Policy Research Working Paper 9886

Estimating Food Price Inflation from Partial Surveys

Bo Pieter Johannes Andrée



Fragility, Conflict and Violence Global Theme December 2021

Policy Research Working Paper 9886

Abstract

The traditional consumer price index is often produced at an aggregate level, using data from few, highly urbanized, areas. As such, it poorly describes price trends in rural or poverty-stricken areas, where large populations may reside in fragile situations. Traditional price data collection also follows a deliberate sampling and measurement process that is not well suited for monitoring during crisis situations, when price stability may deteriorate rapidly. To gain real-time insights beyond what can be formally measured by traditional methods, this paper develops a machine-learning approach for imputation of ongoing subnational price surveys. The aim is to monitor inflation at the market level, relying only on incomplete and intermittent survey data. The capabilities are highlighted using World Food

Programme surveys in 25 fragile and conflict-affected countries where real-time monthly food price data are not publicly available from official sources. The results are made available as a data set that covers more than 1200 markets and 43 food types. The local statistics provide a new granular view on important inflation events, including the World Food Price Crisis of 2007–08 and the surge in global inflation following the 2020 pandemic. The paper finds that imputations often achieve accuracy similar to direct measurement of prices. The estimates may provide new opportunities to investigate local price dynamics in markets where prices are sensitive to localized shocks and traditional data are not available.

This paper is a product of the Development Data Group, Development Economics and the Fragility, Conflict and Violence Global Theme. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The author may be contacted at bandree@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Estimating Food Price Inflation from Partial Surveys

By Bo Pieter Johannes Andrée*

JEL: C01, C14, C25, C53, O10.

Keywords: Inflation, Food Security, Financial Stability, Machine

Learning.

^{*} The World Bank, Development Economics, Data Group, Analytics and Tools Unit. The author may be contacted at bandree(at)worldbank.org. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. This work is part of the program "Building the Evidence on Protracted Forced Displacement: A Multi-Stakeholder Partnership". The program is funded by UK aid, it is managed by the World Bank Group (WBG) and was established in partnership with the United Nations High Commissioner for Refugees (UNHCR). The scope of the program is to expand the global knowledge on forced displacement by funding quality research and disseminating results for the use of practitioners and policy makers. This work does not necessarily reflect the views of the UK government, the WBG or UNHCR. The author would like to thank Paolo Verme, Aart Kraay, Phoebe Spencer, Olivier Dupriez, Kamwoo Lee, Andres Chamorro, Nadia Piffaretti, Benjamin Stewart, Roberta Gatti, Daniel Lederman, and Talip Kilic for detailed feedback, inputs and rich discussions on early results.

Low inflation and price stability have historically been associated with improved growth and development outcomes (The World Bank, 2019; Ha et al., 2019b,a). Inflation dis-proportionally taxes the purchasing power of low incomes (Easterly and Fischer, 2001), which is in contrast with the proportional income effects of economic growth itself (Dollar and Kraay, 2002). As such, price stability and low inflation are crucial to achieve the Sustainable Development Goals and price data are critical for economic monitoring.

Food prices specifically play a critical role in determining the immediate ability of households to consume food (Andreyeva et al., 2010; Waterlander et al., 2019), and the capacity of farmers to plan the investments needed for reliable future production of food (Koomen et al., 2015; Andrée et al., 2017; Diogo et al., 2017). Food price inflation thus serves as an important metric to inform economic policy and is closely watched by both economists and humanitarians during times of financial turmoil and economic crisis.¹

The traditional figures on inflation used for these purposes have primarily been produced at an aggregate level using data from few, highly urbanized, areas. As such they do not directly describe price evolution in rural or poverty-stricken areas, or in regions where vast populations are internally displaced and live in fragile situations.² For example, food crises are often characterized by great spatial heterogeneity and geographic specificity (Maxwell et al., 2020), while traditional price indicators do not provide insights beyond the major (urban) markets where prices are formally measured.³ Traditional indicators are also often produced with delay, or not at all, particularly during crises when economic variables deteriorate rapidly.⁴

Recognizing these limitations, international institutions have invested heavily

¹The significance of price inflation as a signal of deteriorating food security can be highlighted by noting that the recent crises in South Sudan and the Republic of Yemen are characterized by high inflation. Historically, the Bengal famine of 1943 followed after a period of hyper-inflation, starvation in the Weimar Republic occurred alongside record hyper-inflation, the 1980 Uganda famine occurred after double to triple digit inflation rates, the 1992 famine in Southern Somalia was preceded by a year of triple-digit inflation, and the 1998 famine in southern Sudan was preceded by years of double digit inflation. The use of food price indicators to track famine risks is documented more substantively by Seaman and Holt (1980); Cutler (1984); Khan (1994); Andrée et al. (2020); Wang et al. (2020).

²Naturally, there are other critiques around inflation calculations, some as old as the methodologies themselves. Some accusations about both up- and downward biases are discussed by Reinsdorf et al. (2009). Often, conflicting views on inflation can be traced back to differences in (baskets of) goods and data sampling locations, and so for some applications improved clarity can be gained by more narrowly defining indexes, which the results of the paper may also help enable.

³See also the recent economic update for South Sudan (World Bank, 2021). The work compares food prices subnationally and finds that increasing prices are the most significant factor driving recent food insecurity but documents strong spatial heterogeneity in both market price dynamics, as well as relations to food insecurity, and provides examples of markets where prices are more sensitive to localized shocks.

⁴For instance, the International Finance Statistics (IFS) data base of the IMF reports monthly price data at Consumer Price Index (CPI) component level, but few developing countries report food price data without substantial delays. Of the 25 countries analyzed by the paper, none presently reports real-time monthly official food price data; half had not reported food price data going back 12 months. Annual statistics are similarly far from complete. Of all countries in the World Bank's WDI, only 60% (40% of the analyzed countries) had an inflation figure for the most recent 2019–2020 period. The IFS data reported similarly on only half of the countries. This was last checked on August 31, 2021; WDI data identifier FP.CPI.TOTL.ZG, and IFS data identifier PCPI_PC_PP_PT.

in subnational price survey systems. However, the capacity to monitor inflation continues to be severely limited by challenges related to data gathering that include such issues as lack of resources, loss of access to markets, or disturbances in ground operations during crises and conflict outbreaks. Despite tremendous effort, missing data remains pervasive, which poses problems when one wants to compute real-time inflation measured over multiple price series. Whenever a single price point is absent, the required index cannot be computed without making assumptions about missing data.

To overcome some of the shortcomings associated with traditional systems, this paper develops an approach for real-time imputation of survey data drawing on multiple imputation and ensemble learning ideas. The imputation is generated as an ensemble prediction from multiple completed price trajectories. Each price trajectory is the result of a stochastic equation chain of machine learning models that leverage correlations between prices of individual food items and similarities between markets to predict missing local price quotes. The idea is similar to a bootstrap aggregation (bag) of base learners (Breiman, 1996), only now randomization in base learners is not achieved by random sampling but from a stochastic simulation of chained model predictions. By augmenting incomplete and intermittent survey data with reliable predictions, subnational food price trends can be monitored continuously.

The paper highlights the new price monitoring capabilities using surveys from the World Food Programme (WFP) gathered in 25 fragile and conflict-affected countries. The final estimates of food price inflation documented in this paper are shown to capture important local inflation events related to conflict; as well as known food crises including the World Food Price Crisis of 2007-08 documented by Baffes et al. (2008); Conceição and Mendoza (2009), and the surge in inflation that followed the 2020 pandemic and subsequent expansion in the global monetary base. The results include inflation estimates for a large number of countries where these data have recently not been available, and the analysis of the results documents new insights into different characteristics of recent and past inflation events.

