Early Monthly Estimation of Mexico's Manufacturing Production Level using Electric Energy Consumption data

Hugo Hernández Ramos

hugo.hernandez@inegi.org.mx

1. Introduction

The Monthly Production Level in Mexico's Manufacturing Sector is a key macroeconomic variable which gives important clues to policymakers about the health status of the National Economy. This work investigates the empirical functional relationship between **electric energy consumption** and **production level** in Mexico's Manufacturing Sector.

Traditional national accounting methods generate a figure for the Monthly Manufacturing Production Level Index (IMAI3133), approximately 40 days after the end of the reference month, by using information from **regular monthly economic surveys**.

On the other hand, **monthly administrative data** on electric energy consumption for industrial, trade and service establishments in Mexico, provided by the Federal Electricity Commission (CFE), and available to INEGI approximately 15 days after the end of reference month, show that, in the manufacturing sector, there is a statistically significant linear relationship between aggregated electric energy consumption and the IMAI3133 index. This enables to produce early estimates for IMAI3133 as a function of electric energy consumption.

2. Modeling design

Back in late 2015, tests were conducted on three linear regression modeling variants:

1. **Simple two variable linear regression model**: $Y_t = \alpha + \beta X_t + \varepsilon_t$
2. **Autoregressive Model with Distributed Lags**: $Y_t = \alpha + \rho Y_{t-1} + \beta X_t + \gamma X_{t-1} + \varepsilon_t$
3. **Logarithmic Differences**:

   $\Delta\, ln(Y_t) = \beta \Delta\, ln(X_t) + \varepsilon_t$, with $\Delta\, ln(X_t) = \ln(X_t) - \ln\,(X_{t-1})$

The objective back then was to select a working linear model for **nowcasting** purposes.

To construct the models, an **Electric Energy Consumption Index (ICEE) was used as explanatory variable X_t**, and the **IMAI3133 indicator was used as response variable Y_t**, spanning the period January 2013 – October 2015 (see Fig. 1).

A battery of hypothesis tests was used in order to assess statistical model adequacy. Such tests include the usual t-tests and F-tests for regression coefficients, the **Durbin Watson** test for 1st order residual autocorrelation, the **Breusch-Pagan** test for residual heteroskedasticity, the **Cramér - Von Mises** test for residual normality, and the analysis of sample **ACF-PACF** graphs for ruling out residual autocorrelation of any order. Forecasting accuracy assessment was done by using out-of-sample ICEE and IMAI3133 values, comparing in **real time** generated forecasts against officially published IMAI3133 values.

2.1 Construction of explanatory variable

Large Manufacturing establishments from INEGI's **Statistical Business Register (SBR)** were linked to CFE data, in order to build a sample for constructing the ICEE index. As of reference month February 2020, this sample contains 17,151 large manufacturing establishments, with economic activity class codes from the SBR and monthly electric energy consumption values (in KWh). The procedure for constructing the ICEE index is as follows:

1. For all establishments in each **manufacturing subsector** s (food industry, automotive industry, chemical industry, textile industry, etc.), and for each month $t = 1,2,\dots,N$, sum their electric energy consumption values, in order to obtain aggregated $c_{s,t}$ quantities.
2. Compute $\mathrm{ICEE}_t = \sum_s w_s c_{s,t}$, where w_s are relative manufacturing subsector weights, determined by National Accounts; these weights are such that $\sum_s w_s = 1$.

Re-scale ICEE_t so that $\mathrm{ICEE}_1 = \mathrm{IMAI3133}_1$; $t = 1$ corresponds to January 2013.

