Journal of Business Research 68 (2015) 2295–2298

Contents lists available at ScienceDirect

Journal of Business Research

Forecasting ICT development through quantile confidence intervals☆

Kun-Huang Huarng *,[a], Tiffany Hui-Kuang Yu [b]

[a] Department of International Trade, Feng Chia University, 100 Wenhwa Road, Seatwen, Taichung 40724, Taiwan
[b] Department of Public Finance, Feng Chia University, Taichung, Taiwan

ARTICLE INFO

Article history:
Received 1 September 2014
Received in revised form 1 March 2015
Accepted 1 April 2015
Available online 10 June 2015

Keywords:
Information criterion
Quantile
Regression
Time series

ABSTRACT

Regression is a common method to calculate relationships between variables. Quantile regression extends the calculation to the coefficients of various quantiles, providing a more complete overview. In addition, quantile forecasting models forecast coefficients. This study proposes a new algorithm to calculate the quantile confidence intervals of the in-sample data to forecast the coefficients of the out-of-sample data. The algorithm analyzes ICT data for 78 countries between 1999 and 2010. Results show that the algorithm provides valid forecasting results and outperforms previous studies. These quantile confidence intervals can also forecast the independent variables' impact trends on the dependent variable. The algorithm is applicable to different domains.

1. Introduction

Information and communication technology (ICT) development affects economy (Huarng, 2010; Huarng & Yu, 2011). Hence, studies probe into the factors affecting ICT adoption or development in different regions including Asia (Wong, 2002), Asia-Pacific (Loo & Wong, 2002), the European Union (Vicente & López, 2006), among others (Huarng, 2011; Pohjola, 2003; Wu & Chu, 2010; Yu, 2011).

ICT is playing a major role in economic development in countries. Hence, many studies focus on how ICT affects economy development in countries including Japan (Zuhdi, 2014), Spain (Estapé-Dubreuil & Torreguitart-Mirada, 2014), Russia (Feiguine & Solovjova, 2014), Indonesia (Zuhdi, 2015), Colombia (Gallego, Gutiérrez, & Lee, 2015), Fiji (Kumar & Singh, 2014), and the Czech Republic (Ministra & Pitnerb, 2015).

To provide a more complete overview than regression by medians, Koenker and Bassett (1978) propose quantile regression to infer the results of the conditional functions for different quantiles. Studies apply the quantile regression model to interpret various problems, such as wages (Machado & Mata, 2005; Martins & Pereira, 2004), survival analysis (Koenker & Geling, 2001), financial analysis (Bassett & Chen, 2001), economic growth (Wang, Yu and Liu, 2013), health care expenditure (Yu, Wang, & Chang, 2011), and small business performance (Seo, Perry, Tomczyk, & Solomon, 2014). Using quantile regression, Yu (2011) analyzes the heterogeneous effects of factors on global ICT adoption.

Scholars advance the quantile regression model to forecast. Taylor (2007) uses exponentially weighted quantile regression to forecast the daily supermarket sales, which performs better than the traditional methods. Banachewicz and Lucas (2008) adopt hidden Markov models to forecast the quantiles of corporate default rates. Cai, Stander, and Davies (2012) propose a Bayesian approach to estimating quantile autoregressive time series model and to forecasting on currency exchange rate data.

Regarding interval forecasting, Yu (2014) suggests using a quantile information criterion (QIC) to assist in forecasting. To improve forecasting performance, Huarng and Yu (2014) propose a new QIC (NQIC) to identify if a variable is predictable and thus improve the forecasting results through quantile regressions.

To further improve forecasting, this study proposes an algorithm to forecast similar ICT data sets by calculating the confidence intervals: first using the in-sample data, and then using the confidence intervals to forecast the quantile coefficients of the out-of-sample data. Section 2 proposes the new algorithm. Section 3 introduces the variables and data. Section 4 provides the empirical results and compares the results with those of the previous studies. Section 5 concludes this study.

