Prediction of electricity prices for Central American countries using dynamic conditional score models

SZABOLCS BLAZSEK, HECTOR HERNÁNDEZ
UNIVERSIDAD FRANCISCO MARROQUÍN, GUATEMALA
Electricity sector in Central America

Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua and Panama

Well-functioning and competitive electricity market

Six Central American countries (all with the exception of Belize) are interconnected and participate in the Regional Electricity Market.

Application of adequate electricity price models is important for potential investors, market participants and regulators.
Spot electricity price data

A unique dataset of daily average spot electricity prices in USD per megawatt hours (MWh) $p_t$ for every calendar day

For three Central American countries:

**El Salvador** (January 2010 to December 2014, $T = 1,826$)

**Guatemala** (January 2005 to December 2014, $T = 3,652$)

**Panama** (July 1998 to December 2014, $T = 6,028$)

*(full data window)*
El Salvador $p_t$, two subperiods

- 2010 to 2011
- 2012 to 2014
Guatemala $p_t$, two subperiods

2005 to 2009

2010 to 2014
Panama $p_t$, two subperiods

1998 to 2006

2007 to 2014
Data windows for out-of-sample forecasts

For each country three **estimation windows:**
Start at the beginning of the full data window

For each country three **forecast evaluation windows:**
2012, 2013, 2014 (correspond to the estimation windows)
Properties of Central American electricity prices

(i) Mean reversion (?)
(ii) Weekly seasonal
(iii) Annual seasonal
(iv) Heteroscedasticity
Mean reversion for electricity prices

Are the Central American electricity prices mean reverting? We do not know (a priori).

In the body of literature of electricity price models the majority of works propose mean-reverting models, but also there are works where non-mean-reverting models are used (see Escribano, Peña, Villaplana, 2011).

We formulate both I(0) and I(1) alternatives for the level of $p_t$. 
Weekly seasonal of electricity prices

As the Central American countries are not industrialized, electricity demand decreases significantly during the weekends.

The carbon-based and geothermic baseload power plants stop on Saturday morning and restart on Sunday evening. Therefore, electricity prices attenuate during the weekend.

On Monday, demand for electricity augments. For baseload plants, about 48 hours are needed to restart a plant. Therefore, on Monday and Tuesday, electricity is generated by more expensive technologies (steam, biomass and diesel power plants), which increase prices.
Annual seasonal of electricity prices

There is an annual seasonal in electricity price time series for Central America due to
(1) significant installed capacity of hydroelectricity generation and
(2) seasonal weather-related water supply conditions
(1) Installed hydropower capacity

<table>
<thead>
<tr>
<th></th>
<th>Central America</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MWh (2014)</td>
</tr>
<tr>
<td>Thermal power</td>
<td>6,714.50</td>
</tr>
<tr>
<td>Hydropower</td>
<td>5,721.30</td>
</tr>
<tr>
<td>Geothermal energy</td>
<td>625.60</td>
</tr>
<tr>
<td>Wind power</td>
<td>589.70</td>
</tr>
<tr>
<td>Solar power</td>
<td>8.40</td>
</tr>
<tr>
<td>Bioenergy</td>
<td>7.70</td>
</tr>
<tr>
<td>Total</td>
<td>13,667.20</td>
</tr>
</tbody>
</table>
(2) Seasonal weather-related water supply

In Central America, the water supply for hydroelectric power plants is different during the winter and summer. The winter season in Central America is approximately from May to October or November. During the winter there is a higher water supply, and due to excess generation the prices are lower.

On the other hand, during the summer due to a lower water supply, more coal is used to generate thermal power and hence prices increase significantly.
Heteroscedasticity in electricity prices

Duffie, Gray, Hoang (1998):
  GARCH may be non-stationary for electricity prices
Knittel and Roberts (2005):
  EGARCH, leverage effect parameter is significant
Escribano, Peña, Villaplana (2011):
  GARCH with time-varying jumps or falls
Heteroscedasticity in electricity prices

(i) We do not consider explicitly jumps or falls in the econometric framework, but replace GARCH with Dynamic Conditional Score (DCS) volatility models. As DCS volatility models are robust to outliers (Harvey 2013), we may obtain appropriate statistical models for electricity prices.

