smooth transition and Markov switching autoregressive models of US unemployment. *Journal of Applied Econometrics* 23(4), 435-462.

Dickey, D. A., Jansen, D. W., and Thornton, D. L., 1991. A primer on cointegration with an application to money and income. Federal Reserve Bank of St. Louis *Review* 73, 58-78.

Doan, T. A., Litterman, R. B., and Sims, C. A., 1984. Forecasting and Conditional Projections Using Realistic Prior Distributions. *Econometric Reviews* 3, 1-100.

Dua, P., and Miller, S. M., 1996. Forecasting Connecticut Home Sales in a BVAR Framework Using Coincident and Leading Indexes. *Journal of Real Estate Finance and Economics* 13, 219-235.

Eisendrath, D., Bernhaerd, B. J., Lucas, A. F., and Murphy, D. J., 2008. An analysis of the effects of September 11, 2001, on Las Vegas strip gaming volume. *Cornell Hospitality Quarterly* 49, 145-162.

Fahlman, S., E., 1988. Faster-learning variations on back-propagation: An empirical study. In Touretsky, D., Hinton, G. E., Sejnowski, T. J. (Eds.), *Proceedings of the 1988 Connectionist Models Summer School*. San Mateo: Morgan Kaufmann, 38–51.

Fausett, L., 1994. *Fundamentals of Neural Networks*. NJ: Prentice-Hall.

Fine, T., L., 1999. *Feedforward Neural Network Methodology*. Berlin: Springer-Verlag.

Foresee, F., D., and Hagan, M., T., 1997. Gauss–Newton approximation to Bayesian regularization. *IEEE International Conference on Neural Networks*, vol. 3 New York: IEEE, 1930-1935.

Funahashi, K., 1989. On the approximate realization of continuous mappings by neural networks. *Neural Networks* 2, 183–192.

Hansen, J., V., McDonald, J., B., and Nelson, R., D., 1999. Time series prediction with genetic-algorithm designed neural networks: An empirical comparison with modern statistical models. *Computational Intelligence* 15, 171-184.

Hayfield, T., and Racine, J. S., 2008. Nonparametric Econometrics: The np Package. *Journal of Statistical Software* 27, 1-32.

Hornik, K., Stinchombe, M., and White, H., 1989. Multilayer feed-forward networks are universal approximators. *Neural Networks* 2, 359–366.

Johansen, S., 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Kaastra, I., and Boyd, M., 1996. Designing a neural network for forecasting financial and economic time series. *Neurocomputing* 10, 215–236.

Lachtermacher, G., and Fuller, J. D., 1995. Backpropagation in time-series forecasting. *Journal of Forecasting* 14: 381–393.

LeSage, J. P., 1999. *Applied Econometrics Using MATLAB*, www.spatial-econometrics.com.

Lisi, F., and Schiavo R. A., 1999. A comparison between neural networks and chaotic models for exchange rate prediction. *Computational Statistics and Data Analysis* 30, 87–102.

Litterman, R. B., 1981. A Bayesian procedure for forecasting with vector autoregressions. *Working Paper*, Federal Reserve Bank of Minneapolis.

Litterman, R. B., 1986. Forecasting with Bayesian vector autoregressions – Five Years of Experience. *Journal of Business and Economic Statistics*, 4(1), 25-38.

Lundbergh, S., and Terasvirta, T., 2002. Forecasting with smooth transition autoregressive models. In Clements, M. P., and Hendry, D. F., (Eds.), *A Companion to Economic Forecasting* Oxford: Blackwell, 485–509.

Luukkonen, R., Saikkonen, P., and Terasvirta, T., 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* 75, 491–499.

MacKay, D. J. C., 1992. A practical Bayesian framework for backpropagation networks. *Neural Computation* 4, 448–472.

Marcellino, M., 2002. Instability and non-linearity in the EMU. Discussion Paper No. 3312, Centre for Economic Policy Research.

Marcellino, M., 2004. Forecasting EMU macroeconomic variables. *International Journal of Forecasting* 20, 359-372.

Masters, T., 1993. *Practical neural networks recipes in C++*. London: Academic Press.

Milas, C., and Naraidoo, R., forthcoming. Financial Market Conditions, Real Time, Nonlinearity and European Central Bank Monetary Policy: In-Sample and Out-of-Sample Assessment. *Computational Statistics and Data Analysis*.

Montgomery, A., Zarnowitz, V., Tsay, R., and Tiao, G., 1998. Forecasting the U.S. unemployment rate. *Journal of the American Statistical Association* 93, 478–493.

Newey, W. K., and West, K. D. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703–08.

Naraidoo, R., and Paya, I., forthcoming. Forecasting monetary rules in South Africa. *International Journal of Forecasting*.

Opsomer, J., Wang, Y., and Yang, Y., 2001. Nonparametric regression with correlated errors. *Statistical Science* 16, 134-153.

Psaradakis, Z., Ravn, M. O., and Sola, M., 2005. Markov switching causality and the money—output relationship. *Journal of Applied Econometrics* 20, 665-683.

Rzempoluck, E. J., 1998. *Neural Network Data Analysis using Simulnet*. New York: Springer.

Ripley, B. D., 1996. *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.

Rothman, P., 1998. Forecasting asymmetric unemployment rates. *Review of Economics and Statistics* 80, 164-168.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986. Learning internal representations by error propagation. In Rumelhart, D. E., and McClelland, J. L., (Eds.), *Parallel Distributed Processing*. Cambridge, MA: MIT Press, 318–362.

Sarle, W. S., 2002. Neural network FAQ. Retrieved 28 May 2010, from ftp://ftp.sas.com/pub/neural/FAQ.html.

Shonkwiler, J. S., 1992. A structural time series model of Nevada gross taxable gaming revenue. *Review of Regional Studies* 22, 239-249.

Sims, C. A., 1980. Macroeconomics and Reality. *Econometrica*, 48(1), 1-48.

Sims, C. A., Stock, J. H., and Watson, M. W., 1990. Inference in Linear Time Series Models with Some Unit Roots. *Econometrica* 58, 113-144.

Smith, M., 1993. *Neural Networks for Statistical Modeling*. New York: Van Nostrand Reinhold.

Spencer, D. E., 1993. Developing a Bayesian vector autoregression model. *International Journal of Forecasting*, 9(3), 407-421.

Stern, H. S., 1996. Neural networks in applied statistics. *Technometrics* 38: 205–220.

Stock, J. H., and Watson, M. W., 1999. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In Engle, R. F., and White, H., (Eds.), *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W. J. Granger.* Oxford: Oxford University Press, 1–44.

Tsay, R. S., 1998. Testing and modeling multivariate threshold models. *Journal of the American Statistical Association* 93, 1188-1202.

Terasvirta, T., 1998. Modelling economic relationships with smooth transition regressions. In Ullah, A., and Giles, D. E. A., (Eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, 507-552.

Teraasvirta, T., 2006. Forecasting economic variables with nonlinear models, *Handbook of Economic Forecasting* Volume 1, 2006, 413-457.

Terasvirta, T., van Dijk, D., and Medeiros, M. C., 2005. "Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination," *International Journal of Forecasting* 21(4), 755-774.

van Dijk, D., Terasvirta, T., and Franses, P. H., 2002. Smooth transition autoregressive models - A survey of recent developments. *Econometric Reviews* 21, 1-47.

Wasserman, P. D., 1989. *Neural Computing: Theory and Practice*. New York: Van Nostrand Reinhold.

White, H., 1988. Economic prediction using neural networks: The case of IBM stock prices. Proceedings of the Second Annual IEEE Conference on Neural Networks, II:451–458. New York: IEEE Press. 17

White, H., 1989. Some asymptotic results for learning in single hidden-layer feedforward network models. *Journal of the American Statistical Association* 84, 1003–1013.

Zellner, A., and Palm, F. 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics* 2, 17-54.