The granular high-frequency results provide an alternative price estimate that is particularly relevant in data-poor, lower-income, regions, where the capacity to maintain in-depth price monitoring programs, that rely on traditional CPI methods, is often limited. The focus on low-income countries is in contrast with past work that, instead, has had a stronger focus on data-rich and higher income areas. For example, past related work has developed forecasting methods for future CPI (Joutz, 1997; Gavin and Mandal, 2002), future commodity spot and futures prices at global markets (Ahumada and Cornejo, 2016; Ouyang et al., 2019), or methods for continuously updating current expectations of future country-level CPI inflation using information from alternative sources (Modugno, 2011; Seabold and Coppola, 2015). Finally, the methods proposed by the paper could also be applied to enhance other data gathering programs, improve their cost-effectiveness

by replacing a subset of expensive surveys with targeted predictions, and advance broader economic monitoring in data-poor regions.

The remainder of the paper has been structured as follows. The next section provides a brief overview of the type of price surveys that can be used to deploy the new monitoring capabilities and discusses some basic challenges encountered when trying to read inflation from the raw data. Section II details the imputation strategy and section III presents results. For those interested in analyzing the subnational price and inflation estimates further, all results for the 25 countries analyzed are available to be interactively explored.⁵ Section IV provides concluding remarks and makes several recommendations for future data gathering.

I. Food Price Survey Data

Subnational food prices have been surveyed in many countries for years by humanitarians to inform their country operations. Well-known data bases are those from the WFP, FEWS NET and the Food and Agricultural Organization (FAO).⁶ The paper focuses on raw monthly data from the WFP, but parts of the discussion, and particularly the methods developed here, could apply to similar data sets.⁷

The paper gathered all end-of-August data available from the WFP Vulnerability Analysis and Mapping (VAM) unit as of September 21, 2021. The WFP data reports prices as measured in different market locations throughout each country. Price quotes are supplied by WFP country and regional offices, and local partners including the FAO. In total, the data base covers over 2,000 markets across 99 countries, and reports monthly data on the prices of goods that vary by country and market. Price monitoring dates back to early/mid 2000s in most countries, while in few countries monitoring began already in the 1990s.

The number of markets varies strongly by country, ranging from 1 to well over a 100. The methods developed by the paper are particularly interesting when there are multiple markets, as otherwise, official CPI data may provide sufficient, if not better, insight. The goods for which prices are collected, and the manner in which these prices are measured, in turn vary widely across markets and may change over time. For example, prices of some goods are collected at the retail level in some markets, while for others they are collected at the wholesale level and include discounts typical for large purchases. Generally, prices are measured in nominal local currency per unit of commodity (e.g. Shillings per 10 kg of maize), but in some countries USD quotes on selected goods may exist alongside

⁵With the publication of the paper, new monthly food price estimates by product and market using the methods and data described in this paper, will be made available on the World Bank Microdata Library. They can be viewed and downloaded at https://doi.org/10.48529/2ZH0-JF55 or, currently, https://microdata.worldbank.org/index.php/catalog/4218. The citation is Andrée (2021).

⁶There are also country-specific data gathered by FSNAU (Somalia) and CLiMIS (South Sudan).

⁶There are also country-specific data gathered by FSNAU (Somalia) and CLiMIS (South Sudan). More recently, the International Food Policy Research Institute has piloted gathering high frequency prices in several countries.

⁷For instance, estimates have also been developed for Papua New Guinea in collaboration with IFPRI in a separate pilot project. Results can be requested from the author.

local currency quotes. As a result, the full data set is highly heterogeneous and challenging to work with. For example, while it contains common staples like rice, sorghum, and maize, that are monitored in many countries, there are also items specific only to one country.

The focus of the paper is purely on monitoring food price inflation, in a reasonably cross-comparable manner.⁸ The strategy has been to extract from the very large set of price data set a stable set of commodity prices that are defined as homogeneously as possible across countries, and that are as widely available as possible across markets, while having the best coverage over time. The period of analysis starts in January 2007, or the next first date at which data was available. After carefully examining the prevalence of price data across commodities, markets, and time for each country, 43 foods were identified for which price data are reasonably abundant across multiple markets in at least one country. Tables A2 and A3 list the selected food items by country. The foods are either staples, agricultural produce, or dairy products. 9 Aside from non-foods, the selection excludes only fish and meat products due to the very high dietary heterogeneity in these foods. The resulting country-specific baskets thus consists mostly of staples, often similar in nature as the type of foods traded at global markets. For example, maize, sorghum, millet, wheat, vegetable oil, to name a few common food items, are also tracked in the World Bank Commodities Price Data (The Pink Sheet) that is used to construct the World Bank Food Price Index used to track international food price developments. In most cases, the different food items can substitute one another to certain degrees, or may act as complements, so that the prices of most foods will be strongly associated with the price developments in others. For example, in Afghanistan there are only four food items: bread, rice, wheat and wheat flour. These prices are likely strongly associated. First, rice and wheat are both cereals and likely to move together in price, as either will replace the other as sources of carbohydrates to a certain degree if their price ratios diverge too strongly. The price of wheat is likely a good predictor of the price of wheat flour, which in turn is used to make bread. Thus, a regression chain, that cycles over the individual food items, can make good imputations for all products as soon as the price of one of these goods is available. When more food items are observed, the chances increase that for each item with a missing price quote, at

⁸In several countries, the data covers non-food items that include common household items, prevailing day labor wages and local exchange rates, but generally the availability of such data is highly specific to the selected area. The proposed statistical methods, however, will not necessarily discriminate between types of commodities and could be applied to generate country-specific results on non-foods.

⁹The items covered are bread, cooking oil, pulses, rice, sugar, wheat, wheat flour, beans, groundnuts, maize, millet, sorghum, bananas, cassava flour, maize flour, onions, potatoes, tomatoes, cassava, cocoyam, cowpeas, eggs, milk, peas, plantains, sesame, cabbage, carrots, cucumbers, dates, fonio, garlic, oranges, tomatoes paste, watermelons, yogurt, bulgur, cheese, chickpeas, lentils, maize meal, gari, and parsley. The selection of these items has been fully codified in a set of flexible rules around minimum data-availability thresholds. The selected items in this way depend on the period of analysis. Generally, if the focus is on recent data, without requiring a long history, a higher number of food items could be selected. The code to scrape and run the price monitor has been made available as open source and can be customized for improved country-specific results.

 $6 \hspace{1cm} ANDR\'E \hspace{1cm} DECEMBER$ - 2021

least one other item can be found that carries very strong predictive power over its price. 10

Table A1 summarizes the raw data availability in each country. The table shows that the data of the analyzed countries on average include 7 food items measured across 27 markets, with around half of the survey data missing. The number of markets is generally a strong improvement in geographic detail when compared to traditional price indicators. For example, World Bank (2021) analyzes official CPI at the subnational component level in South Sudan. The analysis highlights the limits of relying purely on primary data collection, as the indexes are only compiled for the three major urban areas and not for any rural areas. This is despite an estimated 80% of the country's population that lives in rural areas, and the fact that only a fraction of the remainder urban population buys goods at these three markets. Many urban households rely on markets that are less connected to international markets than primary markets in the capital city.

In total, the application uses data from 676 different markets. There are an additional 547 markets without observations, which are markets where WFP tracks commodities that are not covered by the application. These locations are not modeled but are spatial interpolated separately. Table A1 provides a country breakdown of data availability, for the country-product combinations analyzed by the paper. Data coverage of around 70%, as a percent of all $time \times location$ with data, is relatively high, while 30% can be considered relatively low. Naturally, these figures depend on the selection of markets, food items, and the period of analysis. Higher data coverage could be achieved by focusing on a shorter time period, working with fewer markets, or tracking a very narrow basket. The statistics are thus not representative of the WFP data base, but represent the data selection of the paper which seeks to balance reasonable data coverage and data availability.