2. A new forecasting algorithm of quantile confidence intervals

The model for a median linear regression is

$$y_j = x_j \times \beta\tau + \varepsilon_{\tau,j}$$

where the assumption is median $(\varepsilon_{\tau,j}|x_j) = 0$. This concept is extendable to any quantile (i.e., $\tau = 25\%$ or 50%). Quantile regressions use the least

☆ The authors thank the National Science Council, Taiwan, ROC for financial support under grants NSC 102-2410-H-035-038-MY2, and NSC 102-2410-H-035-041-MY2.
* Corresponding author.
E-mail addresses: khhuarng@mail.fcu.edu.tw (K.-H. Huarng), hkyu@mail.fcu.edu.tw (T.H.-K. Yu).

http://dx.doi.org/10.1016/j.jbusres.2015.06.014
0148-2963/

Table 1
Quantile forecasting of x_1.

y's quantile	*90% quantile confidence interval*		*Coefficients*	*Forecasting result*
	1999–2004		*2005*	
20th	$-\infty$	0.0020	0.0005	Correct
40th	−0.0370	0.0110	0.0003	Correct
60th	0.0003	0.0490	0.0026	Correct
80th	−0.0010	∞	0.0027	Correct
100th	0.0001	0.0110	0.0002	Correct
	1999–2005		*2006*	
20th	$-\infty$	0.0030	0.0006	Correct
40th	−0.1040	0.0040	0.0003	Correct
60th	0.0003	0.0650	0.0023	Correct
80th	−0.0003	∞	0.0031	Correct
100th	−0.0001	0.0040	0.0002	Correct
	1999–2006		*2007*	
20th	$-\infty$	0.0023	0.0006	Correct
40th	−0.1030	0.0050	0.0003	Correct
60th	0.0004	0.0600	0.0002	Incorrect
80th	−0.0004	∞	0.0013	Correct
100th	0.0000	0.0050	0.0002	Correct
	1999–2007		*2008*	
20th	$-\infty$	0.0025	0.0006	Correct
40th	−0.0983	0.0059	0.0002	Correct
60th	0.0002	0.0543	0.0002	Correct
80th	0.0004	∞	0.0001	Incorrect
100th	0.0000	0.0053	0.0002	Correct
	1999–2008		*2009*	
20th	$-\infty$	0.0040	0.0005	Correct
40th	−0.1080	0.0043	0.0002	Correct
60th	0.0003	0.0667	0.0001	Incorrect
80th	0.0004	∞	−0.0001	Incorrect
100th	−0.0023	0.0038	0.0001	Correct
	1999–2009		*2010*	
20th	$-\infty$	0.0013	0.0000	Correct
40th	−0.1728	0.0031	0.0000	Correct
60th	0.0002	0.2137	−0.0001	Incorrect
80th	0.0003	∞	0.0003	Correct
100th	−0.0023	0.0028	−0.0001	Correct

absolute deviations method to minimize the absolute values of the errors. Koenker and Bassett (1978) propose a new method to find the quantile as follows:

$$\operatorname{argmin}\left\{\sum\nolimits_{j\in\{j|y_j\geq\varepsilon_{\tau,j}}\tau\left|y_j-\varepsilon_{\tau,j}\right|\sum\nolimits_{j\in\{j|y_j<\varepsilon_{\tau,j}}(1-\tau)\left|y_j-\varepsilon_{\tau,j}\right|\right\}.$$

Suppose dependent variable y and independent variables x_j, ($j = 1, \ldots, p$). To calculate the quantile regression between each $x_j$and y, set confidence interval to a percentile; each regression has$\tau 1$, $\tau 2\ldots$ and τk quantiles.

For a time series,time periods, $t = 1, .. m - 1, m, m + 1, \ldots n - 1, n$.For x_j, the data of $t = 1$ to m are in-sample and data of $t = m + 1$ are out-of-sample. This study proposes an algorithm to calculate quantile confidence intervals on the basis of the in-sample data to forecast if the out-of-sample quantile coefficient is within the corresponding quantile confidence intervals. The time span of the in-sample data increases by 1 in the following time period repeatedly until the time series exhausts.

