Merging mandatory communications and real time information to forecast LM dynamics

Nowcasting employment

Massimiliano Tancioni
Sapienza Università di Roma

Tesi Aliaj - Paolo D'Imperio - Federica Vassalli
Sapienza Università di Roma

Plan of the presentation

- Premises and goals of the project
- Methodological background
- Labor market dynamics and mandatory communications (MC)
- Proposed methodology
- Predictive capabilities of MCs: a tentative application

Premises

- Forecasting employment dynamics (possibly at a disaggregated level) is a major concern for policy makers and welfare institutions. Gains in policy implementation, management costs and effectivenes can be exploited with very short term forecasts of LM trends
- Official statistical information about the LM is released with a substantial delay (Istat: about two months and a half for quarterly LM statistics; about one month for aggregate employed and unemployed monthly provisional data)
- Merging these official releases with an high dimensional set of possibly high frequency data can help estimation of the current LM dynamics and understanding of the underlying developments
- MCs could be included in the set of f observables to improve LM predictability

Objectives

- The objective of this project is to merge MCs data with other high and standard frequency data in a VAR model to provide very high frequency forecasts (nowcasting) of specific segments of employment
- Two major methodological issues arise in dealing with large data sets composed of informarion observed at different frequencies:
  1. Solving the curse of dimensionality problem which emerges because of the n.of coefficients in VARs increasing with the square of the n. of variables
  2. Merging data observed at different frequencies, i.e. quarterly, monthly, daily and, possibly, real time data

Methodological background: dealing with big data

- Two major approaches
  1. Factor-augmented VARs (FAVAR): Factor models reduce the dimensionality issue by extracting the information about the (few) unobserved common factors driving the VAR variables. The largest the information, the more likely the identification of common factors ("dimensionality blessing", Marcellino, 2017): Doz, Giannone and Reichlin (2012-REStat) apply these mehodologies and provide the ML estimator for large approximate Dinamic Factor Models (DFM)
  2. Large-dimensional Bayesian VARs (BVAR): BVARs exploit prior restrictions (generally Minnesota-like) to shrink the VAR parameterization: Banbura, Giannone and Reichlin (2010-JApplEconometrics); Carriero, Kapetanios and Marcellino (2012-JBankFin); Koop (2013-JApplEconometrics)

Methodological background: dealing with heterogeneous frequencies

- Giannone, Reichlin and Small (2008-JME) and Angelini, Banbura and Runstler (2007-ECB WP) deal with data released at different frequencies, by using monthly data to forecast quarterly GDP
- Banbura and Modugno (2012-JApplEconometrics) propose a modified Expectation Maximization (EM) algorithm for missing observations to generalize Doz, Giannone and Reichlin (2012-REStat)'s DFM approach to the case of large samples characterized by heterogeneous frequencies and varying release dates. They implement the procedure to nowcast GDP using FAVARs. Modugno (2013-IntJForec) apply this methodology to nowcast inflation.

Labor market dynamics and MCs: theoretical relations

- Employment law of motion in the DMP model:

$$n_t = (1 - \rho_t)\, n_{t-1} + m_t$$

where n_t denotes employment t, ρ_t is the separation rate (retirement, firing, etc.) and m_t denotes new matches

- Cobb-Douglas matching function:

$$m_t = \sigma_m v_t^{\sigma_n} u_t^{1-\sigma_n}$$

where σ_m denotes the matching efficiency parameter, v_t are vacancies (labor demand), $u_t = 1 - n_t$ is the unemployment rate (labor supply, once LF is normalized to one), σ_n is the CD share parameter

Labor market dynamics and MCs: observables

- The employment law of motion has the following sample counterpart

$$n_t^{obs} - n_{t-1}^{obs} = \Delta n_t^{obs} = m_t^{obs} - sep_t^{obs}$$

where $sep_t^{obs} = \rho_t n_{t-1}$ are the separations observed in time t and m_t^{obs} are the time t observed matches, both provided by MCs. Among the variables included in the matching function, only v_t is not provided by official sources (Google search?), such that labor demand could be implicitly identified.

- Basically, MCs could provide high frequency (real time?) information to produce high frequency forecasts of employment
- Given its highly detailed structure, forecasts can in principle be obtained for specific segments of interest of employment/unemployment

Proposed methodology: DFM basics

We adopt the methodology proposed in Banbura and Modugno (2012), based an approximate DFM, to produce forecasts of employment observed by official LM statistics at the quarterly frequencies. The standard model representation is as follows:

$$\mathbf{y}_t = \mathbf{C}\mathbf{f}_t + \varepsilon_t, \quad \varepsilon_t \sim iid\left(\mathbf{0}, {}^{\circ}\right)$$

where $\mathbf{y}_t$ is the $m \times 1$ vector of observables, $\mathbf{C}$ is an $m \times r$ matrix of loadings and $\mathbf{f}_t$ is an $r \times 1$ vector of unobserved factor components, for which a VAR structure is assumed, i.e.:

$$\mathbf{A}(L)\mathbf{f}_t = \mathbf{u}_t, \quad \mathbf{u}_t \sim iid\left(\mathbf{0}, \mathbf{Q}\right)$$