The raw price data of these selected food items remains a challenging source to deal with. Available data is regularly contaminated by outliers, which could be due to incorrect survey entries (e.g. a misplaced digit). Moreover, many local price series are incomplete and price quotes of different products may become available at different times and locations. The data availability constraints become more problematic when the interest is in tracking inflation, which requires prices quotes to be matched with historical quotes using a fixed time interval. For any given country, if the data selection was made such that the data coverage was 100%, the selection would collapse to zero cases unless the focus would be on an extremely narrow selection of data points that introduces strong sampling

¹⁰There may also be nonlinear patterns that carry important information about unobserved prices. First, a change in price ratio between two prices may be an indicator that a certain food of interest will likely trade against an elevated price ratio to some other food. Second, governments may use price controls and so certain prices may remain fixed for long periods. This can for example be observed in the bread prices in Afghanistan here (Andrée, 2021). These fixed prices regularly break once input prices have risen too sharply, and such level shifts in the data can signal that all price levels must seek new price equilibria. To that regard, there may also be time-varying relations between various prices that, for example, change depending on key level shifts in the prices of certain foods.

biases. Such a selection can also only be made *ex-post* and so it is only useful for historical analysis of prices and not for real-time monitoring.

II. Methods

To overcome some of the shortcomings associated with the raw data, this section develops an approach for real-time imputation. The idea is to create an ensemble prediction from multiple completed market-level price trajectories, produced by simulating stochastic equation chains of predictive machine learning models that leverage correlations between prices of individual food items and similarities between markets. The completed data can then be used to construct regional or local food price indexes.

A. The missing data problem

First, table 1 visualizes the general missing data problem.

Table 1—Example of the missing data problem.

 b_6 c_6

Note: Example of the missing data problem, three hypothetical vectors A, B, and C that represent price series, with elements a_t , b_t and c_t being individual price quotes indexed by time periods t. Blank entries represent missing observations. The challenge is to estimate change rates ΔP of the basket price vector P = A + B + C that spans all $t = 1, \ldots, 6$. Element b_4^* is an example outlier price which needs to be removed and replaced with an estimate.

Source: Example has been prepared by the author for this paper.

The literature has put forward many solutions for time series interpolation and extrapolation. It is useful to discuss briefly why some standard tools cannot reliably be used. Simple data gaps in univariate time series can often be linearly interpolated or solved with last-observation-carried-forward imputation. Both approaches can be sufficiently adequate if there is just one observation missing in a sequence. When several values are missing in a row, the results might rapidly become unrealistic. For example, a_4 can only be carried forward since a_5 and a_6 are missing. This introduces a lag with which price increases are observed. When prices are rising, a last-observation-carried-forward imputation introduces a downward bias in the price average, for example using this method we would approximate $\hat{p}_6 \sim a_4 + b_6 + c_6 << a_6 + b_6 + c_6$ when $a_6 >> a_4$. As shall become clear from the results, price levels in low-income countries can move fast in either

direction so that old price data quickly loses its relevance. 11

Kalman filtering is a popular method that is known to produce results that are optimal in common settings. These methods are discussed at length by Durbin and Koopman (2013). The state of the methodologies at present, however, is such that the current application brings together too many issues to derive a state space specification that generalizes across all data situations. Essentially, a good approach may be specified for narrow studies of prices within any given country individually, but likely not across all prices and all countries.¹²

B. The predictive imputation framework

The literature on missing data often distinguishes between different mechanisms of missing data: MCAR (missing completely at random), MAR (missing at random), and MNAR (missing not at random). This taxonomy, introduced by Rubin (1976) and well-explained by van Buuren (2012), becomes important when missing data occurs at the covariate side, and different corrections are needed for unbiased inference depending on the type of missing data problem. In the current application, it is required that the data is at least MAR and possibly MNAR and that the value of one food item can be accurately predicted by making reference to prices of other food items and the location in space and time. ¹³ As such, the

 11 Closely related interpolation methods are plagued by similar issues. For example, linear interpolation takes information from the future to the past. The linear interpolation estimate for c_5 can only be made after c_6 has already been observed. Alternatively, moving average imputation based on past values results in biases when used for real-time monitoring. When prices are rising, a moving average imputation of a_6 would be based on an average of lower prices and a real-time monitoring system using moving average imputation would regularly suggest false trend reversals. These simple interpolation approaches are also highly influenced by outliers. Interpolating linearly from b_2 to b_6 , through outlier b_4^* , replicates information into two new outliers b_3^* and b_5^* . More flexible spline methods are known to introduce spurious cycles in time series data and may result in explosive extrapolation due to their quadratic or higher order nature, particularly in the presence of outliers.

¹²Optimality in a Kullback-Leibler or Root-Mean-Square sense is only under restrictive specification assumptions which require tedious design effort and diagnosing. Generally, Kalman filtering estimates time varying states and uses current state estimates and observations to estimate future states. When observations are missing, the imputation will rely fully on the transition equation. This means that the dynamic properties deteriorate rapidly during periods of large data gaps. This may in part be tackled by considering multivariate state space approaches so that realistic predictions for one sequence can be made by utilizing recent information from another sequence. Fast multivariate filtering implementations that could be deployed at scale are for example explored by Koopman and Durbin (2000). Due to outliers, and general non-Gaussian behavior, non-Gaussian outlier models may need to be considered. The univariate context is for example explored by Shephard (1994); Koopman et al. (2019). The above methods all rely on restrictive linearity assumptions. Nonlinear filtering is explored by Wan and Van Der Merwe (2000). Deciding between approaches is complicated by the variability in quality and quantity of data, the patterns of missingness, and the dynamic properties of the data-generating process. The considerations are likely unique to each country-specific study. Finally, predictions with time-series methods are based on historical information and not designed to smooth data that is simultaneously missing in multiple trajectories based on contemporaneous relations between intermittently observed time series with limited overlap.

¹³From the traditional missing data perspective, the MCAR situation is an easy problem as it effectively implies simply dropping incomplete cases will not bias subsequent inference. On the other side of the spectrum are MNAR problems in which the probability of a case being missing systematically varies for reasons that are unknown. In this case, data is needed that explains why certain observations are missing. From an inference perspective, MNAR problems are not even salvaged by predictive regression-based methods that make predictions for missing data based on other covariates. These methods simply

focus is entirely on specifying highly accurate prediction models for the expected prices.

A standard regression-based prediction approach would fill missing entries in A using data from (B,C), fill missing entries in B using information from (A,C), and so on. There are now two main issues that need to be solved to move forward with this idea. First, few matchable entries may exist. For example, only 8 out of 24 entries are missing in table 1, so the data is 67% complete. However, there are zero cases with elements in A, B and C that can be contemporaneously matched, so a regression cannot be fit even though price information is relatively abundant. Second, if missing entries in A are updated, then the information would certainly be relevant for the imputations in B and C. The information held in imputations in B and C, on the other hand, may lead one to find a better model to adjust the imputations in A and so on. The problems of simultaneity and few matchable entries are solved using a chained equations approach. The application here essentially adapts the Multiple Imputation using Chained Equations framework discussed extensively by Rubin (1996); van Buuren and Groothuis-Oudshoorn (2011); van Buuren (2012); Murray (2018), for predictive accuracy using ensemble methods to produce a stable prediction for missing data. In particular, the approach starts by estimating the following regression function:

(1)
$$(a_1, a_3, a_4) = f(b_1, \hat{b}_3, \hat{b}_4^*, \hat{c}_1, c_3, \hat{c}_4).$$

In this regression, \hat{b}_4^* is an estimate that has replaced outlier b_4^* and \hat{b}_3 , \hat{c}_1 , \hat{c}_4 are previously generated imputations. These values need to be initialized around some initial value. A suitable initialization and outlier-replacement method will be discussed. For now, let's assume these elements are initialized with some plausible guess work. After estimating the regression, the function \hat{f} can be used to update entries t=2,5,6 in variable A by generating the predictions

(2)
$$(\hat{a}_2, \hat{a}_5, \hat{a}_6) = \hat{f}(b_2, \hat{b}_5, b_6, c_2, \hat{c}_5, c_6).$$