2.2 Variables for building models

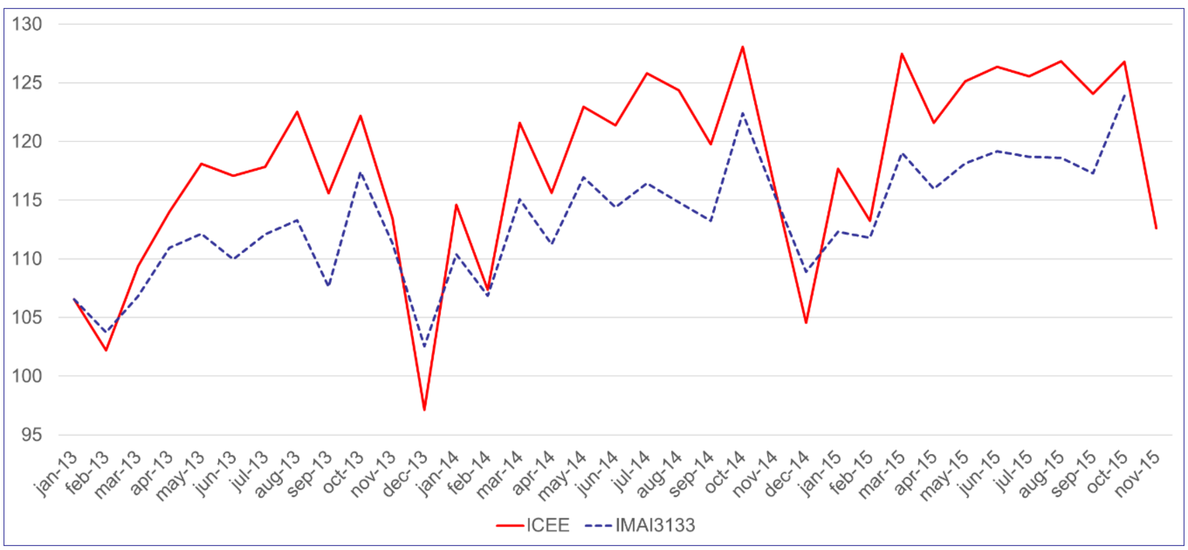

Fig. 1. Explanatory variable X_t: ICEE (solid line); response variable Y_t: IMAI3133 (dashed line), back in late 2015. Linear correlation coefficient between X_t, Y_t : 0.93.

3. Results from Modeling design

After conducting all hypothesis tests and assessments of empirical forecasting accuracy, **model 3 (Logarithmic Differences) was selected as the working linear model**. Model 3 passes all statistical tests, it is a parsimonious model, produces the most accurate forecasts, and has shown to be the most robust model when adding additional training data.

As of April 2020, the working linear model 3 has the following form:

$$\Delta\, \boldsymbol{ln}(\boldsymbol{Y_t}) = \boldsymbol{\beta_1}\Delta\, \boldsymbol{ln}(\boldsymbol{X_t}) + \boldsymbol{\beta_2 i_{oct}} + \boldsymbol{\beta_3 i_{apr}} + \boldsymbol{\varepsilon_t},$$

$$\boldsymbol{\varepsilon_t} = \boldsymbol{\rho\varepsilon_{t-1}} + \boldsymbol{v_t}; \qquad \boldsymbol{v_t} \sim \boldsymbol{i.i.d.}\ \ \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma^2})$$

Note that **model 3 estimates monthly growth rates** instead of IMAI3133 levels. Residual term $\boldsymbol{\varepsilon_t}$ for this model has a 1st order autoregressive structure; the **Cochrane-Orcutt** autoregressive correcting procedure is used to estimate model parameters. Throughout the August 2015 – April 2020 period, seasonal indicator variables $\boldsymbol{i_{oct}}$, $\boldsymbol{i_{apr}}$ have been incorporated to model 3; however, the effect of explanatory variable $\Delta\, \boldsymbol{ln}(\boldsymbol{X_t})$ is 11 times larger than the effect of $\boldsymbol{i_{oct}}$ and 33 times larger than the effect of $\boldsymbol{i_{apr}}$ on response variable $\Delta\, \boldsymbol{ln}(\boldsymbol{Y_t})$. All regressor variables are significant, $\widehat{\boldsymbol{\rho}} \cong -0.5$, and adjusted $R^2 \cong 0.9$.

4. Nowcasts: real time assessment

In this project, nowcast is a forecast corresponding to reference month N, generated by the working model during month $N + 1$. The official IMAI3133 value for reference month N is published by National Accounts during month $N + 2$, as a preliminary value. Fig. 2 shows graphically the historical comparison between nowcasts and official IMAI3133 values.

