Inflation Prediction Using Scanner Data

Sonan Memon*

This version: July 2022

Abstract

Predicting inflation using scanner data from super markets.

Keywords: Inflation Prediction. Scanner Data From Retail Markets. Crisis of Hyperinflation. Russian and Ukraine Crisis. Machine Learning.
JEL Classification:

---

*Research Economist, PIDE, Islamabad. smemon@pide.org.pk

CONTENTS

1. INTRODUCTION

Accurate forecasting of inflation is a concern for both market players and central banks. On the one hand, market participants want to update their inflation expectations in line with new information revelation so that their investment strategies are optimal. Meanwhile, central banks typically have mandates for price stability and they routinely collect data on inflation expectations and forecasts.

While central banks collect data on consumer price indices, the frequency of such data does not allow accounting for sudden swings in expectations and their dramatic effects on inflation. Some examples of standard measures include the HICP (Harmonized Consumer Price Index) data used in the Euro area and the CPI (consumer price index) data from USA. Such data typically tends to be quarterly in worst cases or in best cases monthly, but results are revealed in the next month. However, when for instance, in a matter of few days and weeks, news about the Ukraine and Russian crisis changed the inflation expectations of many products, conventional price indices had little forecasting potential. The Ukraine and Russian crisis led to unexpected, sudden hyperinflation shock in many countries and similar inflation shocks can emerge from sudden change of State Bank governors or governments, terrorism episodes or political turmoil, especially in developing economies, where inflation tends to more volatile (see for instance Vuletin and Zhu (2011)).

The State Bank of Pakistan (SBP) has also done some work on inflation forecasting by using neural network type machine learning methods and monthly year on year (YoY) inflation rate of Pakistan from Jan 1958 to Dec 2017 Hanif et al. (2018). Similarly, the SBP has worked on *nowcasting* GDP using large scale manufacturing growth (LSM) in Pakistan Hussain et al. (2018) and LASSO type[1] ML methods. However, lack of high frequency data on the order of days or weeks poses a limitation in forecasting inflation using these methods. Hence, I argue that we need more granular data, on the order of days or weeks for enhancing forecasting. Next, I discuss methods for collecting such high frequency data, used at the current frontier of research on inflation.

https://www.centralbanking.com/central-banks/economics/data/3315546/big-data-in-central-banks-2017-survey

In recent literature, the daily consumer price index (CPI) produced by the Billion Prices Project (BPP CPI) of Cavallo and Rigobon (2016) offers a glimpse of the direction taken by consumer price inflation in *real time*. For instance, Figure 1 is based on web scrapping

---

[1]Least Absolute Shrinkage Operator, Ridge Regressions and Elastic Nets.

online inflation data for Argentina Cavallo and Rigobon (2016). It shows that the official CPI significantly under-stated actual inflation, when measured by web scrapping. An added benefit of such data is that it reveals the partisan measurement and particularly disclosure of CPI data in developing economies such as Argentina, where central bank independence is low. Should we expect a similar lack of correspondence between official inflation data of the SBP (State Bank of Pakistan) and non-partisan research measures?
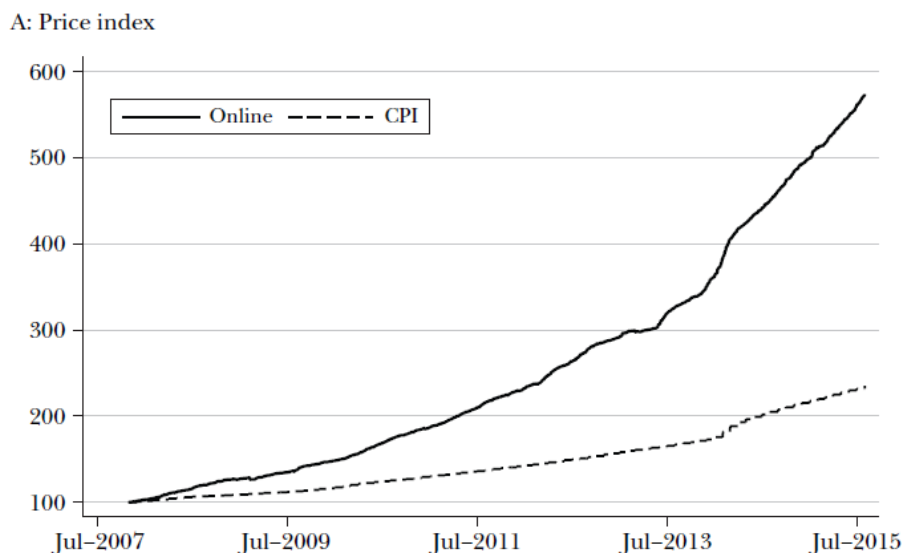