(ii) We consider DCS volatility models with leverage effects.
Econometric models

(1) Standard financial time series model

(2) DCS model for location, stochastic seasonal and DCS volatility
(1) Standard financial time series model

\[ p_t = \mu_t + s_{w,t} + s_{a,t} + \nu_t = \mu_t + s_{w,t} + s_{a,t} + \sqrt{\lambda_t} \epsilon_t \]
\[ \epsilon_t \text{ i.i.d.} \]

Two alternatives for the level:

ARMA(1,1): \( \mu_t = c + \varphi \mu_{t-1} + \theta \nu_{t-1} \) with \( |\varphi| < 1 \)

RW: \( \mu_t = \mu_{t-1} + \theta \nu_{t-1} \)
(1) Standard financial time series model

$s_{w,t}$ includes weekday dummies with time-invariant parameters
$s_{a,t}$ includes month dummies with time-invariant parameters
(parameters sum to zero; parameter for the first observation is zero)

$\nu_t$ is normal-GARCH (Bollerslev, 1986; Taylor, 1986) with leverage effects (Glosten, Jagannathan, Runkle, 1993)
DCS model for location and stochastic seasonal, body of literature

Harvey (2013) and Harvey and Luati (2014):
Dynamic Student-\(t\) location plus stochastic seasonal DCS model

Application to United Kingdom (UK) rail travel data
Log of number of kilometers travelled by UK passengers
Monthly data
DCS model for location and stochastic seasonal, our extension

We extend the dynamic Student-\( t \) location plus stochastic seasonal DCS model of Harvey (2013) and Harvey and Luati (2014), as we consider

(i) alternatives of the \( t \)-distribution for the error term \( \varepsilon_t \)

(ii) exponential DCS volatility specification
(i) Distributions of the error term

\( \varepsilon_t \sim t(\nu) \) (Harvey and Chakravarty, 2008)
\( \varepsilon_t \sim \text{GED}(\nu) \) (General Error Distribution) (Harvey, 2013)
\( \varepsilon_t \sim \text{EGB2}(0,1, \xi, \zeta) \) (Exponential Generalized Beta distribution of the second kind) (Caivano and Harvey, 2014)
\( \varepsilon_t \sim \text{Gen-}t(\nu, \eta) \) (Generalized-\( t \)) (Harvey and Lange, 2015)
(ii) Exponential DCS volatility models

Beta-\( t \)-EGARCH (Harvey and Chakravarty, 2008)
Gamma-GED-EGARCH (Harvey, 2013)
EGB2-EGARCH (Caivano and Harvey, 2014)
Beta-Gen-\( t \)-EGARCH (Harvey and Lange, 2015)

*For all volatility models we measure leverage effects.*
(2) DCS model for location, stochastic seasonal and exponential DCS volatility

Our DCS model for location, stochastic seasonal and DCS volatility Central American of electricity prices:

\[ p_t = \mu_t + s_{w,t} + s_{a,t} + \nu_t = \mu_t + s_{w,t} + s_{a,t} + \exp(\lambda_t)\varepsilon_t \]

\[ \varepsilon_t \text{ i.i.d.} \]
Level

Two alternatives for the level:
QAR(1): $\mu_t = c + \varphi \mu_{t-1} + \theta e_{t-1}$ with $|\varphi| < 1$
RW: $\mu_t = \mu_{t-1} + \theta e_{t-1}$

We initialize $\mu_t$ by parameter $\mu_0$ or $(1/7) \sum_{t=1}^{7} p_t$.