Next, a new regression with the elements (b_1, b_2, b_6) from B on the left-hand side and matching entries from A, C on the right-hand-side can be constructed. As seen in table 1, the updated values \hat{a}_2, \hat{a}_6 generated with equation 2, using patterns learned in equation 1, would now be on the right-hand side, so the

amplify existing correlations and, acting as low-pass filters, reduce variability in the data. This leads to over-confident inferences. Stochastic methods that specify the mechanism of missingness and estimate a distribution for each missing data point are then needed (see van Buuren (2012)). In the current paper, this taxonomy is less relevant as the interest is in obtaining the single most likely value for a missing data point and not in corrections in standard errors of some final regression model. From this perspective, of a pure prediction problem, the MCAR problem is actually the hardest as it implies that no useful information exists to make predictions. Instead, prediction requires that the data is MAR or MCAR and that the values of a missing data point can accurately be predicted using the values of other covariates. The only word of caution is then that the predicted values are based on patterns in the observed data and underestimate total variance in the unobserved data. Section B.B4 develops methods to model the conditional heteroskedastic variance throughout the data to provide some plausible estimate of price variance at lower time frames.

covariates have been improved by the previous modeling step. This process can be repeated iteratively until all elements have been updated several times and new updates do not improve subsequent predictions any further. For each food item $d \in 1, ..., D$, in this example D = 3, this involves a chain of regression functions $\hat{f}_{d,i}$, where $i \in 1, ..., I$ indexes the updating iteration. The iterative sequence of regressions involved in updating the imputations is referred to as a regression chain because the predictions made by the regression fitted in the previous step, feed into the inputs of the next regression model.

Depending on several factors, such as the values at which the missing data values in the first regression equation are initialized, the order in which the algorithm cycles through the data, random elements of modeling (for example, taking bootstrap samples or using stochastic methods to generate synthetic training data), the result after I iterations will each time be different. Hence, the process is repeated M times, thus simulating a function space indexed by $\hat{f}_{d,i,m}$. Because the properties of simulated values for missing data change throughout the chain, the chained equations method allows to find various prediction models beyond what could be estimated with observed data alone. More details are provided in the appendix B.B1.

An ensemble predictor is finally constructed by averaging the M prediction results from the regression model at the tail of the simulation chain $\hat{f}_{d,I,m}$. In particular, after generating M imputations, the final imputation for a price element $\hat{x}_t \in (A, B, C)$ is generated by calculating the ensemble average

(3)
$$\hat{x}_t = \frac{1}{M} \sum_{1}^{M} (\hat{x}_t^{i=I,m}).$$

Since there are stochastic elements to the iterative algorithm, each prediction at iteration I will be generated from a different prediction model. Increasing M improves both the stability of the stochastic result as well as the accuracy of the ensemble prediction. This is similar to bootstrap aggregating, a prediction improvement technique central to the Random Forest algorithm of Breiman (2001), in which multiple random simple base learners are combined to improve stability and accuracy by canceling out random prediction errors. The key difference is that bootstrap aggregation produces multiple learners by taking random draws of the training data, while the randomization in learners at iteration I result from different stochastic simulations of chained model predictions, see section B.B2 for more details on the derivation of the stochastic ensemble predictor.

C. The regression specification

The methods are applied to each country individually, so the desired flexibility of the prediction models varies. Moreover, the properties of the data change throughout the simulation, thus the desired flexibility of the model should change accordingly, gradually exchanging robustness for flexibility. 14

The paper considers two implementations depending on data availability in the country. When data is scarce, an elastic net model is used (implemented by Hastie et al. (2021)), which helps reduce the impact of uninformative predictors (Friedman et al., 2010). When data is abundant, a cubist regression is used following Quinlan (1992); Witten et al. (2016). This is a piece-wise linear model that combines decision trees, boosting, and neighborhood smoothing, to capture smoothed versions of Random Forest-type of nonlinearities (Kuhn et al., 2012). The model differs from a Random Forest regression by using simple linear regressions at terminal nodes, thereby resulting in smoother transitions, that are more typical for numeric data, and enabling short-range out-of-sample extrapolation based on local regressions. Due to their linear nature, these local extrapolations generally remain reasonably stable, which is as opposed to, say, neural network predictions that can rapidly turn explosive outside observed data intervals. Cu-

¹⁴See for example the discussions in (Andrée et al., 2019; Andrée, 2020) on the relation between the sample size, strength of nonlinearity, and model size. Dynamic assumptions about the behavior of the process being modeled can be imposed by using non-parametric approaches whose parameterizations can arise flexibly. However, being overly flexible and over-fitting f would lead to terrible predictions. To a certain degree, the ensemble prediction will counter some of the prediction errors that arise from over-fitting. For example, similar to a Random Forest algorithm, if the prediction errors of models $m=1,\ldots,M$ at step I are uncorrelated, the ensemble simply cancels them out, thereby increasing robustness when M increases. Regardless, over-fitting can be problematic. Particularly important is to avoid an overly flexible model in the first estimation step. When f is an overly flexible model, an incorrect initialization can be learned, particularly if there is a pattern of missingness that can locally be correlated to the levels in other prices. For example, if missing entries in A occur during elevated levels in B, then initializing missing entries in A at the unconditional mean and using a flexible model to parameterize f, can lead to a model that simply learns to use the mean of A as a predictor within distinct regions of elevated levels in B. In such a situation, the imputations do not improve well across updating iterations. Essentially, across iterations, the model memorizes the initialization. To avoid the algorithm getting stuck, it is best to introduce a stochastic element around a reasonably correct initialization and use techniques that can control the flexibility of f across the iterations. In addition, it may be preferable to use only models for f that result in reasonably smooth transitions in nonlinearity across levels, rather than in very sharp cut-offs in local data associations. Since the entries on the covariate side are updated iteratively, there could also be more nuances fit by the function f as the algorithm iterates. This means while the flexibility of f should be relatively low in the first step, it could be beneficial to increase it across iterations. Finally, a variable selection mechanism is also advisable as the importance of predictive features may change across iterations, see (Graham et al., 2012).

¹⁵In total there are 186 different regression problems (number of food items summed across countries), 72 of these are solved with a linear model. Thus, 61% of problems are solved nonlinearly. Both models achieve good results, for instance in the Republic of Yemen and the Syrian Arab Republic, most of the regression problems are solved linearly but the cross-validated accuracy results are still good. The exact rules to determine whether a linear or nonlinear model is used are codified based on human judgment and best viewed in the source code that has been made available. It is possible to determine this by letting both models compete in a cross-validation exercise, but the runtime increases beyond what is practical. The guiding principle has been that when few markets are available, or long temporal gaps exist, and the results need to be extrapolated far across the data dimensions, the linear model is used.

¹⁶Cubist is an extension of M5 regression trees that incorporates pruning, neighborhood smoothing and boosting. Essentially is uses a computationally efficient strategy to recursively partition the data space and fit simple piece-wise linear prediction models within each partition, whose predictions are combined using neighborhood averaging of local model predictions. The advantages over M5 are that it can produce smoother transitions across numeric outputs, and much faster runtime. Both being of high importance to the current application. The advantages over Random Forest are that the cubist model has linear regressions at terminal nodes and so it can extrapolate slightly out of range, while Random Forests can only interpolate using medians or averages of typical values associated withing ranges of the input data.

bist models have done well on a variety of spatially oriented prediction problems, often reaching accuracy not far below that of deep learning methods while maintaining full model interpetability (Morellos et al., 2016; Ng et al., 2019; Sbahi et al., 2021). While the short-range extrapolation capabilities of cubist models are useful in the current setting, the most beneficial feature for this application here is that it runs much faster than its M5 cousin or other boosting or ensembling methods such as Gradient Boosting Machines and eXtreme Gradient Boosting methods (Hagenauer et al., 2019).