Table 1: VAR, BVAR, VEC, and BVEC Forecast Results: Taxable Sales

FH					w=0.3, d=0.5			w=0.2,d=1			w=0.1,d=1			w=0.2, d=2			w=0.1,d=2		
	AR	VAR	VEC	RVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC
1	**0.26**	7.10	0.77	2.16	0.91	1.29	0.77	0.92	1.34	0.77	0.94	1.33	0.77	0.95	1.48	0.77	0.98	1.43	0.77
2	0.56	1.95	**0.48**	1.78	0.88	1.35	0.48	0.89	1.49	0.48	0.91	1.58	0.50	0.92	1.77	0.50	0.96	1.78	0.52
3	0.42	2.56	**0.33**	1.70	0.90	1.50	0.33	0.91	1.72	0.33	0.92	1.89	0.33	0.93	2.13	0.34	0.95	2.19	0.36
4	0.49	3.15	**0.41**	1.33	0.84	1.45	0.41	0.85	1.66	0.41	0.88	1.84	0.41	0.90	2.04	0.42	0.93	2.12	0.44
5	0.54	2.40	**0.50**	1.13	0.83	1.49	0.52	0.84	1.73	0.53	0.86	1.93	0.53	0.88	2.13	0.54	0.91	2.23	0.56
6	0.48	1.81	**0.40**	0.95	0.82	1.50	0.43	0.83	1.75	0.43	0.85	1.96	0.44	0.87	2.14	0.44	0.90	2.24	0.47
7	1.02	1.89	**0.53**	0.84	0.81	1.44	0.57	0.82	1.67	0.57	0.84	1.86	0.58	0.86	2.01	0.58	0.90	2.10	0.61
8	1.03	1.32	**0.57**	0.85	0.81	1.44	0.61	0.82	1.67	0.61	0.84	1.85	0.61	0.86	1.99	0.62	0.90	2.07	0.64
9	0.81	0.99	**0.49**	0.78	0.81	1.39	0.56	0.82	1.61	0.56	0.84	1.78	0.56	0.86	1.90	0.57	0.90	1.97	0.60
10	1.32	1.11	**0.55**	0.69	0.80	1.35	0.59	0.81	1.55	0.59	0.84	1.71	0.60	0.86	1.81	0.60	0.90	1.87	0.63
11	1.01	0.85	**0.47**	0.63	0.79	1.33	0.55	0.81	1.53	0.56	0.83	1.68	0.56	0.86	1.77	0.57	0.90	1.83	0.60
12	1.08	0.97	**0.48**	0.52	0.78	1.30	0.53	0.80	1.49	0.53	0.83	1.63	0.54	0.85	1.71	0.54	0.90	1.76	0.58
13	1.00	0.87	0.58	**0.44**	0.78	1.29	0.66	0.79	1.47	0.67	0.83	1.60	0.67	0.85	1.67	0.67	0.89	1.72	0.70
14	1.63	0.53	0.58	**0.44**	0.78	1.29	0.61	0.80	1.47	0.62	0.82	1.59	0.62	0.85	1.66	0.63	0.89	1.70	0.65
15	2.29	0.48	0.61	**0.41**	0.78	1.26	0.70	0.79	1.43	0.70	0.83	1.54	0.71	0.85	1.60	0.71	0.89	1.63	0.73
16	2.43	**0.34**	0.63	0.40	0.78	1.24	0.65	0.79	1.40	0.65	0.83	1.51	0.65	0.85	1.56	0.66	0.89	1.59	0.68
17	2.31	**0.25**	0.57	0.40	0.78	1.24	0.69	0.79	1.40	0.69	0.82	1.50	0.70	0.85	1.55	0.70	0.89	1.57	0.73
18	2.09	0.41	0.61	**0.38**	0.77	1.22	0.60	0.79	1.37	0.60	0.82	1.46	0.61	0.85	1.50	0.61	0.89	1.53	0.63
19	2.56	**0.28**	0.61	0.33	0.77	1.21	0.75	0.79	1.35	0.75	0.82	1.44	0.76	0.85	1.48	0.76	0.89	1.50	0.78
20	1.07	**0.25**	0.60	0.29	0.77	1.21	0.54	0.79	1.35	0.54	0.82	1.43	0.55	0.85	1.46	0.55	0.89	1.48	0.58
21	3.00	**0.28**	0.64	0.29	0.77	1.20	0.82	0.79	1.32	0.82	0.82	1.40	0.83	0.85	1.43	0.82	0.89	1.44	0.84
22	3.41	**0.12**	0.63	0.30	0.77	1.19	0.51	0.79	1.32	0.52	0.82	1.39	0.52	0.85	1.42	0.53	0.89	1.43	0.56
23	3.33	**0.14**	0.63	0.29	0.77	1.19	0.89	0.79	1.31	0.90	0.82	1.37	0.90	0.85	1.40	0.90	0.89	1.41	0.91
24	3.29	**0.13**	0.65	0.30	0.76	1.18	0.43	0.78	1.29	0.44	0.82	1.35	0.44	0.85	1.37	0.45	0.90	1.38	0.49
Mean	1.56	1.26	**0.56**	0.73	0.80	1.31	0.59	0.82	1.49	0.59	0.85	1.61	0.60	0.87	1.71	0.60	0.91	1.75	0.63

Note: Numbers represent the ratio of the root mean squared error (RMSE) of the model relative to the RMSE of the random-walk (RW) model. Thus, a ratio less (more) than one implies that the model in question performs better (worse) than the RW model in forecasting at a particular forecast horizon. FH means forecast horizon. AR, VAR, and VEC mean autoregressive, vector autoregressive, and vector error-correction models. BVAR, UBVAR, and BVEC mean Bayesian VAR, univariate BVAR, and Bayesian VEC models. Shaded cells with bold numbers represent the best performing model at forecasting at each forecast horizon. Ave calculates the averages of the 1 to 24 month forecast horizons for each model. For the BVAR and BVEC models, *w* and *d* represent the tightness values chosen for the hyper parameters in the Bayesian specification of the standard errors of the priors on the parameters in the Minnesota prior. We adopt the standard specification for the weighting matrix.

Table 2: VAR, BVAR, VEC, and BVEC Forecast Results: Gross Gaming Revenue

FH					w=0.3, d=0.5			w=0.2,d=1			w=0.1,d=1			w=0.2, d=2			w=0.1,d=2		
	AR	VAR	VEC	RVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC	UBVAR	BVAR	BVEC
1	**0.35**	1.23	0.81	0.41	0.81	0.79	0.81	0.82	0.80	0.81	0.87	0.83	0.81	0.89	0.83	0.81	0.96	0.86	0.81
2	0.78	0.93	1.40	**0.47**	0.82	0.79	1.40	0.83	0.78	1.40	0.87	0.80	1.40	0.89	0.81	1.39	0.96	0.83	1.37
3	0.71	1.12	2.07	**0.54**	0.88	0.84	2.07	0.89	0.84	2.07	0.92	0.84	2.05	0.93	0.86	2.04	0.99	0.85	1.98
4	**0.53**	0.99	2.39	0.57	0.83	0.75	2.39	0.85	0.78	2.39	0.90	0.81	2.37	0.92	0.82	2.36	0.99	0.82	2.29
5	0.62	1.21	2.17	**0.60**	0.82	0.72	2.23	0.84	0.75	2.23	0.90	0.79	2.21	0.93	0.81	2.19	1.01	0.82	2.14
6	0.75	1.38	2.17	**0.66**	0.83	0.71	2.34	0.85	0.76	2.31	0.90	0.80	2.29	0.94	0.83	2.26	1.02	0.84	2.11
7	**0.63**	0.76	16.14	0.74	0.86	0.74	17.86	0.88	0.78	17.86	0.92	0.82	17.71	0.96	0.86	17.43	1.04	0.86	16.57
8	0.91	0.79	4.12	**0.72**	0.89	0.75	4.39	0.90	0.79	4.39	0.94	0.82	4.34	0.98	0.85	4.32	1.05	0.85	4.17
9	1.22	0.96	5.70	**0.70**	0.87	0.70	6.70	0.89	0.74	6.70	0.94	0.77	6.60	0.98	0.79	6.50	1.06	0.78	6.15
10	1.07	**0.61**	6.60	0.71	0.85	0.66	7.20	0.87	0.70	7.20	0.93	0.73	7.12	0.97	0.75	7.08	1.06	0.74	6.80
11	1.33	0.85	6.80	0.72	0.85	**0.67**	7.88	0.88	0.72	7.88	0.93	0.75	7.80	0.98	0.77	7.72	1.07	0.76	7.40
12	1.12	0.82	4.47	0.73	0.84	**0.63**	4.86	0.87	0.69	4.86	0.93	0.73	4.81	0.98	0.75	4.79	1.08	0.73	4.63
13	1.25	0.89	5.42	0.71	0.83	**0.61**	6.88	0.86	0.67	6.83	0.93	0.70	6.75	0.97	0.72	6.67	1.08	0.71	6.29
14	1.96	0.57	**0.51**	0.73	0.84	0.66	0.66	0.87	0.72	0.65	0.94	0.74	0.63	0.98	0.77	0.63	1.09	0.75	0.56
15	2.15	**0.41**	1.06	0.73	0.83	0.62	1.64	0.87	0.68	1.63	0.94	0.71	1.60	0.99	0.73	1.56	1.09	0.70	1.45
16	2.94	**0.39**	1.69	0.73	0.83	0.61	1.84	0.86	0.67	1.82	0.93	0.70	1.80	0.98	0.71	1.79	1.09	0.70	1.67
17	3.83	**0.42**	3.31	0.75	0.84	0.64	4.79	0.88	0.70	4.77	0.94	0.72	4.72	0.99	0.74	4.64	1.10	0.72	4.36
18	2.46	**0.46**	1.46	0.73	0.83	0.62	1.39	0.87	0.67	1.39	0.94	0.70	1.36	0.99	0.71	1.36	1.10	0.69	1.27
19	3.42	**0.39**	3.29	0.73	0.82	0.62	4.31	0.86	0.68	4.30	0.93	0.70	4.26	0.99	0.71	4.21	1.10	0.69	4.03
20	3.92	0.93	1.36	0.73	0.83	**0.64**	0.96	0.87	0.69	0.96	0.93	0.71	0.95	0.99	0.72	0.95	1.10	0.70	0.91
21	3.48	**0.34**	0.87	0.73	0.83	0.64	2.04	0.87	0.69	2.03	0.94	0.71	1.99	0.99	0.72	1.93	1.10	0.70	1.76
22	8.19	0.29	1.02	0.71	0.82	0.63	**0.21**	0.86	0.69	**0.21**	0.93	0.70	0.21	0.99	0.71	0.24	1.10	0.69	0.25
23	6.80	**0.31**	1.12	0.74	0.82	0.66	3.06	0.86	0.71	3.04	0.93	0.73	2.99	0.98	0.73	2.89	1.10	0.71	2.62
24	7.31	0.31	3.97	0.73	0.83	0.67	**0.05**	0.87	0.71	0.08	0.94	0.72	0.10	0.99	0.73	0.28	1.10	0.70	0.51
Mean	2.41	0.72	3.33	**0.68**	0.84	0.68	3.66	0.87	0.73	3.66	0.92	0.75	3.62	0.97	0.77	3.58	1.06	0.76	3.42

Note: See Table 1.