$e_t$ is proportional to the conditional score with respect to $\mu_t$:
$e_t \propto \partial \ln f(p_t | p_1, \ldots, p_{t-1}) / \partial \mu_t$
Level

\[ e_t = \left(1 + \frac{\varepsilon_t^2}{\nu}\right)^{-1} \exp(\lambda_t) \varepsilon_t \text{ for } \nu > 0, \varepsilon_t \sim t(\nu) \]

\[ e_t = \text{sign}[\exp(\lambda_t)\varepsilon_t] \times |\varepsilon_t|^{\nu-1} \text{ for } \nu > 1, \varepsilon_t \sim \text{GED}(\nu) \]

\[ e_t = (\xi + \zeta) \frac{\exp[\exp(\lambda_t)\varepsilon_t]}{1+\exp[\exp(\lambda_t)\varepsilon_t]} - \xi \text{ for } \xi, \zeta > 0, \varepsilon_t \sim \text{EGB2}(0,1, \xi, \zeta) \]

\[ e_t = \frac{(\nu+1)\varepsilon_t|\varepsilon_t|^{\eta-2}}{\exp(\lambda_t)(\nu+1|\varepsilon_t|^{\eta})} \text{ for } \nu, \eta > 0, \varepsilon_t \sim \text{Gen-t}(\nu, \eta) \]
$e_t$ as a function of $\varepsilon_t$, update of the level
Weekly seasonality

\[ s_{w,t} = D'_{w,t} \delta_t \]

\[ D'_{w,t} = (D_{Mo,t}, D_{Tu,t}, D_{We,t}, D_{Th,t}, D_{Fr,t}, D_{Sa,t}, D_{Su,t}) \]

\[ \delta_t = \delta_{t-1} + \kappa_{w,t} e_t \]

\[ \kappa'_{w,t} = (\kappa_{w,1t}, \ldots, \kappa_{w,7t}) \]

\[ \kappa_{w,jt} = \begin{cases} 
\kappa_{w,j} & \text{for } D_{w,jt} = 1 \\
-\frac{\kappa_{w,j}}{(7-1)} & \text{for } D_{w,jt} = 0 
\end{cases} \]
Annual seasonality

\[ s_{a,t} = D'_{a,t} \gamma_t \]
\[ D'_{a,t} = (D_{Jan,t}, \ldots, D_{Dec,t}) \]
\[ \gamma_t = \gamma_{t-1} + \kappa_{a,t} e_t \]
\[ \kappa'_{a,t} = (\kappa_{a,1t}, \ldots, \kappa_{a,12t}) \]
\[ \kappa_{a,jt} = \begin{cases} 
\kappa_{a,j} & \text{for } D_{a,jt} = 1 \\
- \frac{\kappa_{a,j}}{(12-1)} & \text{for } D_{a,jt} = 0 
\end{cases} \]
Initialize weekly and annual seasonality

We initialize the filters $\delta_t$ and $\gamma_t$ by estimating a linear regression model with constant, linear time trend and all weekday and month dummies (Harvey and Luati, 2014). We assume that the sum of all weekday parameters is zero. We assume that the sum of all month parameters is zero. We estimate the linear regression by using the Non-linear Least Squares (NLS) method. The estimates of dummy parameters give $\delta_0$ and $\gamma_0$. 
Exponential DCS volatility models

\[ \lambda_t = \omega + \beta \lambda_{t-1} + \alpha u_{t-1} + \alpha \cdot \text{sign}(v_{t-1}) (u_{t-1} + 1) \]

We initialize this filter by parameter \( \lambda_0 \).