Two regression specifications are considered. Both regression models simply perform a spatio-temporal interpolation, leveraging temporal price trends, geographic proximity, and spatial trends, to make time-varying spatial interpolations between prices of related food items. More precisely, the linear model is of the following form:

(4)
$$P^{A} = \beta_0 + P^{-A}\beta_1 + X\beta_2 + G\gamma_1 + S\gamma_2 + \varepsilon,$$

while the cubist approximates the nonlinear function:

(5)
$$P^{A} = f(P^{-A}, X, S) + \varepsilon.$$

In these equations, P^A is a vector that stacks all the prices of commodity A observed at all $markets \times time$ combinations within a country. Similarly, P^{-A} is a matrix that has the prices of all other food items. In the previous example, it would simply bind columns (B,C), and the iterations would simply cycle over specifications according to the scheme $P^B \sim P^{-B} + \varepsilon_B$ and $P^C \sim P^{-C} + \varepsilon_C$. In this example, there are just three price vectors, but table A1 shows that the application considers problems with up to 15 price predictors.

Prices on both sides of the equation are modeled in logarithms, omitted from notation here. The vector β_1 captures the linearized relationships between log prices, e.g. the price ratios. The matrix G are group dummies for market-level fixed effects and administrative-level fixed effects, the matrix S are seasonal dummies. Finally, X is a matrix of additional covariates. These include logarithmic coordinates to capture spatial trends, and price trend features engineered from (A, B, C, \dots) that capture important temporal variation. The price trend features are engineered by first taking the individual market price trajectories contained in (A, B, C, ...) that have been observed with at least 95% data coverage, the up to 5% missing price points are imputed with a seasonal Kalman filter using a Basic Structural Model. Second, a commodity-specific country price trend is constructed by Kalman interpolating all market trajectories and taking a weighted average based on data coverage. In particular, trajectories that had above 75percentile data coverage are averaged, weighting by normalized data coverage rates. As an example, if there are three markets with above 75-percentile data coverage that respectively have 50%, 75% and 100% data coverage, the weights are (0, 0.5, 1).

Particularly within the matrix X there can be highly correlated predictors, while the matrix G may contain multiple identical indicators, for instance if there is only one market within an administrative unit. The following hard-coded variable selection rules are used. First, linear combinations are removed to avoid dummy problems. Highly collinear variables with > .95 correlation are removed by iteratively recalculating the correlation matrix and removing the variable with the highest overall correlation. Finally, variables with near-zero variance are removed. All predictors are centered and scaled.

Note that equation 4 is estimated using an elastic net model that uses L_1 and L_2 penalties to shrink the predictor space, see again (Friedman et al., 2010; Hastie et al., 2021). The cubist regression specification of equation 5 is essentially an observation specific counterpart of the linear model. As such, only the regional dummies are removed. The cubist model can partition the data and fit local regressions. Since the coordinates are supplied, the model can learn spatial fixed effects using adaptive neighborhood sizes as well as spatial interaction effects by partitioning by coordinates rather than relying on explicit spatial dummies, see again (Quinlan, 1992; Kuhn et al., 2012; Witten et al., 2016).

D. Validation of imputation accuracy

There are no countries in the data where the food price data is complete and the true inflation of a basket of food items is fully known. This makes validation against true data difficult.¹⁸ Cross-validation techniques are used to adjust the flexibility and assess the predictive accuracy of the models in each iteration. The elastic net model optimizes over the standard L_1 and L_2 penalties and the mixing parameter.¹⁹ The cubist model tunes over the neighborhood size used for smoothing, and the boosting iterations, using grid of all $Neighborhood \times Committees = (2, 4, 8) \times (1, 25, 50, 100)$ combinations.

The training and out-of-sample validating are as usual on mutually exclusive draws of observations. A 4-fold validation is used to limit the computational burden of the application to manageable levels.²⁰

¹⁷Food price series with zero variance do not occur in the data, but the cross-validation sampling might in theory land a draw that has extremely low price variation, particularly if the code would be applied to other countries.

¹⁸For instance, if a set of countries where the WFP data are complete would have been available, an interesting exercise would be to construct the *true* food basket price and accompanying inflation. Then, in this same sample, randomly a subset of the data could be removed to simulate a pattern of missing observations and the final basket price imputation could be imputed and compared to the true data.

 $^{^{19}}L_1$ penalizes likelihood by the absolute sum of coefficients, and L_2 by the sum of squared parameter values, thereby discouraging large parameter estimates but having very different impacts when redundant parameters approach 0. In particular, L_1 penalization can set parameter estimates to 0 as the penalty remains influential when a parameter approaches 0. The mixing parameter determines whether only L_1 , L_2 or a mix of penalties, is used, balancing between the well-known Lasso Regression and Ridge regression, or mixing them (elastic net).

²⁰Note that for D foods, a total of $countries \times M \times I \times \bar{D}_{country} \times (folds + 1) \times \Theta$ regressions need to be estimated, with $\theta \in \Theta$ indexing the different tuning configurations and $\bar{D}_{country}$ being the average number of food items in a country.

The training data contains in addition to the observed prices a small draw of 10% of the imputations generated from the previous regression estimates. These synthetic data points help balance prediction performance in thinner regions of the sample space. When the training data consists only of actual observations, the predictions may be biased toward purely observed value ranges. Adding synthetic cases using previous predictions makes the estimation problem slightly more representative of the missing data. The validation sample always consists only of actual data points, thus excluding imputations.²¹

The model-tuning focuses on a Normalized Mean-Absolute-Error criterion. The reason for this criterion is that the MAE is less impacted by outliers than the common Root-Mean-Squared Error measure and has a stable interpretation across applications. Note that an attempt will be made to replace outliers by missing data values and imputing them, but there is no guarantee that all outliers are captured and so an outlier-robust prediction validation metric is a safe option. There are M validation results for each food item, but the interest is primarily in the robustness of the final imputed price index which combines all the predictions. The validation results are condensed for better presentation. In particular, a cross-comparable confidence score is constructed from the individual validation estimates. The metric is defined as follows. First, a normalized MAE for food item d is constructed as the ratio of the MAE of the model for that food item to the MAE obtained by a simple mean prediction. Since each MAE estimate represents an average point percent error rate due to the log nature of the price data, the individual MAE values are averaged geometrically.

(6)
$$\widehat{NMAE}_d = \frac{\left(\prod_{m=1}^M \widehat{MAE}_{m,d}\right)^{\frac{1}{M}}}{MAE_d|\mu}$$

where $\widehat{MAE}_{m,d}$ is a cross-validation estimate of MAE using the standard for-

 21 Note that throughout the simulation, the quality of synthetic training data improves. Such techniques are also being explored elsewhere. For example, Lee and Braithwaite (2020) use a regression chain of image recognition models and feature-based models that update one another's training data which improved their learning results. In the current paper, f is parameterized using a piece-wise linear approach and adding bootstrap draws from previous imputations helped improve validation performance on actual data by allowing the models to train on denser representative example data near the edges of the sample space. This helps stabilize extrapolations outside of the sample space.

²²With simple arithmetic one can show that $MAE \leq RMSE \leq \sqrt{n}MAE$, which reveals that the upper-limit of RMSE varies with sample size and has different interpretations across applications.

 $^{23}M \times D \times I$ cross-validation metrics can be computed in each country application. The result at iteration I is used to diagnose the quality of imputations, but the full validation sequence provides a diagnostic to determine a relevant value for I. In particular, throughout the simulation, the cross-validation metrics should improve. The number of iterations for stochastic imputation is often determined by observing whether the means and variances of the predictions start changing in a purely random fashion. In the current case, the stopping criterion that the cross-validation performance does not improve further is a slightly less vague criterion. The multiple imputation implementation of van Buuren and Groothuis-Oudshoorn (2011) has default values of I=M=5. In the current application, diagnosing the performance indicated I=8 was usually sufficient. M=5 was kept due to the high computational load of the application.

mula for MAE, $MAE_d|\mu$ is the MAE calculated using observed data and the unconditional mean. Since the true data range is not observed, and averaging is known to improve ensemble performance, the quantity from equation 6 is likely a conservative estimate.²⁴

The focus next is on the quantity $1 - \widehat{NMAE}$, which is the share of the total absolute variation in the demeaned data explained by the imputation model. The D values are averaged as follows.