Table 3: Non-Parametric and Semi-Parametric Forecast Results

Forecast Horizon	Model 1: Non-Parametric		Model 2: Non-Parametric		Model 3: Semi-Parametric	
	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue
1	4.555	2.100	4.429	**1.492**	**4.426**	1.773
2	3.652	2.276	3.418	**1.682**	**3.353**	1.710
3	3.578	2.094	3.053	**1.956**	**2.841**	2.156
4	2.883	1.931	2.569	1.723	**2.420**	**1.619**
5	2.851	2.229	2.509	1.774	**2.216**	**1.551**
6	2.689	2.500	2.316	1.822	**2.018**	**1.657**
7	2.369	2.796	2.065	1.871	**1.798**	**1.536**
8	2.219	2.743	1.879	1.745	**1.568**	**1.476**
9	1.962	2.350	1.663	1.382	**1.360**	**1.189**
10	1.802	2.201	1.527	1.307	**1.160**	**1.023**
11	1.649	2.289	1.418	1.344	**0.963**	**1.051**
12	1.484	2.333	1.363	1.283	**0.898**	**1.070**
13	1.370	2.011	1.362	1.235	**0.848**	**0.941**
14	1.203	2.239	1.324	1.277	**0.854**	**0.951**
15	1.061	51.252	1.242	1.186	**0.818**	**0.937**
16	0.967	49.756	1.170	1.135	**0.686**	**0.803**
17	0.931	49.582	1.143	1.094	**0.642**	**0.674**
18	0.892	46.818	1.116	1.042	**0.617**	**0.678**
19	0.838	46.377	1.091	1.080	**0.574**	**0.738**
20	0.777	45.326	1.066	1.088	**0.605**	**0.692**
21	0.724	44.029	1.041	1.008	**0.594**	**0.674**
22	0.668	42.788	1.030	0.996	**0.518**	**0.550**
23	0.581	43.724	1.011	1.046	**0.500**	**0.658**
24	0.542	42.454	0.985	1.001	**0.515**	**0.597**
Average	1.760	20.592	1.741	1.357	**1.366**	**1.113**

Note: See Table 1.

Table 4: Smooth Transition Autoregressive Forecast Results

Forecast Horizon	STAR		MSTAR	
	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue
1	**0.102**	0.130	0.108	**0.128**
2	**0.170**	**0.234**	**0.170**	0.240
3	0.194	**0.236**	**0.183**	0.323
4	**0.128**	**0.198**	0.143	0.271
5	**0.148**	**0.212**	0.185	0.317
6	**0.165**	**0.215**	0.179	0.367
7	**0.135**	**0.208**	0.167	0.407
8	**0.143**	**0.180**	0.197	0.381
9	**0.135**	**0.179**	0.160	0.397
10	**0.127**	**0.152**	0.171	0.366
11	**0.134**	**0.151**	0.170	0.395
12	**0.135**	**0.157**	0.151	0.410
13	**0.134**	**0.131**	0.182	0.369
14	**0.136**	**0.147**	0.175	0.433
15	**0.133**	**0.124**	0.166	0.418
16	**0.140**	**0.123**	0.188	0.390
17	**0.143**	**0.144**	0.169	0.453
18	**0.141**	**0.114**	0.160	0.418
19	**0.144**	**0.114**	0.191	0.414
20	**0.145**	**0.128**	0.182	0.448
21	**0.145**	**0.112**	0.175	0.424
22	**0.148**	**0.099**	0.186	0.419
23	**0.149**	**0.130**	0.180	0.479
24	**0.149**	**0.106**	0.175	0.392
Average	**0.143**	**0.155**	0.171	0.377

Note: See Table 1. STAR and MSTAR mean smooth transition autoregressive and multivariate STAR models.

Table 5: Artificial Neural Network Forecast Results

Forecast Horizon	AR-ANN		MAR-ANN	
	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue
1	**0.127**	**0.128**	0.130	**0.128**
2	**0.165**	**0.237**	0.181	0.240
3	**0.190**	**0.236**	0.209	0.323
4	**0.130**	**0.192**	0.153	0.271
5	**0.146**	**0.208**	0.190	0.317
6	**0.163**	**0.214**	0.185	0.367
7	**0.135**	**0.203**	0.168	0.406
8	**0.142**	**0.175**	0.187	0.381
9	**0.135**	**0.178**	0.166	0.397
10	**0.127**	**0.152**	0.181	0.366
11	**0.134**	**0.148**	0.199	0.394
12	**0.135**	**0.154**	0.198	0.410
13	**0.133**	**0.126**	0.230	0.369
14	**0.135**	**0.143**	0.223	0.433
15	**0.132**	**0.120**	0.196	0.417
16	**0.138**	**0.119**	0.266	0.389
17	**0.142**	**0.139**	0.219	0.453
18	**0.139**	**0.109**	0.219	0.417
19	**0.142**	**0.110**	0.255	0.413
20	**0.143**	**0.124**	0.220	0.448
21	**0.142**	**0.110**	0.272	0.424
22	**0.145**	**0.097**	0.311	0.418
23	**0.147**	**0.128**	0.205	0.479
24	**0.146**	**0.104**	0.224	0.391
Average	**0.142**	**0.152**	0.208	0.377

Note: See Table 1. AR-ANN and MAR-ANN mean autoregressive artificial neural network and multivariate AR-ANNN models.

Table 6: Comparison AR-ANN Performance to Other Models

Forecast Horizon	AR-ANN		AR-ANN/SP		AR-ANN/VEC	AR-ANN/RVEC	AR-ANN/STAR	
	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue	Taxable Sales	Gaming Revenue
1	0.127***	0.128***	0.029***	0.072***	0.166***	0.313***	1.245*	0.985
2	0.165***	0.237***	0.049***	0.139***	0.342***	0.505***	0.971	1.013
3	0.190***	0.236***	0.067***	0.109***	0.578**	0.436***	0.979	1.000
4	0.130***	0.192***	0.054***	0.119***	0.317***	0.337***	1.016	0.970
5	0.146***	0.208***	0.066***	0.134***	0.290***	0.347***	0.986	0.981
6	0.163**	0.214***	0.081***	0.129***	0.410***	0.325***	0.988	0.995
7	0.135***	0.203***	0.075***	0.132***	0.255***	0.275***	1.000	0.976
8	0.142***	0.175***	0.090***	0.119***	0.249***	0.243***	0.993	0.972
9	0.135***	0.178***	0.099***	0.150***	0.274***	0.255***	1.000	0.994
10	0.127***	0.152***	0.110***	0.149***	0.233***	0.214***	1.000	1.000
11	0.134***	0.148***	0.139***	0.141***	0.286***	0.206***	1.000	0.980
12	0.135***	0.154***	0.150***	0.144***	0.282***	0.211***	1.000	0.981
13	0.133***	0.126***	0.157***	0.134***	0.228***	0.178***	0.993	0.962
14	0.135***	0.143***	0.158***	0.150***	0.232***	0.196***	0.993	0.973
15	0.132***	0.120***	0.161***	0.128***	0.214***	0.164***	0.992	0.968
16	0.138***	0.119***	0.201***	0.148***	0.218***	0.162***	0.986	0.967
17	0.142***	0.139***	0.221***	0.207***	0.246***	0.186***	0.993	0.965
18	0.139***	0.109***	0.225***	0.161***	0.230***	0.150***	0.986	0.956
19	0.142***	0.110***	0.248***	0.149***	0.233***	0.151***	0.986	0.965
20	0.143***	0.124***	0.236***	0.180***	0.240***	0.170***	0.986	0.969
21	0.142***	0.110***	0.240***	0.163***	0.221***	0.151***	0.979	0.982
22	0.145***	0.097***	0.280***	0.175***	0.230***	0.136***	0.980	0.980
23	0.147***	0.128***	0.293***	0.195***	0.233***	0.173***	0.987	0.985
24	0.146***	0.104***	0.284***	0.175***	0.225***	0.143***	0.980	0.981
Average	0.142	0.152	0.104	0.137	0.256	0.224	1.001	0.979

Note: See Table 1. The columns headed by AR-ANN measures the RMSE for the AR-ANN relative to the RW model. The columns headed by AR-ANN/SP measures the RMSE for the AR-ANN model to the SP models in Table 3. The column headed by AR-ANN/VEC measures the RMSE for the AR-ANN model to the VEC model in Table 1. The column headed by AR-ANN/RVEC measures the RMSE for the AR-ANN model relative to the RVEC model in Table 2. The column headed by AR-ANN/STAR measures the RMSE for the AR-ANN model relative to the STAR model in Table 4. Finally, *(**)[***] indicates 10%, (5%), [1%] level of significance for the *ENC-T* test of Clark and McCracken (2001).