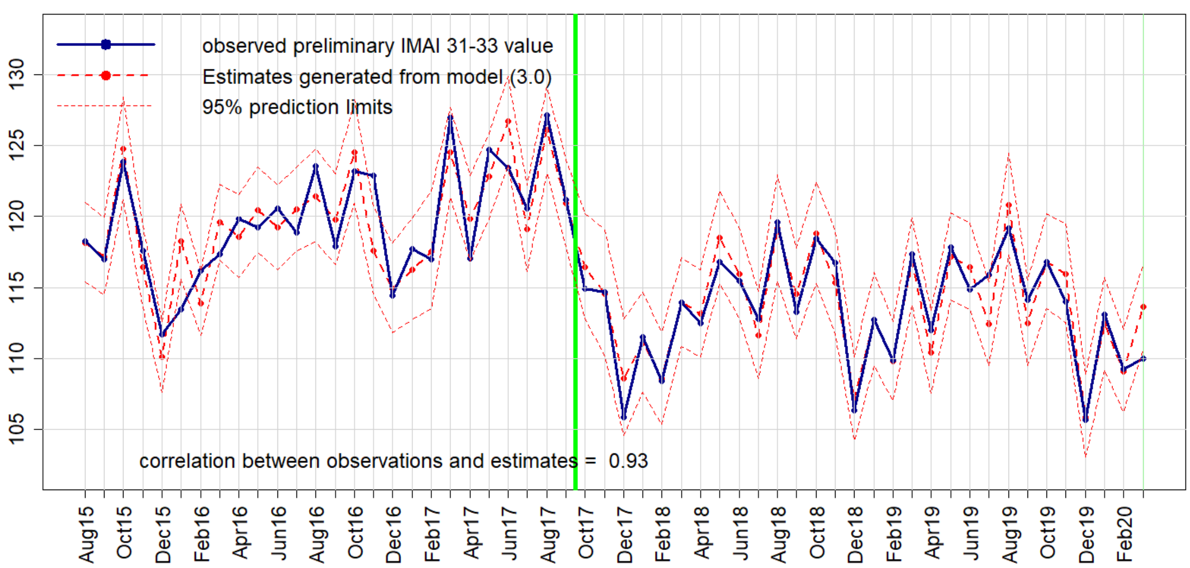

Fig. 2. Officially published IMAI3133 values (solid line); model 3 nowcasts (dashed lines), generated in real time before official IMAI3133 publication. Nowcasts are computed as prediction intervals with a 95% confidence level. IMAI3133 level nowcasts are recovered, via an exponential transformation, from monthly growth rate forecasts generated by model 3. Vertical bold line between September 2017 and October 2017 indicates a change of base

year, from 2008 to 2013, effected by National Accounts. This change of base year does not affect nowcasting methodology.

4.1 Nowcast for April 2020

	Lower Limit	Point Estimate	Upper Limit
IMAI3133	81.80	85.11	88.55
IMAI3133 annual growth rate	-31.69%	-27.73%	-23.76%

This nowcast was computed in May 2020. The official IMAI3133 figure for April 2020 will be published in June 2020.

5. Concluding remarks

From Fig. 2, it can be observed that 91% of the time (51 out of 56 months), the official IMAI3133 value is located inside the prediction interval, which was computed with a 95% confidence level; this means that, in this case, observed empirical accuracy approached the theoretical confidence level. This is empirical evidence in favor of the structural stability of model 3.0.

Overall, although there is not a formal Memorandum of Understanding (MoU) between CFE and INEGI, there have been no months during the realization of this project in which no data has been received from CFE; this has enabled the successful realization of the project. There is work in progress to establish a formal MoU between CFE and INEGI.

These IMAI3133 nowcasts are now being communicated publicly, since May 2020, as experimental statistics at the INEGI internet site.

References

1 Sistema de Cuentas Nacionales de México. Cuentas de Corto Plazo y Regionales: Fuentes Metodológicas. 20 de Agosto de 2013. INEGI. Link (in Spanish, accessed on August 14, 2019): https://www.inegi.org.mx/contenidos/programas/imai/2008/metodologias/SCNM_Metodologia_06.pdf

2 Davidson, R., & MacKinnon, J. G. (1993). Estimation and inference in econometrics. OUP Catalogue. Oxford University Press, number 9780195060119

3 Kutner, M. H.; Nachtsheim, C. J.; Neter, J. (2004). Applied Linear Regression Models (4th ed.). McGraw-Hill Irwin

4 D. Cochrane & G.H. Orcutt (1949). Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms. Journal of the American Statistical Association. Volume 44, Issue 245: pages 32–61.