(7)
$$CV\text{-}score = \left(Z^{-1} \left(\frac{\sum_{d}^{D} w_{d} Z \sqrt{\left(1 - \widehat{NMAE}_{d}\right)}}{\sum_{d}^{D} w_{d}}\right)\right)^{2}$$

where Z is the Fischer Z-transformation and Z^{-1} its inverse, and w are the relative weights of the price component in the final price index. The final score from equation 7 roughly has the interpretation of the average R^2 of the food price index, using a robust calculation of out-of-sample errors. The unit of measurements are harmonized so that each food item has equal weight after in the index once scaled to a comparable unit of measurement. Specifically, the food item specific predictions are aggregated into food price indexes by summing the prices of all foods in the basket, after bringing the prices to comparable units of measurement (1 kg for foods, 1 liter for fluids and vegetable oils, 1 dozen for single packaged eggs, 1 unit for foods that come in other units of measurement – such as some fruits that come in bundles). As a simple example, if the index consists of 1 kg of sorghum and 100 Kg of maize, the latter price is multiplied by 0.01 and an equally weighted index is constructed with the result. Note

 24 Note that $MAE|\mu$ only uses the observed data range while the true data range is likely larger. As such, the estimated \widehat{NMAE}_d is likely larger than the true value since the denominator is underestimated in equation 6. There are further related challenges to the validation due to the fact that the general objective of imputing unobserved data can only be validated using observations. Some further rationale for the divergence metric and how it relates to the objective is provided in section B.B3.

 25 Note that $(1-\widehat{NMAE}) \geq (1-\widehat{NRMSE})$, where the second value equals the out-of-sample calculation of the R^2 . Keep in mind, however, that the data for validation may still contain outliers, see again section B.B3, so that $(1-\widehat{NRMSE}) > (1-\widehat{NRMSE})$. Regardless, when \widehat{NMAE} is low, such as in nearly all countries as the results will show and particularly in the outlier case, the value $(1-\widehat{NMAE})$ approaches the out-of-sample R^2 given by $(1-\widehat{NRMSE})$. Moreover, when \widehat{NMAE} is small, then model fit is good and average bias must be small, which means that $1-\widehat{NMAE}$ approaches the in-sample R^2 which is the squared Pearson-correlation between the predictions and the observations. It is well known that the arithmetic mean of multiple correlation coefficients underestimates the total correlation and that the distribution of the correlation coefficient must first be normalized before averaging to obtain a less biased estimate, particularly when the number of coefficients is small. See (Corey et al., 1998) on this matter and the use of the Fischer Z-transformation in this context.

 26 The standardization in the units of measurement is automated and works by parsing the text in the WFP data base that describes the food items and inferring multipliers that standardize the prices to comparable units. In particular, each text string is parsed as $amount \times unit$ and a multiplier for food item d is calculated as $w_d = \frac{1}{amount}c$, where c is a conversion factor from unit to either Kilograms or Liters. For example, the text string 10 Pounds of Rice would result in a multiplier of $w_{rice} = \frac{1}{10} \times \frac{1}{0.453592} = 0.2204624$, while 1 kg of wheat would simply result in a weight of $w_{wheat} = 1$. This strategy is usually

that from a nutritional perspective, some foods should be weighted more strongly to model a food price index that reflects preferable consumption ratios.²⁷ The simple approach here is just to ensure that an item measured in larger quantities does not dominate the final index simply because it has a larger price range, or dominate the validation result simply because the unit of measurement is small. This is not ideal, but sensible, given that expenditure shares for specific food items used to construct traditional CPI are not widely available in the countries analyzed by the paper. Recall that the estimation exercise of the paper is in fact motivated by the unavailability of traditional data. Tables A2 and A3 list for all countries the food-specific Index Weights used to scale price levels to correspond to prices for comparable units of measurement.

E. The initialization

A two-step approach is taken to initialize the regression chain. First, univariate imputation methods are used to pre-impute the starting values of missing entries.²⁸ The iterative modeling then initializes randomly around the pre-imputed values by adding an additive disturbance term to the initialization drawn from a uniform distribution centered around 0, scaled to 10% of the range in the data. The variance of the prices is stabilized with logarithms, so this disturbance term impacts the initialization roughly evenly across levels.

It is important to clean the data from outliers before calculating the preimputations and applying the imputation algorithm. In particular, outliers can lead to explosive predictions or generally bad model fit depending on the types of learning methods used. Since the outliers need to be removed from incomplete time series, standard time series methods such as those relying on non-parametric

reasonable and avoids that a food commodity measured in bags of $100\ kg$ dominates the food basket price, but is obviously not smart enough to deal with all situations effectively. A specific rule is specified to deal with eggs, which may be described in some countries as $12\ units$, while in other countries eggs are measured as $1\ dozen$. In the case of ' $12\ units$ ', a conversion to $1\ dozen$ is made. All the price estimates from the paper are provided for analysis, and the index weights are provided with the index estimates. If it is suspect that conversion factors may have an impact on the final inflation rate estimates, researchers are encouraged to construct their own indexes relevant to the studies at hand using the individual food price estimates.

²⁷For example, 1 kg of salt and 1 kg of sorghum are given equal weight in the simple index generated here, whereas the latter may carry more dietary importance. The reason for using equal weights is that it takes expert judgment to define weights based on dietary needs, which is not easily automated given the heterogeneity of the data. In an ideal world, future price survey programs would be accompanied with surveys on household expenditure shares or food specific trading volumes.

 $^{28} \rm First,$ minor gaps of up to 3 consecutive months are filled with a univariate Seasonal Kalman smoother using a Basic Structural Model. Then food-market combinations with at least 67% data coverage are completed using the same method. Next, all markets trajectories completed in this way are averaged to construct a preliminary country mean. All other market level series with >50% data coverage are then imputed using predictions from a Generalized Additive Model with observed price points as dependent variable and the preliminary country mean as the predictor. The parameters are estimated by Restricted Maximum Likelihood. The country average is recalculated, if there are any further markets with missing data, but >20% data coverage, the same method is again applied. Any market which has <20% data coverage, is pre-imputed by using an inverse distance weighing interpolation based on the other imputations. The spatial interpolation uses a neighborhood size cutoff which is selected using cross-validation.

smoothers with adaptive bandwidths cannot be applied. The approach here first applies a linear interpolation to the raw data, then calculates returns and detects outliers in the returns sequence using the approach by Boudt et al. (2008). The outliers are then turned into missing data points which will be replaced by imputation. In most cases, severe outliers are simply measurement errors. Minor outliers that simply reflect the natural extreme price variation of illiquid markets, should already be reasonably stabilized by the log transformations.

III. Results

The methods are applied country-by-country to 25 FCS countries. There are a high number of results, all commodity-specific validation and predictions results are available in the online repository of this paper. Since for multiple countries in the sample no official food CPI are available for the periods of analysis, the section here shall discuss some of the highlights.

The cross-validation results of the final price estimates are summarized in the last column of table A4 and highlight that, even on a monthly basis, the market-level price predictions are fairly accurate. The Fisher-average CV-score across countries is 0.85. The lowest result is at 0.69 still a reasonably well informed prediction result given the natural volatility in prices that make direct measurement similarly prone to error.

To better understand the determinants of prediction reliability, a simple linear regression was performed with results in A5. This indicated that the share and relative dimension of available data (number of foods, number of markets) are not linear predictors of prediction accuracy, instead the volatility and strength of inflation determine how well the statistical methods impute missing data. Jointly, average inflation and volatility explain almost half of the variation in cross-validation performance. The signs of the coefficients suggest that when volatility increases, prediction performance deteriorates and data collection becomes more important. When the inflation trend is strong relative to volatility, then the price trend is clearer in the data and imputation accuracy increases, suggesting that robust inflation tracking is possible even when little ground truth data is available. Purely based on signs, there is some indication that an increased number of food items is beneficial and an increased number of markets makes the prediction task more difficult. This is in line with the idea that a higher number of food items increases the chances that at least one strongly related food item can be leveraged for prediction, while an increased number of markets increases spatial heterogeneity in the data thereby increasing the difficulty of accurately predicting all local price levels.