\[ u_t \propto \partial \ln f(p_t|p_1, \ldots, p_{t-1}) / \partial \lambda_t \]
Exponential DCS volatility models

\[ u_t = \frac{(v+1)\varepsilon_t^2}{v+\varepsilon_t^2} - 1 \text{ for } v > 0, \varepsilon_t \sim t(v) \]

\[ u_t = \frac{v}{2} |\varepsilon_t|^v - 1 \text{ for } v > 1, \varepsilon_t \sim \text{GED}(v) \]

\[ u_t = (\xi + \zeta) \frac{\varepsilon_t \exp[\varepsilon_t]}{1+\exp[\varepsilon_t]} - \xi \varepsilon_t - 1 \text{ for } \xi, \zeta > 0, \varepsilon_t \sim \text{EGB2}(0,1,\xi,\zeta) \]

\[ u_t = \frac{(v+1)|\varepsilon_t|^\eta}{v+|\varepsilon_t|^\eta} - 1 \text{ for } v, \eta > 0, \varepsilon_t \sim \text{Gen-t}(v, \eta) \]
$u_t$ as a function of $\varepsilon_t$, update of the scale
Statistical inference

Maximum Likelihood (ML) method

Robust sandwich estimator for standard errors

Delta method for transformed parameters

Start parameter values for DCS model are obtained by a preliminary four-step ML estimation procedure for \((\mu_t, s_{w,t}, s_{a,t}, \lambda_t)\).

Criterion for effective ML estimation is convergence to the maximum LL with convergence tolerance for the gradient \(10^{-5}\).
Statistical inference

For the standard financial time series model, the ML estimation procedure always converged effectively.

For the DCS models with $\varepsilon_t \sim t(\nu)$, the ML estimation procedure converged effectively for most of the data windows.

For the DCS models with $\varepsilon_t \sim \text{GED}(\nu)$, $\varepsilon_t \sim \text{EGB2}(0, 1, \xi, \zeta)$ and $\varepsilon_t \sim \text{Gen-t}(\nu, \eta)$, the ML estimation procedure never converged effectively. We do not report results for these DCS models.
<table>
<thead>
<tr>
<th>Panel B. Mean BIC</th>
<th>Standard ARMA</th>
<th>Standard RW</th>
<th>DCS QAR</th>
<th>DCS RW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full data window</td>
<td>7.8153</td>
<td>7.8132</td>
<td>NA</td>
<td>7.6702*</td>
</tr>
<tr>
<td>First subperiod</td>
<td>8.0866</td>
<td>8.0812</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Second subperiod</td>
<td>7.5450</td>
<td>7.5383</td>
<td>7.3420</td>
<td>7.3320*</td>
</tr>
<tr>
<td>First estimation window</td>
<td>8.0866</td>
<td>8.0812</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Second estimation window</td>
<td>8.0113</td>
<td>8.0083</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Third estimation window</td>
<td>7.9545</td>
<td>7.9542</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
## Likelihood-based performance, Guatemala

<table>
<thead>
<tr>
<th>Panel B. Mean BIC</th>
<th>Standard ARMA</th>
<th>Standard RW</th>
<th>DCS QAR</th>
<th>DCS RW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full data window</td>
<td>7.0469</td>
<td>7.0437</td>
<td>NA</td>
<td>6.9991*</td>
</tr>
<tr>
<td>First subperiod</td>
<td>6.1875</td>
<td>6.1846</td>
<td>NA</td>
<td>6.1021*</td>
</tr>
<tr>
<td>Second subperiod</td>
<td>7.8408</td>
<td>7.8357</td>
<td>NA</td>
<td>7.7520*</td>
</tr>
<tr>
<td>First estimation window</td>
<td>6.5455</td>
<td>6.5426</td>
<td>NA</td>
<td>6.4953*</td>
</tr>
<tr>
<td>Second estimation window</td>
<td>6.7435</td>
<td>6.7400</td>
<td>NA</td>
<td>6.6671*</td>
</tr>
<tr>
<td>Third estimation window</td>
<td>6.9236</td>
<td>6.9202</td>
<td>NA</td>
<td>6.8684*</td>
</tr>
</tbody>
</table>
## Likelihood-based performance, Panama