Figure 1: CBER-DETR Nevada Coincident Employment Index

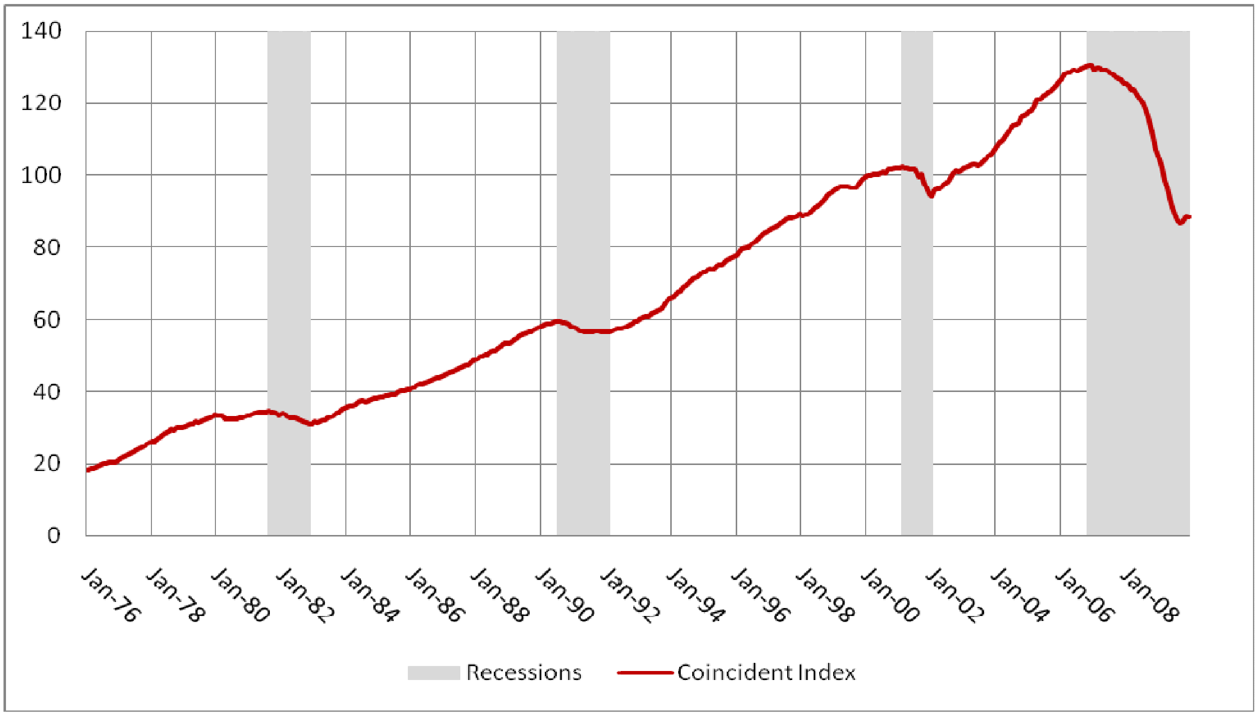


Figure 2: CBER-DETR Nevada Leading Employment Index

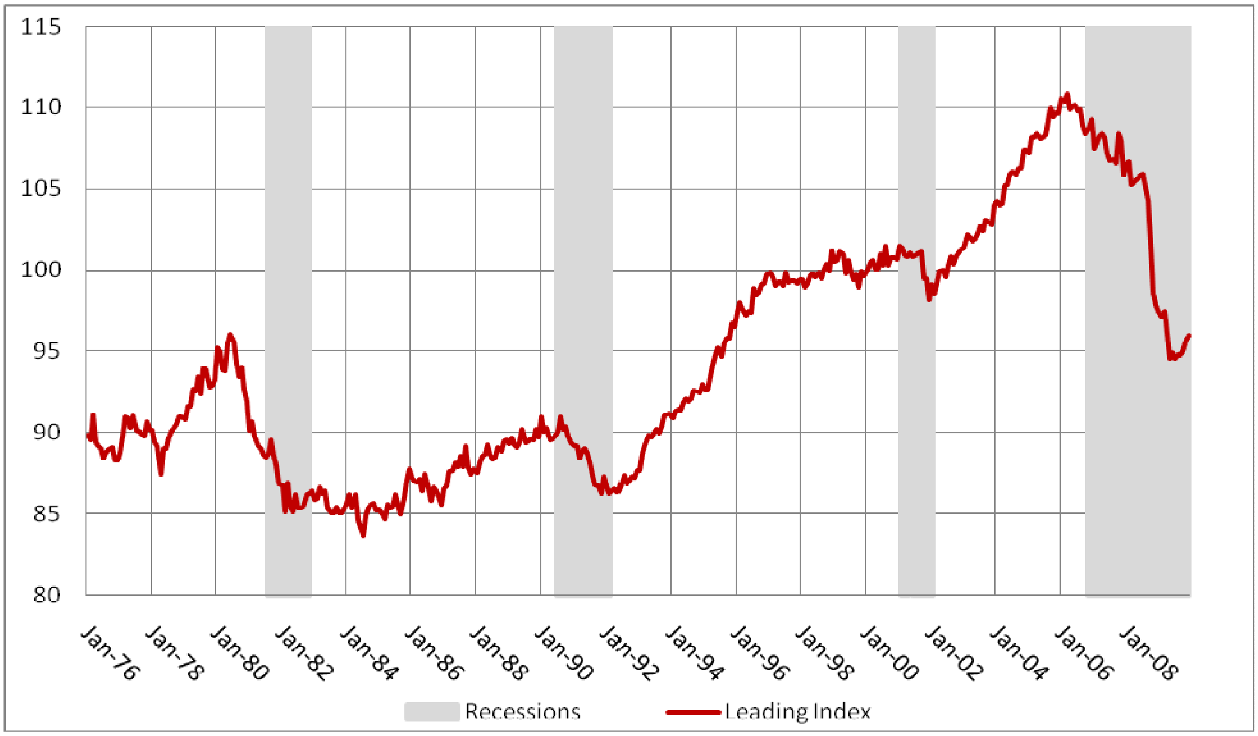


Figure 3: Forecast and Actual Taxable Sales: Optimal VEC Model

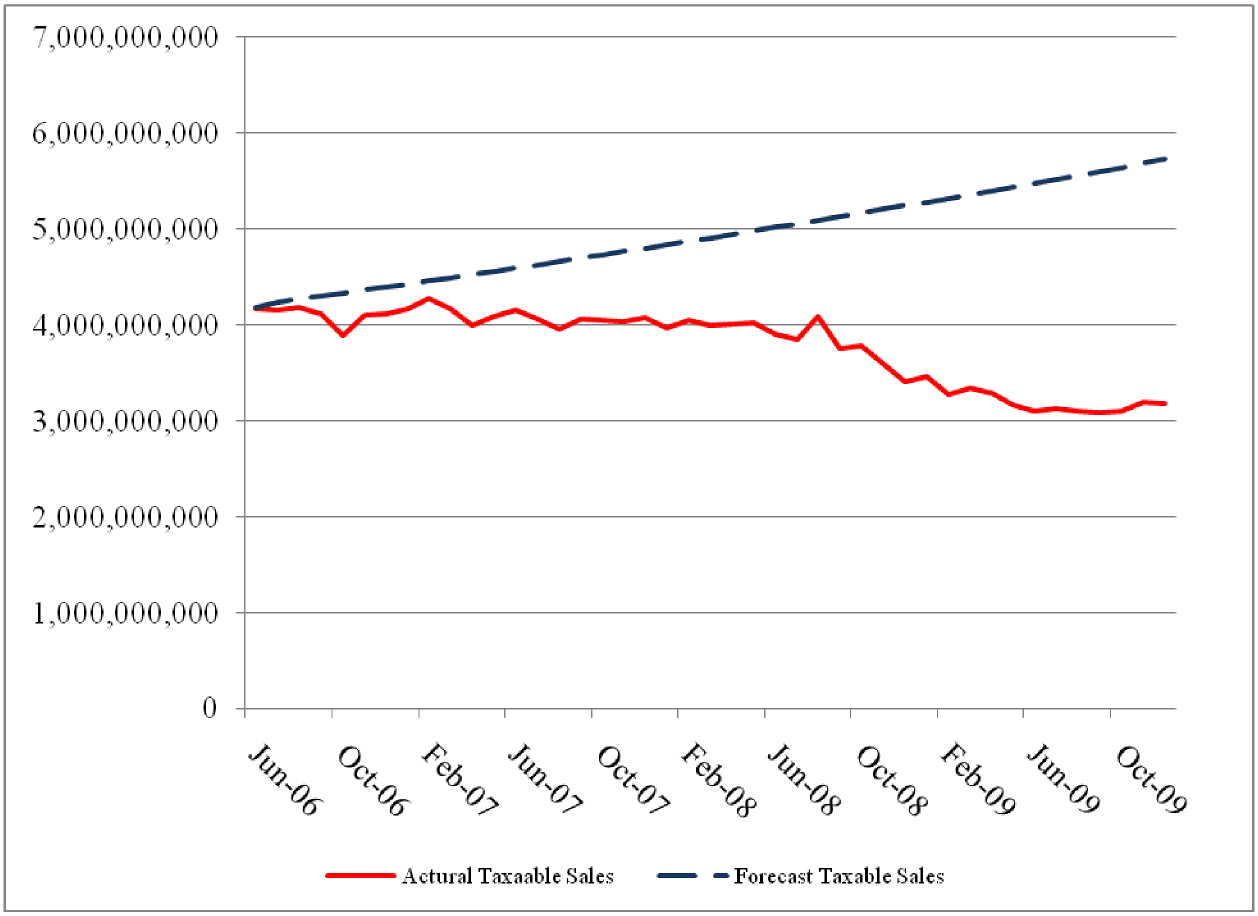