Appendix shows the algorithm in detail.

3. Variables and data set

Huarng (2011) uses Internet users per 100 inhabitants as the technical variable and GDP per capita as the economic variable to cluster the ICT development in 121 economies. Yu (2011) examines the heterogeneous effects of population density, GDP per capita, and telephone lines per 100 inhabitants on global ICT adoption (Internet users per 100 inhabitants). Yu (2014) and Huarng and Yu (2014) propose in-sample data to generate QIC and NQIC, respectively, to forecast the same data set as Yu (2011).Using similar variables, Huarng (2015) applies fuzzy set/Qualitative Comparative Analysis to conduct ICT development analysis and forecasting.

Following these studies, this research uses the variables as follows:

$$y : Internet\_users\_per\_100\_inhabi\,\text{tants} = \frac{total\_(fixed)\_Internet\_users}{population} \times 100\infty$$
$$x_1 : population\_density$$
$$x_2 : GDP\_per\_capita$$
$$x_3 : telephone\_lines\_per\_100\_inhabi\,\text{tants} = \frac{total\_main\_lines}{population} \times 100.$$

The data are from the World Telecommunication/ICT Indicators database by International Telecommunication Union, covering 78 countries, ranging from 1999 to 2007. This study uses the same data set but extending the data to 2010.

4. Empirical analysis

4.1. Forecasting results

This study applies SAS to calculate the confidence intervals using ranks (Chen, 2005). In this analysis, the quantiles for y are 20th, 40th, 60th, 80th, and 100th, respectively. All the analyses have a 90% confidence interval.

First, the in-sample data from 1999 to 2004 serves to calculate the quantile confidence intervals that then forecast the out-of-sample quantile coefficients of 2005. This process progressively completes all

Table 2
Quantile forecasting of x_2.

y's quantile	*90% quantile confidence interval*		*Coefficients*	*Forecasting result*
	1999–2004		*2005*	
20th	−0.0004	0.0002	0.0003	Incorrect
40th	0.0001	0.0002	0.0003	Incorrect
60th	0.0000	0.0012	0.0004	Correct
80th	−0.0003	0.0010	0.0004	Correct
100th	0.0001	0.0012	0.0003	Correct
	1999–2005		*2006*	
20th	0.0002	0.0002	0.0002	Correct
40th	0.0001	0.0001	0.0003	Incorrect
60th	−0.0001	0.0010	0.0003	Correct
80th	−0.0002	0.0008	0.0005	Correct
100th	0.0001	0.0012	0.0003	Correct
	1999–2006		*2007*	
20th	0.0001	0.0002	0.0002	Correct
40th	0.0001	0.0011	0.0002	Correct
60th	−0.0001	0.0010	0.0002	Correct
80th	−0.0002	0.0006	0.0003	Correct
100th	0.0001	0.0012	0.0002	Correct
	1999–2007		*2008*	
20th	0.0001	0.0004	0.0001	Correct
40th	0.0001	0.0005	0.0002	Correct
60th	0.0001	0.0004	0.0002	Correct
80th	0.0000	0.0006	0.0001	Correct
100th	0.0001	0.0006	0.0002	Correct
	1999–2008		*2009*	
20th	0.0001	0.0003	0.0001	Correct
40th	0.0001	0.0004	0.0002	Correct
60th	0.0000	0.0003	0.0003	Correct
80th	0.0000	0.0005	0.0002	Correct
100th	0.0001	0.0005	0.0003	Correct
	1999–2009		*2010*	
20th	0.0001	0.0004	0.0003	Correct
40th	0.0001	0.0004	0.0003	Correct
60th	0.0000	0.0003	0.0002	Correct
80th	0.0001	0.0004	0.0002	Correct
100th	0.0001	0.0004	0.0003	Correct