DFM mplementation

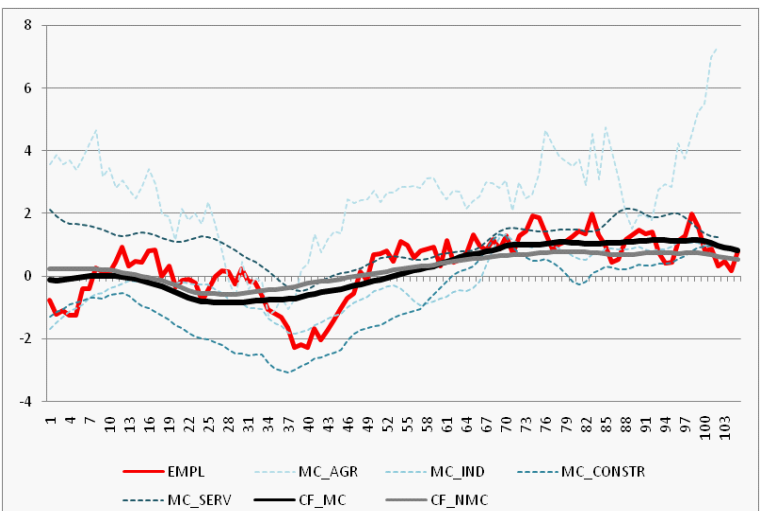
- We consider a set of 22 monthly and quarterly time series observed over the sample 2009-2018. Following Bock, Caratelli, Giannone, Sbordone and Tambalotti (2017), we assume a factor grouping composed of four blocks, namely:
  1. Global block, which is shared by all 22 monthly and quarterly variables
  2. Real block, specific to real variables only (monthly observations: industrial production index, economic sentiment index, energy production index; quarterly observations: value added in the four macro-sectors)
  3. Price block, specific to nominal variables (monthly observations only: HICP, oil prices, the real exchange rate, interest rates on 10 years government bonds and on the short term lending rate to non financial corporations)
  4. Labor block, specific to labor market variables (monthly observations: the istat's aggregate employment stock monthly estimate, the employment index obtained from recorded hirings and firings in MCs in the four macro-sectors; quarterly: the aggregate employment stock and the employment stocks in the foru macro-sectors from the istat's labor force survey

Evaluation strategy

- We simulate an high frequency (actually one month) forecasting environment by considering repeated forecasts from 2015q1 to 2018q4 of quarterly employment, in which forecasters do and do not consider MC information
- In a first step, we estimate the model over the different vintages of samples including MC information to perform one-step-ahead forecasts targeting the quarterly employment stock
- In a second step, we estimate the model over different vintages of samples excluding the MCs-based data, and repeat the one-step-ahead forecasts
- We then compare the forecasting performances of both model specifications to the true values over time and to forecasts obtained with a benchmark (ARIMA) forecasting device

Results: Common Factor - Global

Common factor, MC data and aggregate employment, monthly



Results: Predictive capabilities and model comparison

Predictive capabilities and models comparison - 2015q1 2018q4

Time	15:1	15:2	15:3	15:4	16:1	16:2	16:3	16:4	17:1	17:2	17:3	17:4	18:1	18:2	18:3	18:4	all
Data	0.70	0.70	1.03	0.93	1.15	1.80	1.04	1.21	1.52	0.57	1.34	1.30	0.72	1.59	0.65	0.43	-
M-MC	0.72	0.61	1.11	1.00	1.04	1.83	0.79	1.28	1.49	0.78	1.64	1.28	1.09	1.62	0.86	0.47	-
R(M)SFE	0.03	0.09	0.08	0.08	0.11	0.02	0.25	0.06	0.02	0.21	0.30	0.02	0.36	0.02	0.22	0.04	0.12
M-NMC	0.49	0.51	0.91	0.69	1.01	1.76	0.60	1.23	1.32	0.31	1.47	1.03	0.60	1.57	0.28	0.55	-
R(M)SFE	0.20	0.18	0.12	0.24	0.14	0.04	0.44	0.02	0.20	0.26	0.14	0.27	0.12	0.02	0.37	0.12	0.18
ARIMA	0.38	0.73	0.3	1.14	0.34	1.48	1.96	0.13	1.29	0.82	0.56	1.09	0.36	1.19	0.91	0.39	-
FE	0.32	0.03	0.73	0.21	0.81	0.32	0.92	1.08	0.23	0.25	0.78	0.21	0.36	0.40	0.26	0.04	0.44