Figure 4: Forecast and Actual Gross Gaming Revenue: Optimal RVEC Model

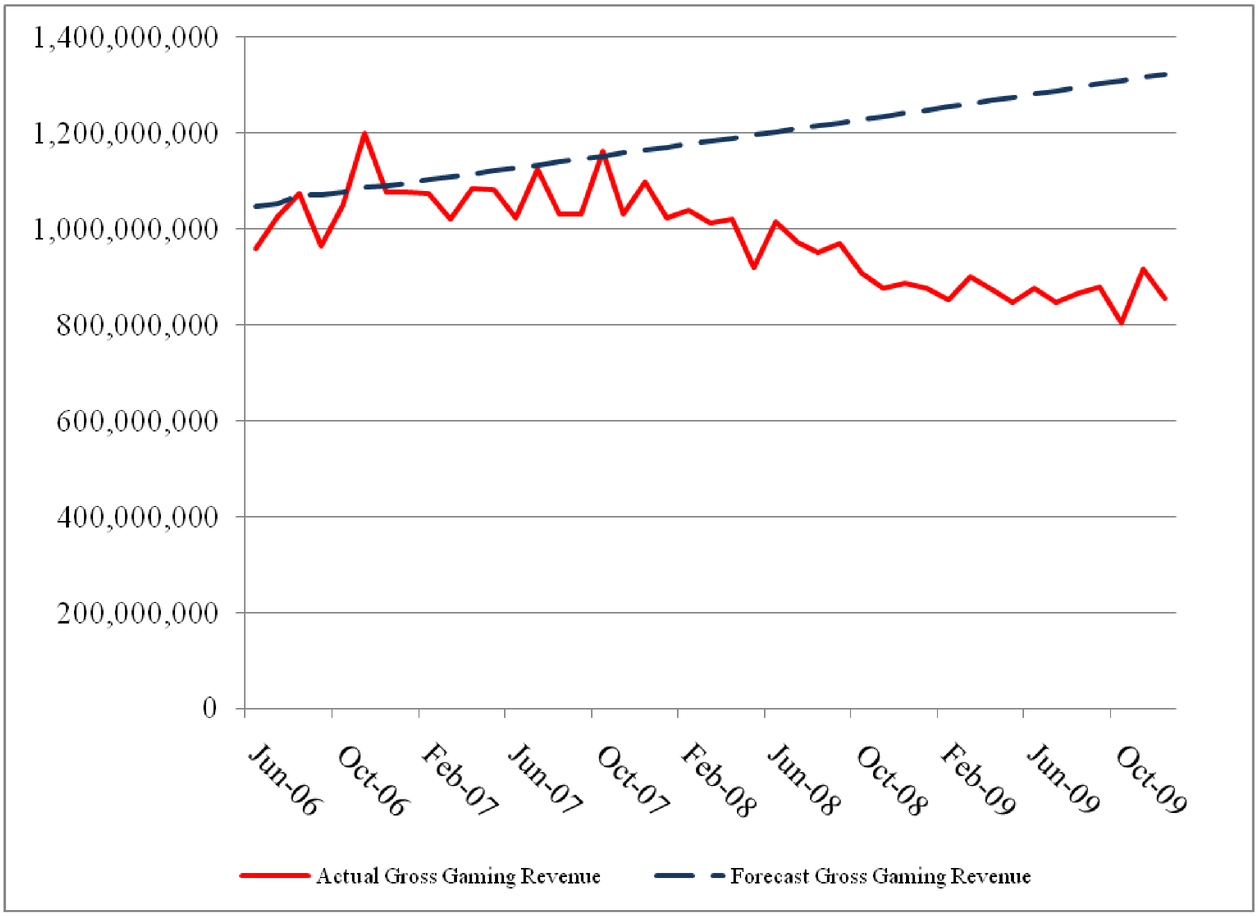


Figure 5: **Forecast and Actual Taxable Sales: Random-Walk Model**

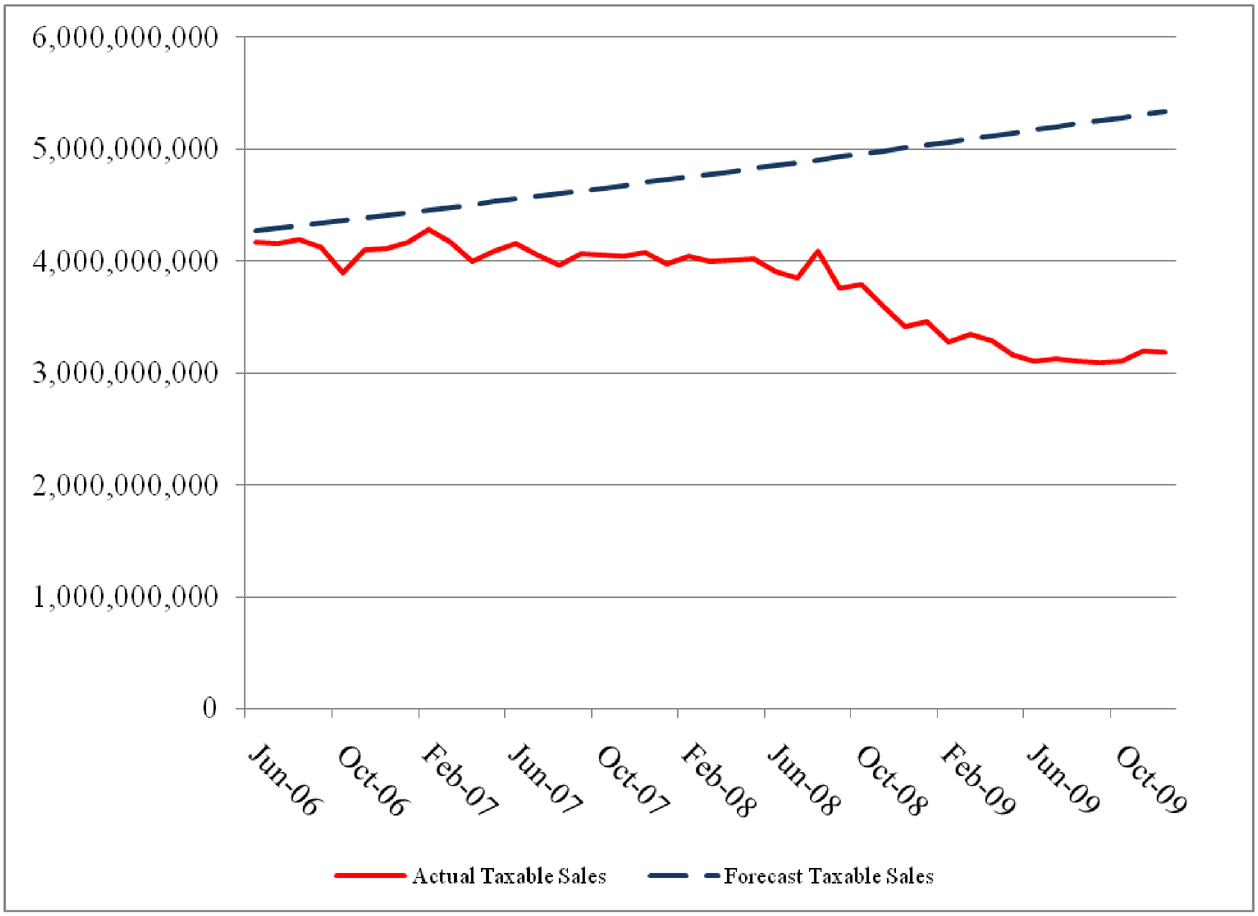


Figure 6: Forecast and Actual Gross Gaming Revenue: Random-Walk Model

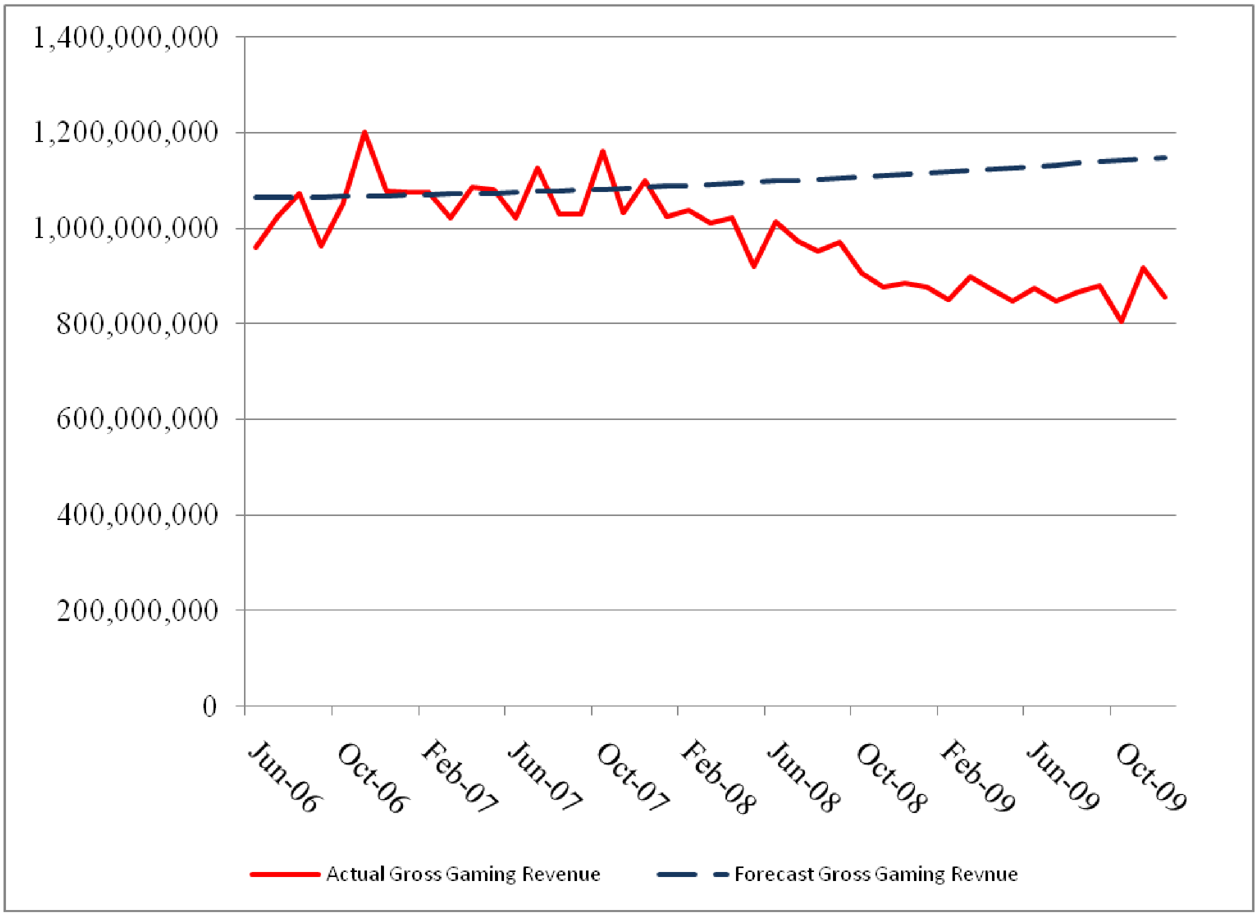