The subnational results are aggregated into national food price indexes by calculating a simple average price giving equal weight to each market within the country. The imputation focuses on the 676 markets where prices are observed, and spatially interpolates the results onto the full set of markets using a Shepard with cross-validated neighbors algorithm. The country average price indexes are

thus based on the full set of markets, but are essentially derived as a weighted average of the imputed markets with more weight given to markets that are more densely surrounded by other markets where WFP maintains or has maintained price monitoring operations.

Table A4 also summarizes the monthly changes in the price indexes over the entire study period using three simple metrics. First, the monthly returns in the price index are annualized. This gives the average rate at which food prices have inflated at an annual basis across all available time periods. The second column provides a maximum draw-down figure, which is the largest negative percentage price change that occurred from top to bottom over the full time period. Apart from price changes, economists are frequently also interested in price uncertainty. The third column in table A4 contains the annualized standard deviation in monthly price returns, which is the average price volatility over the entire period of study. For example, in Afghanistan prices have increased at an average annual rate of 5.76%, the largest deflationary event was a measured price move of -43.86\% from top to bottom, and the average standard deviation with which price changes fluctuated was 8.69%. The simple aggregation of results highlights that food price inflation is problematic in many FCS countries. Inflation-targeting countries often aim for a positive inflation rate just below 2%. Only 6 out of the 25 analyzed countries meet this criterion for food when measured over the entire study period. Moreover, in 7 countries, active price increases in food have been so strong that they outweighed price uncertainty.

Historical famines have frequently been associated with high price increases that led to speculative purchasing of food as an investment (Gráda, 2007). Investors typically aim for investments in high-return low-volatility assets and often track the performance of investments by focusing on risk-adjusted returns (Markowitz, 1991). A commonly applied metric is the Sharpe ratio, which is the annualized return divided by the annualized standard deviation of returns. The ratio is a z-score type of metric that describes the excess return received for the extra volatility endured when holding an asset. In general, for an asset to be investmentgrade, one likes the ratio to be above 1 so that the price increase outweighs the price uncertainty.²⁹ The Sharpe ratios can be obtained from table A4 by dividing the first column by the third. It is important to note that this ratio is constructed in hindsight and does not fully characterize all the factors involved with an investment decision, nor does it take into consideration the scarcity of information and the complexity of the market environment during inflation events. Nevertheless, it provides a simple and well-understood way to standardize price changes in a way that allows comparing the significance of price changes across markets associated with different typical volatilities. For example, a 4% inflation surge is more significant if the typical price change over that time period is 1% as opposed to, say, 10%. Contrasting these ratios, highlights that some inflation

²⁹As a reference, the average Sharpe ratio of the dollar denominated S&P500 index has been around 1 over the past 25 years, while good performing hedge funds often achieve annual Sharpe ratios of 1.5.

events have been more significant than others. For example, in Afghanistan, the average ratio between price increase and price uncertainty in local currency as measured over all time steps equals 0.66. At this value, inflation is positive but price uncertainty outweighs price increase and so there is no strong incentive to hold on to food to hedge against currency depreciation risks. This is opposed to Sudan, South Sudan and the Syrian Arab Republic, where the increase in food prices has not only been extremely high, but has also outweighed food price uncertainty by roughly a factor 2. Alternatively, food prices in Afghanistan rose by 14.09% annualized in 2020, which seems low when compared to some other high inflation events, but highly significant when compared to the relatively low price volatility over that same time period (inflation was 2.67 times volatility).

Inflation-targeting countries also aim for prices to be stable on low time frames. Tables A6 and A7 condense the high frequency results by presenting annualized figures. An FCS-wide average is included at the bottom. In total, price estimates for 1223 markets and 43 foods underly these figures. For all $year \times countries$ combinations for which estimates are produced (361), only 48 annualized inflation rate estimates fall within the 0-2% targeting range, while 51 estimates cross a critical inflation threshold of 50%. The FCS-wide results highlight the elevated inflation levels during the period associated with the World Food Price Crisis of 2007-08 and the aftermath of the pandemic. The annualized monthly price change in 2008 peaks at geometric average of 19.52% across the full set of countries, while remaining mostly within a modest several percentage point range over the 2009 to 2018 period. The estimated average inflation rate then spikes again at 27.46%in the year of the pandemic, while the preliminary 2021 inflation estimate of 22.65% remains well above levels of the World Food Price Crisis of 2007-08. The results in table A7 highlight that price uncertainty has remained more constant. A striking result is that average price uncertainty was higher during the World Food Price Crisis of 2007-08, peaking at an annualized rate of 13.92%, while hitting 10.52\% during the pandemic shock and even dropping to 8.1\% in the preliminary 2021 estimate. This means that the recent high inflation event was not only stronger in terms of price increase, but also more significant when compared to natural price volatilities in the two periods. It is from this perspective again interesting to transform the estimates to Sharpe ratios. Both the World Food Price Crisis of 2007-08 and the post-pandemic inflation surge was associated with prices increasing relative to price uncertainty, and this price development is more pronounced in the post-pandemic inflation event. For example, from 2009 to 2018 the FCS-wide average annual Sharpe ratio for food prices was 0.55, but during 2007-08 it averaged 1.30. During 2020, this ratio hits 2.70, while the preliminary annualized 2021 estimate sits at 2.80, meaning that returns in food prices outweigh price volatility risks of holding food as an asset by nearly a factor 3. This means the ratio of inflation to price uncertainty is almost double that of the World Food Price Crisis of 2007-08.

The monthly price estimates indicate that prices can rise and fall dramatically

within very short time periods in FCS countries. Figure A1 presents estimated monthly year-on-year inflation on a line chart for the Republic of Yemen together with intra-month price ranges on a candle chart. The intra-month volatility algorithm is detailed in section B.B4 and uses a conditional autoregressive heteroskedasticity model to estimate the time-varying properties of the monthly price variance process. The estimates are used to calculate Expected Shortfall in the price returns, which is used to construct the wicks and bodies of the candles. The open values are defined as the estimated conditional expectation of monthly prices, the highs are the average intra-month price levels in the worst 50% of estimated conditional price increase events, the lows are equal to the average intra-month price levels in the worst 50% of estimated conditional price decrease events. The colored parts of the candles are the estimated price ranges within which the majority of monthly prices are estimated to be, with red candles indicating months in which the end-of-month prices closed below the start-of-month price estimates. Strong inflationary events are thus visualized as consecutive green candles. Periods of high volatility are visualized by candles with large wicks. Results for all other countries are available in figures A2 to A4. The charts in figures A2 to A4 highlight that there have been several notable high inflation events. The World Food Price Crisis of 2007-08 is visible as a price up-tick in most countries, most notably in Afghanistan, Burkina Faso, Chad, Haiti, Niger and Somalia. In some of the charts, the price action seems modest as it is suppressed by recent price action. In these instances, it may be more useful to look at prices on a logarithmic chart. Prices have also visibly surged in many countries in the year following the pandemic.

IV. Discussion and concluding remarks

Food price inflation is an important metric to inform economic policy and is closely watched by both economists and humanitarians, yet official figures are often missing, lacking in spatial detail, or published with delay. More important, crisis situations, or vulnerable populations in general, are often characterized by high geographic specificity while traditional price indicators do not provide insights beyond the major (urban) markets where prices are formally measured. Traditional CPI data in these situations, if available, may be insufficiently unbundled into distinct price estimates for it to be put into a relevant context. Recognizing this, international institutions have invested substantially in subnational price surveys. However, the capacity to monitor inflation has continued to be severely limited by challenges related to missing data. To overcome some of these shortcomings, this paper proposed an approach for real-time imputation of survey data drawing on multiple imputation and ensemble learning ideas.

The paper highlighted the new price monitoring capabilities using survey data gathered in 25 fragile and conflict-affected countries. The final estimates of food price inflation documented in this paper were shown to accurately capture important inflation events including the World Food Price Crisis of 2007-08 and the

surge in inflation that followed the 2020 pandemic and subsequent expansion in the global monetary base.