Figure 1: Inflation in Argentina (Source is Cavallo and Rigobon (2016))

Meanwhile, another branch of emerging literature uses scanner-based data (see for instance Beck et al. (2020)) on prices rather than web scrapping. In Figure 2 below, recent scanner-based price indices for Germany are disclosed from the work of Beck et al. (2022). The data compares trajectories in 2022 (red and orange solid lines below) with their historical averages from 2019 to 2021 (blue and purple solid lines) along with historic minimum and maximum values (shaded areas). The data indicates a very strong increase in prices for sunflower oil and flour in light of the Ukraine conflict, accompanied by temporarily higher sales. The price increase of sunflower oil was rather gradual and already started as of early February. In contrast, prices for flour increased very sharply, but only more than two months after the invasion. However, in both cases, sales went far beyond their average levels, suggesting increased demand and possibly stockpiling behavior from pessimistic consumers (see Cavallo and Kryvtsov (2021)). Concerning the more recent period up to June 2022, prices for both products seemed to have stabilized at a very high level, whereas quantities have converged back to their average levels.
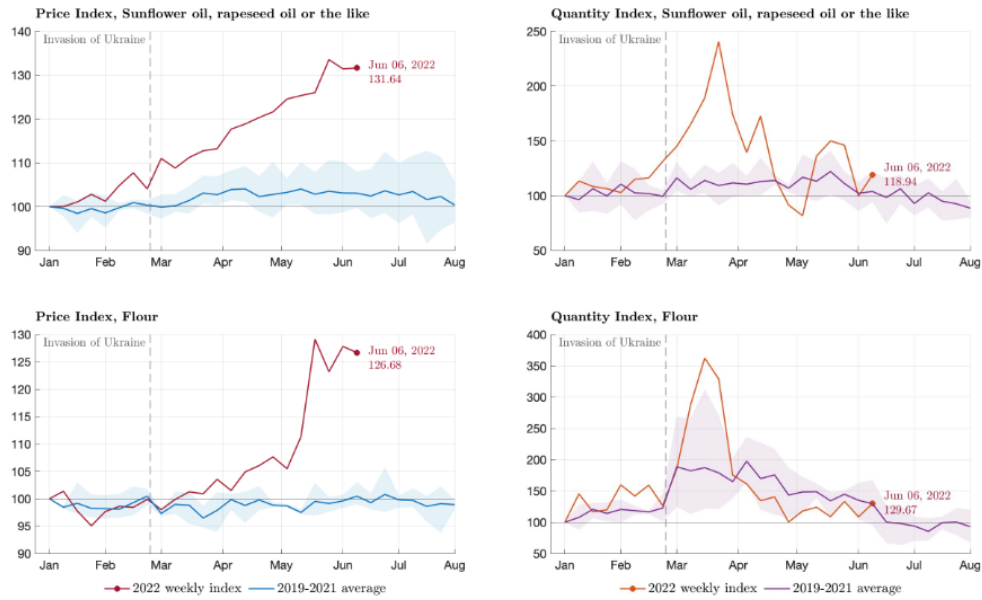
Figure 2: Source is Beck et al. (2022)

2. MORE ON SCANNER DATA

Modugno (2013).

Beck et al. (2020).

3. ML and Prediction

4. CONCLUSION

REFERENCES

Beck, Guenter W, Hans-Helmut Kotz, and Natalia Zabelina, "Price gaps at the border: Evidence from multi-country household scanner data," *Journal of International Economics*, 2020, *127*, 103368.

Beck, Guenter W., Kai Carstensen, and Jan-Oliver Menz, "Real-time food price inflation in Germany in light of the Russian invasion of Ukraine," *VOXEU*, 2022.

Cavallo, Alberto and Oleksiy Kryvtsov, "What can stockouts tell us about inflation? evidence from online micro data," Technical Report, National Bureau of Economic Research 2021.

___ **and Roberto Rigobon**, "The billion prices project: Using online prices for measurement and research," *Journal of Economic Perspectives*, 2016, *30* (2), 151–78.

Hanif, Muhammad Nadim, Khurrum S Mughal, and Javed Iqbal, "A Thick ANN Model for Forecasting Inflation," Technical Report, State Bank of Pakistan, Research Department 2018.

Hussain, Fida, Kalim Hyder, and Muhammad Rehman, "Nowcasting LSM growth in Pakistan," *State Bank of Pakistan*, 2018, p. 1.

Modugno, Michele, "Now-casting inflation using high frequency data," *International Journal of Forecasting*, 2013, *29* (4), 664–675.

Vuletin, Guillermo and Ling Zhu, "Replacing a "disobedient" central bank governor with a "docile" one: A novel measure of central bank independence and its effect on inflation," *Journal of Money, Credit and Banking*, 2011, *43* (6), 1185–1215.