Figure 7: Forecast and Actual Taxable Sales: Optimal Semi-Parametric Model

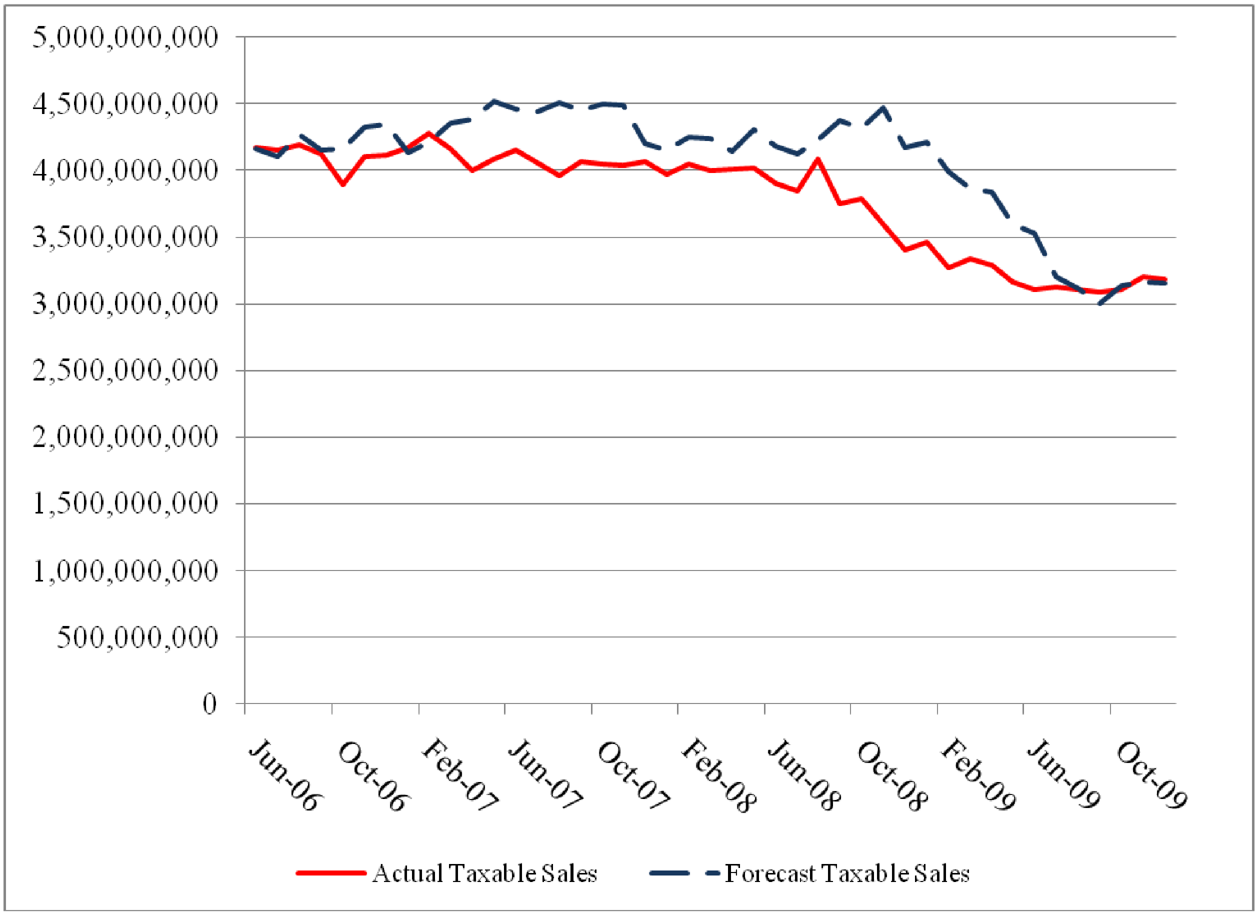


Figure 8: Forecast and Actual Gross Gaming Revenue: Optimal Semi-Parametric Model

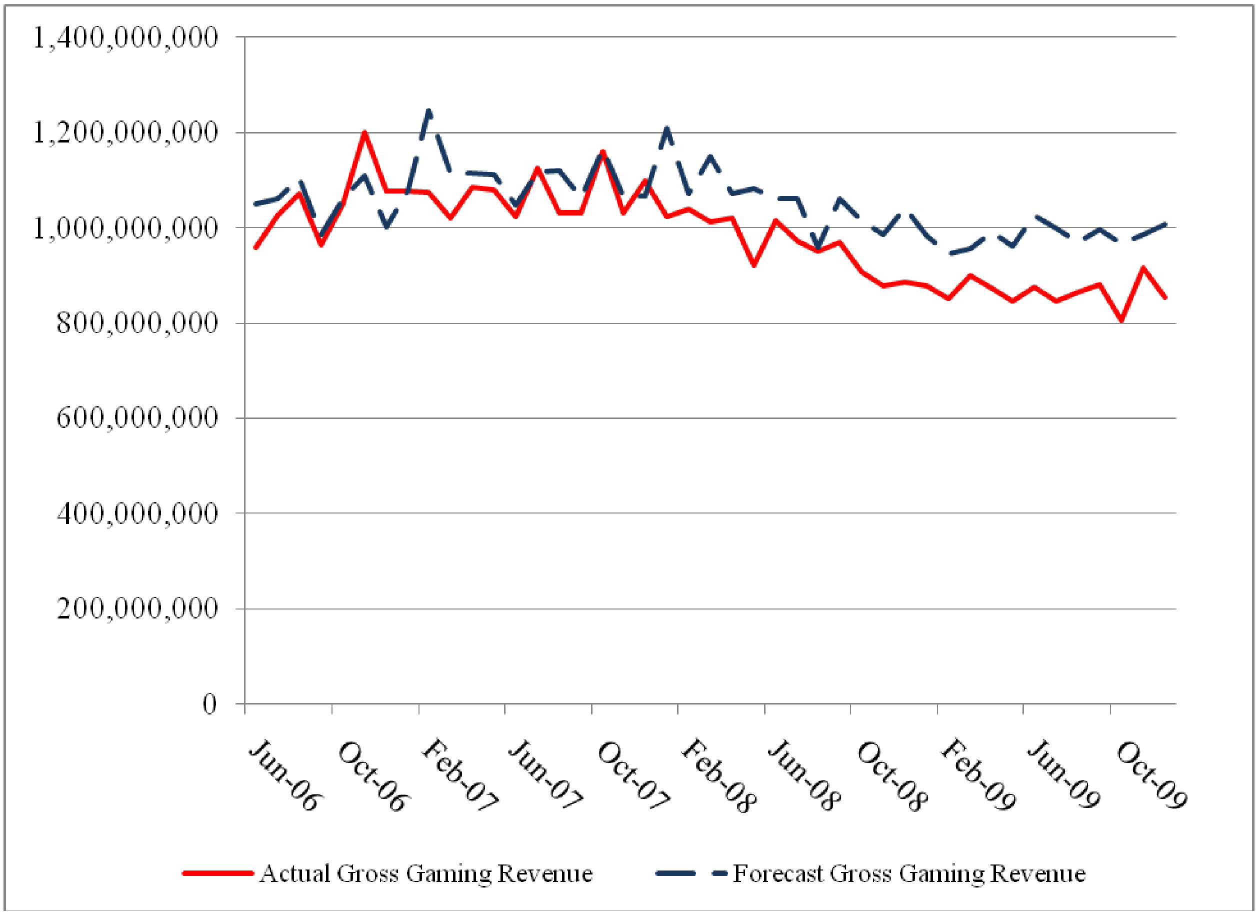


Figure 9: Forecast and Actual Taxable Sales: Smooth Transition Autoregressive Model