<table>
<thead>
<tr>
<th>Panel B. Mean BIC</th>
<th>Standard ARMA</th>
<th>Standard RW</th>
<th>DCS QAR</th>
<th>DCS RW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full data window</td>
<td>7.1087</td>
<td>7.1462</td>
<td>NA</td>
<td>6.9846*</td>
</tr>
<tr>
<td>First subperiod</td>
<td>5.8375</td>
<td>5.8346</td>
<td>5.4638*</td>
<td>5.4701</td>
</tr>
<tr>
<td>Second subperiod</td>
<td>8.3599</td>
<td>8.3738</td>
<td>8.1241*</td>
<td>8.1351</td>
</tr>
<tr>
<td>First estimation window</td>
<td>6.7788</td>
<td>6.8093</td>
<td>NA</td>
<td>6.6291*</td>
</tr>
<tr>
<td>Second estimation window</td>
<td>6.8780</td>
<td>6.9046</td>
<td>NA</td>
<td>6.7812*</td>
</tr>
<tr>
<td>Third estimation window</td>
<td>6.9836</td>
<td>7.0087</td>
<td>6.8551*</td>
<td>6.8575</td>
</tr>
</tbody>
</table>
### Out-of-sample point forecast performance

**Panel A. Point forecasts performance**

<table>
<thead>
<tr>
<th>Country</th>
<th>RMSE, standard</th>
<th>RMSE, DCS</th>
<th>DM statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guatemala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First forecast evaluation window</td>
<td>13.9273</td>
<td>13.7405</td>
<td>0.3308</td>
</tr>
<tr>
<td>Second forecast evaluation window</td>
<td>16.0387</td>
<td>15.7209</td>
<td>0.6213</td>
</tr>
<tr>
<td>Third forecast evaluation window</td>
<td>15.3719</td>
<td>14.2891</td>
<td>2.1019**</td>
</tr>
<tr>
<td>Panama</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First forecast evaluation window</td>
<td>16.0743</td>
<td>19.6480</td>
<td>−0.8197</td>
</tr>
<tr>
<td>Second forecast evaluation window</td>
<td>18.3090</td>
<td>20.3240</td>
<td>−0.3637</td>
</tr>
<tr>
<td>Third forecast evaluation window</td>
<td>22.8993</td>
<td>21.2295</td>
<td>1.8163*</td>
</tr>
<tr>
<td></td>
<td>mean ln $f$, standard</td>
<td>mean ln $g$, DCS</td>
<td>AG statistic</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>Guatemala</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First forecast evaluation window</td>
<td>$-4.1043$</td>
<td>$-3.8584$</td>
<td>$-2.2998^{**}$</td>
</tr>
<tr>
<td>Second forecast evaluation window</td>
<td>$-4.6228$</td>
<td>$-4.0552$</td>
<td>$-2.5651^{**}$</td>
</tr>
<tr>
<td>Third forecast evaluation window</td>
<td>$-4.1009$</td>
<td>$-3.9492$</td>
<td>$-3.0872^{***}$</td>
</tr>
<tr>
<td><strong>Panama</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First forecast evaluation window</td>
<td>$-5.2971$</td>
<td>$-4.1994$</td>
<td>$-1.9584^{*}$</td>
</tr>
<tr>
<td>Second forecast evaluation window</td>
<td>$-7.3279$</td>
<td>$-4.1431$</td>
<td>$-2.0952^{**}$</td>
</tr>
<tr>
<td>Third forecast evaluation window</td>
<td>$-4.4754$</td>
<td>$-4.2400$</td>
<td>$-3.9067^{***}$</td>
</tr>
</tbody>
</table>
Weekly season, El Salvador, 2010 to 2014
Annual season, El Salvador, 2010 to 2014
Weekly season, Guatemala, 2005 to 2014
Annual season, Guatemala, 2005 to 2014
Weekly season, Panama, 1998 to 2014
Annual season, Panama, 1998 to 2014
Thank you for your attention!