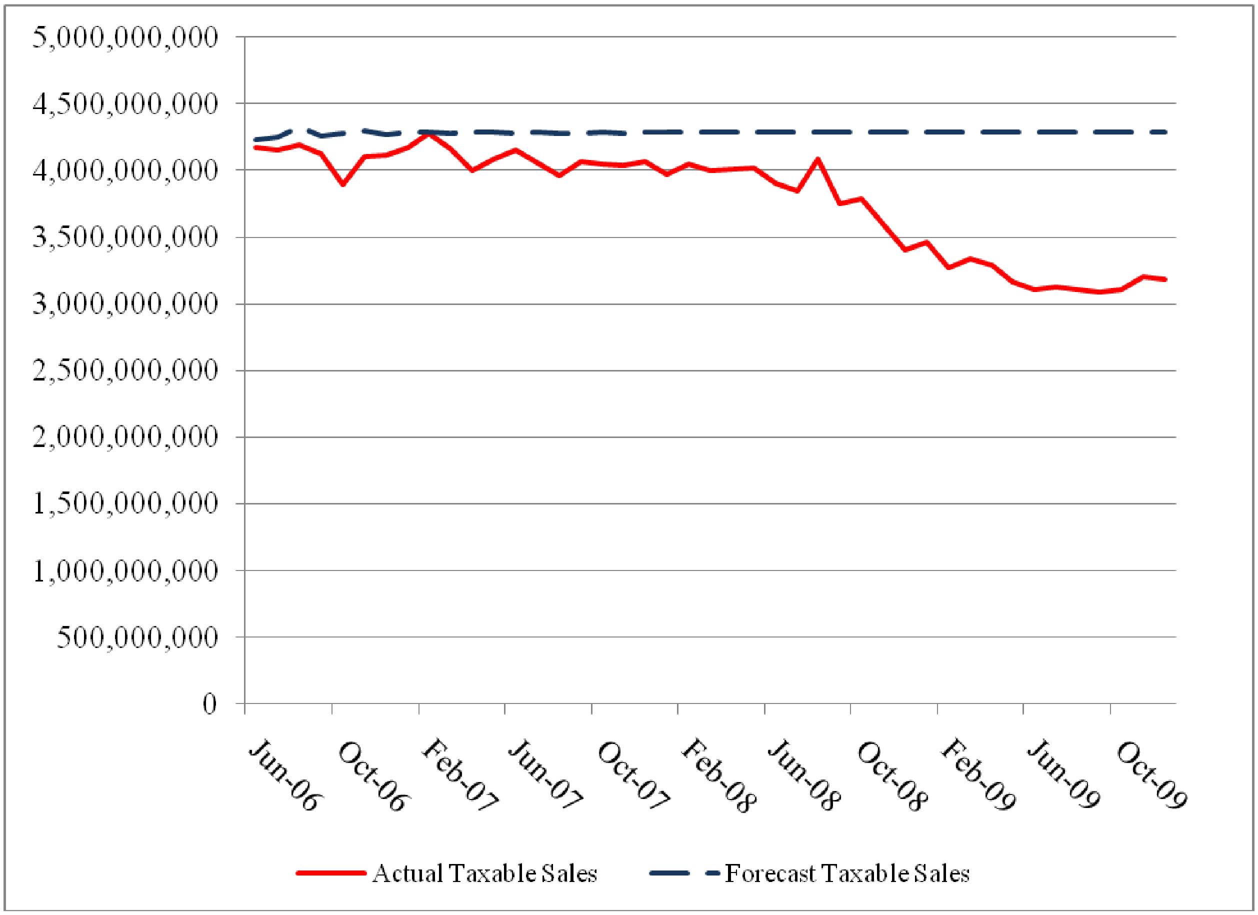


Figure 10: Forecast and Actual Gross Gaming Revenue: Smooth Transition Autoregressive Model

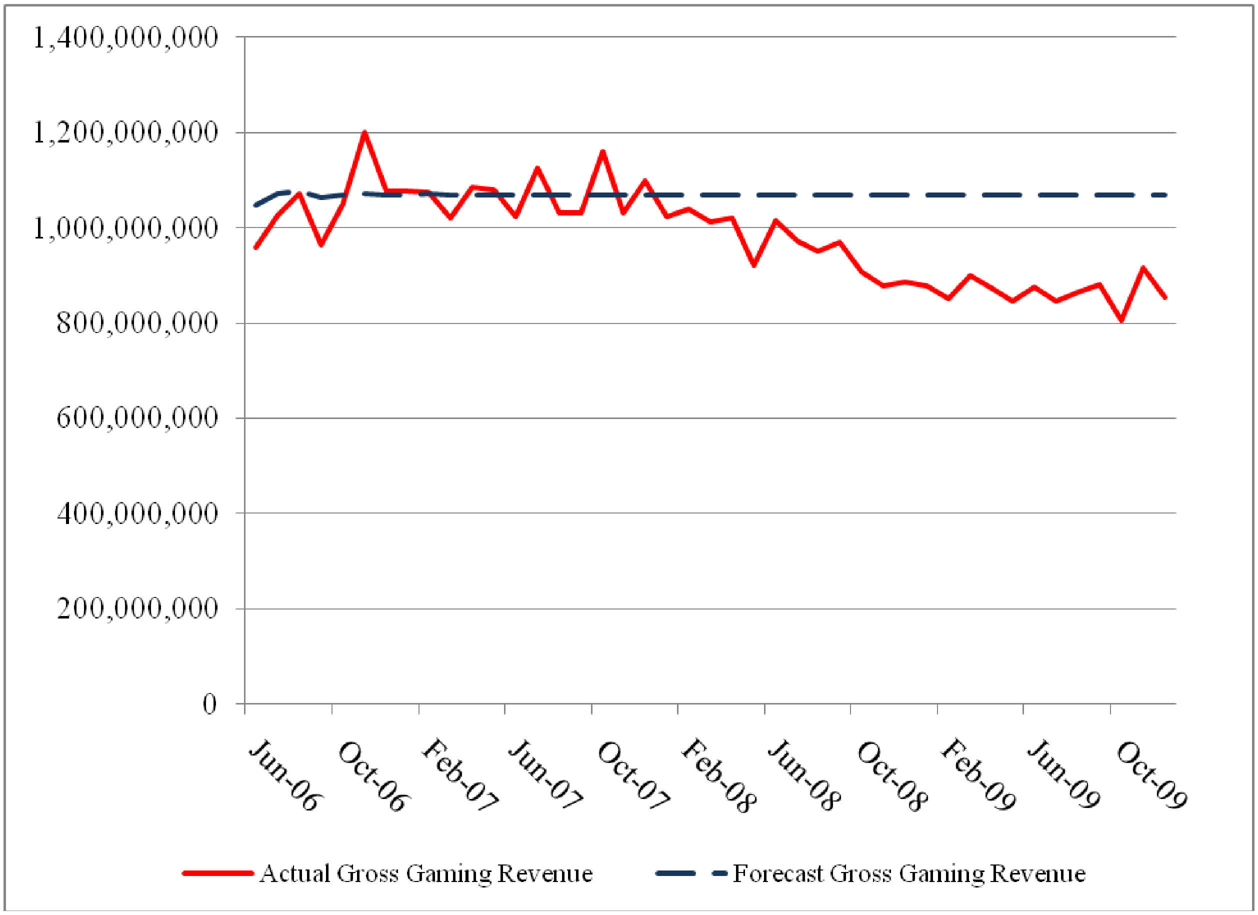


Figure 11: Forecast and Actual Taxable Sales: Artificial Neural Network Autoregressive Model

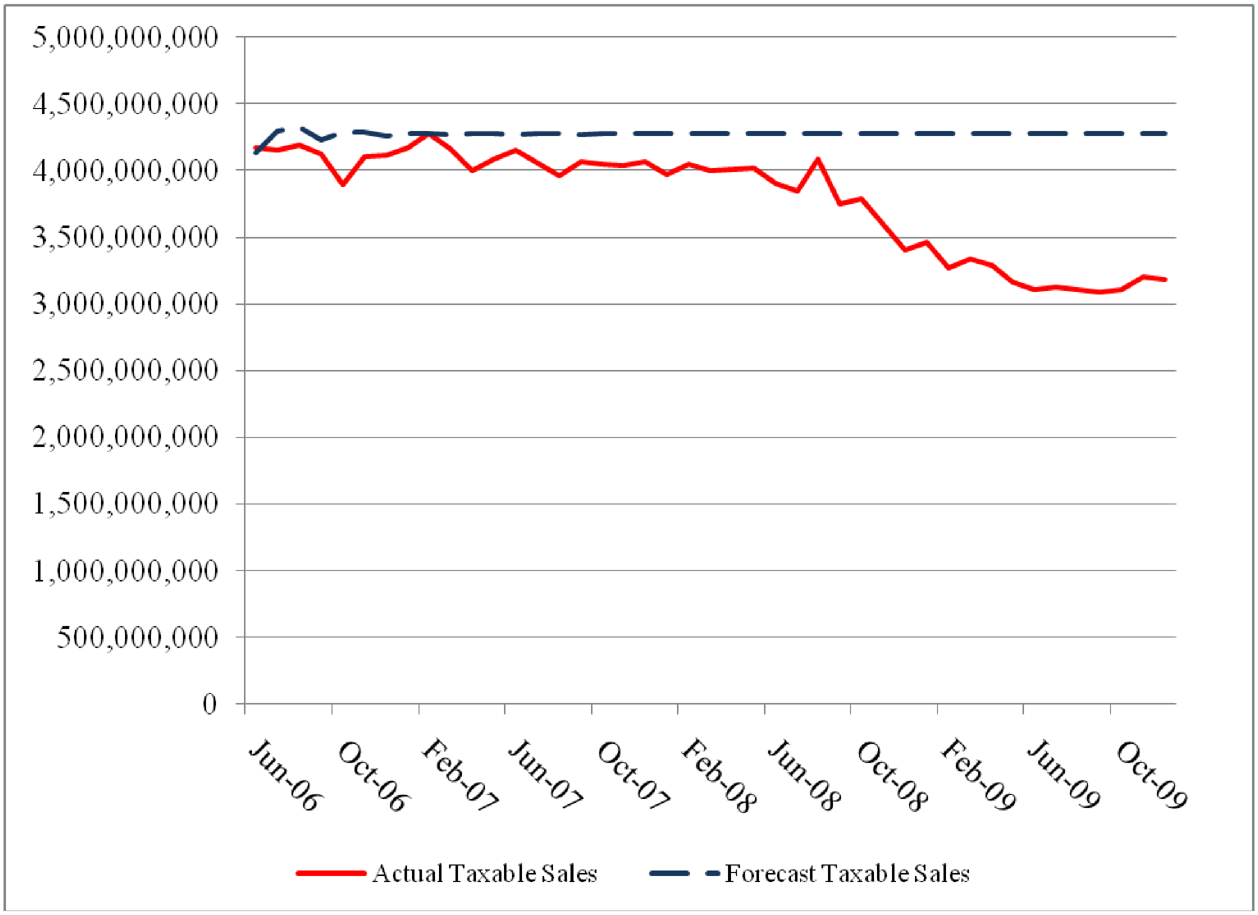
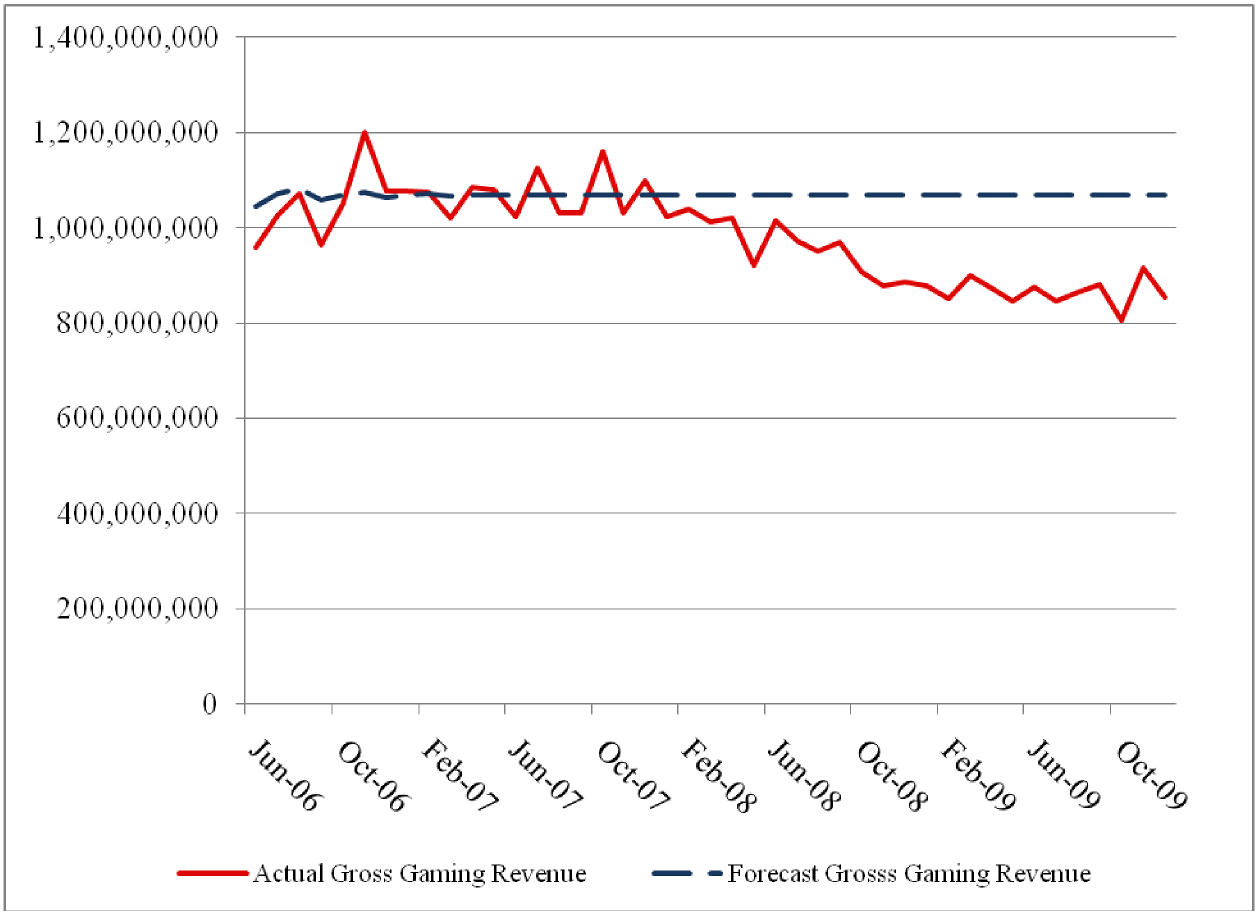


Figure 12: Forecast and Actual Gross Gaming Revenue: Artificial Neural Network Autoregressive Model



Appendix:

For each of the one- to twenty-four-months-ahead forecasts, we test whether the gain (loss) in the RMSE from the AR-ANN model relative to the five other alternative models (random walk, VEC, RVEC, SP, STAR) proves significant using the *ENC-T* test of Clark and McCracken (2001). This test applies to nested models, given that the AR-ANN model nests the random walk model, and, is in turn, nested by the VAR, VECM, SP, STAR models.[25]

The test statistic is defined as follows:

$$\text{(A.1)} \qquad ENC-T = (P-1)^{1/2} \frac{\overline{c}}{(P^{-1}\sum_{t=R}^{T-1}(c_{t+h}-\overline{c}))^{1/2}},$$

where, $c_{t+h} = \hat{v}_{0,t+h}(\hat{v}_{0,t+h} - \hat{v}_{1,t+h})$ and $\bar{c} = \sum_{t=R}^{T-1} c_{t+1}$, R denotes the estimation period, P is the prediction period, f is some generic loss function ($f(v_{0,t+h}) = v_{0,t+h}^2$, in our case), $h \geq 1$ is the forecast horizon, $\hat{\upsilon}_{0,t+h}$ and $\hat{\upsilon}_{1,t+h}$ are h-step ahead prediction errors for models 0 and 1 (where 0 stands for the AR-ANN model and 1 stands, in turn, for the other five models), constructed using Newey and West (1987) type consistent estimators.

The hypotheses of interest are:

$$\text{(A.2)} \qquad H_0 : E(f(\upsilon_{0,t+h}) - f(\upsilon_{1,t+h})) = 0, \text{ and } H_A : E\left(f\left(\upsilon_{0,t+h}\right) - f\left(\upsilon_{1,t+h}\right)\right) > 0.$$

The limiting distribution is $N(0, 1)$ for $h = 1$. The limiting distribution for $h > 1$ is non-standard, as discussed in Clark and McCraken (2001). As long as a Newey and West (1987) type estimator is used when $h > 1$, however, then the tabulated critical values closely approximate the $N(0, 1)$ values (Bhardwaj and Swanson, 2006).

---

[25] In an earlier version, we use the Diebold and Mariano (1995) test for forecast error comparison across the optimal models. The Associate Editor, however, correctly noted that the Diebold and Mariano (1995) test is not appropriate for comparing nested models, since it no longer has a limiting Normal distribution under the null of non-causality. Hence, we use the *ENC-T* test proposed by Clark and McCraken (2001) in this version.