The paper used out-of-sample validation techniques to estimate the reliability of the augmented data. The share of missing data (20.4% to 79.18%), the number of markets (3 to 77), and the number of food items (3 to 16), varied widely across the country-specific applications, but the imputation methods were shown to remain robust across these aspects, even when data coverage was relatively low. A linear regression that correlated the cross-validated prediction performance with key properties of the country applications showed that data coverage itself was not a correlate of prediction accuracy. Instead, the strength of inflation and the natural volatility in prices are strong predictors of the imputation accuracy.

The accuracy of the imputations was judged by estimating the total price variation explained by the imputation models. On average, across the countries, the models predicted 85% of the observed price variation. In individual countries, the errors were in the 5% to 30% range of observed prices. This puts the accuracy of imputations in a similar range of that of the direct measure of prices at major urban markets in countries with well-established CPI methods. In particular, direct measure of prices is prone to error with respect to true price levels simply because of the natural volatility in prices. For example Lebow and Rudd (2003) estimated that measurement error in CPI change rates in the United States, a country with strong statistical capacities, could have been as high as 0.3% to 1.4% points over a period where the change rates were typically in a 3%-6% range, which, similarly, places the measurement error of official CPI change in a 5% to 45% range of true values.

An important result is thus that as long as incomplete and intermittent survey data has a at least a 20% to 40% rate of completeness, it can be augmented reliably with predictions to monitor subnational food price trends continuously. This contributes to a previously non-existing capacity to provide insights beyond the major (urban) markets where prices are formally measured with traditional methods. In situations with great spatial heterogeneity, and localized vulnerabilities — such as in fragile environments or during crisis — this can provide important new insight. Additionally, in those markets where prices are very sensitive to localized shocks, the statistical estimates may provide new opportunities to investigate local price dynamics with a similar confidence as one would have obtained otherwise with measured prices.

A cross-country analysis of inflation trends, and comparison with inflation trends at global markets such as captured by the FAO or World Bank Food Price Index, would be interesting for future analyses, but is beyond the scope of the paper as it would require dealing with the differences in exchange rates. In a few countries, such as the Republic of Yemen, local unofficial exchange rates are also surveyed by WFP. In a separate analysis, the methods of the paper have been applied to track prices of 23 items covering foods and non-foods in the Repub-

lic of Yemen.³⁰ This highlighted that it is possible to track dollar-denominated prices, broadly for the five food categories that comprise the FAO index, and draw comparisons with price events in global markets when the exchange rate is also monitored. In this paper, it is simply noted that figures A2 to A4 provide some visual guidance on how global price change events propagate at the national level, as nearly all the country graphs display inflation events in 2020. The figures indicate that there may be substantial cross-country heterogeneity in the timing and amplitude of the price surges, which likely relates to differences in factors such as whether a country is a net importer or exporter of food and how the local currency is managed during times of crisis. Finally, since not all products matter the same for household well-being, future analysis may use the estimates developed by this paper to explore food-specific price dynamics or produce inflation estimates based on food baskets that incorporate information on expenditure shares.

Deploying statistical methods to enhance data gathering may help improve price estimates more widely. For instance, additional investments could be redirected to broaden the scope of data gathering to include additional sampling locations and food items rather than be used to strengthen data coverage of existing narrow monitoring operations. This could give a more complete view on subnational inflation than can be achieved by pure data gathering alone. Moreover, gathering data on additional food items can help produce more accurate predictors for missing observations. The imputations primarily utilize contemporaneous relations between the prices of different items, and so priority in data gathering should be given to ensuring that in each month prices of at least some items at some markets are observed. Ensuring that some data is available most of the time periods is more important than ensuring that most of the data is available in some of the time periods. In addition, predictions are more reliable in high inflation episodes and less reliable in high volatility episodes, so data gathering processes may also use statistical methods to produce real-time price estimates and use the results in turn to determine how much new data gathering is needed.

References

- Ahumada, H. and Cornejo, M. (2016). Forecasting food prices: The case of corn, soybeans and wheat. *International Journal of Forecasting*, 32(3):838–848.
- Andrée, B. P. J. (2020). Theory and Application of Dynamic Spatial Time Series Models. Rozenberg Publishers and Tinbergen Institute, Amsterdam.
- Andrée, B. P. J. (2021). Monthly food price estimates by product and market. In WLD_2021_RTFP_v02_M, Version 2021-12-02. World Bank Microdata Library, Washington, DC.

 $^{^{30}}$ Results and adapted code are available from the author upon request.

- Andrée, B. P. J., Chamorro, A., Spencer, P., Koomen, E., and Dogo, H. (2019). Revisiting the relation between economic growth and the environment; a global assessment of deforestation, pollution and carbon emission. *Renewable and Sustainable Energy Reviews*, 114:109221.
- Andrée, B. P. J., Diogo, V., and Koomen, E. (2017). Efficiency of second-generation biofuel crop subsidy schemes: Spatial heterogeneity and policy design. *Renewable and Sustainable Energy Reviews*, 67:848–862.
- Andrée, B. P. J., Kraay, A., Chamorro, A., Spencer, P., and Wang, D. (2020). Predicting Food Crises. World Bank Policy Research Working Papers.
- Andreyeva, T., Long, M. W., and Brownell, K. D. (2010). The impact of food prices on consumption: A systematic review of research on the price elasticity of demand for food.
- Baffes, J., Mitchell, D., Riordan, E. M., Streifel, S., Timmer, H., and Shaw, W. (2008). *Global Economic Prospects: Commodities at the Crossroads 2009*. World Bank Publications, Washington, DC.
- Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.
- Baillie, R. T., Han, Y. W., and Kwon, T.-G. (2002). Further Long Memory Properties of Inflationary Shocks. *Southern Economic Journal*, 68(3):496.
- Boudt, K., Peterson, B., and Croux, C. (2008). Estimation and decomposition of downside risk for portfolios with non-normal returns. *The Journal of Risk*, 11(2):9.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*.
- Breiman, L. (2001). Random forests. Machine Learning, 45(1):5–32.
- Chang, C.-L., McAleer, M., and Tansuchat, R. (2012). Modelling Long Memory Volatility In Agricultural Commodity Futures Returns. *Annals of Financial Economics*, 07(02):1250010.
- Conceição, P. and Mendoza, R. U. (2009). Anatomy of the global food crisis. Third World Quarterly.
- Corey, D. M., Dunlap, W. P., and Burke, M. J. (1998). Averaging correlations: Expected values and bias in combined pearson rs and fisher's z transformations. Journal of General Psychology, 125(3):245–261.
- Cutler, P. (1984). Famine forecasting; Prices and peasant behaviour in Northern Ethiopia. *Disasters*, 8(1):48–56.

- Diogo, V., Reidsma, P., Schaap, B., Andrée, B. P. J., and Koomen, E. (2017). Assessing local and regional economic impacts of climatic extremes and feasibility of adaptation measures in Dutch arable farming systems. Agricultural Systems, 157:216–229.
- Dollar, D. and Kraay, A. (2002). Growth Is Good for the Poor. *Journal of Economic Growth*.
- Durbin, J. and Koopman, S. J. (2013). Time Series Analysis by State Space Methods. Oxford University Press.
- Easterly, W. and Fischer, S. (2001). Inflation and the Poor. *Journal of Money*, *Credit and Banking*, 33(2):160.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22.
- Gavin, W. T. and Mandal, R. J. (2002). Predicting inflation: food for thought. The Regional Economist, (Jan.):4–9.
- Ghalanos, A. (2020). Introduction to the rugarch package. (Version 1.4-3).
- Gráda, C. Ó. (2007). Making Famine History. *Journal of Economic Literature*, 45:5–38.
- Graham, J. W., Van Horn, M. L., and Taylor, B. J. (2012). Dealing with the Problem of Having Too Many Variables in the Imputation Model. In *Missing Data*, pages 213–228. Springer New York, New York, NY.
- Ha, J., Ivanova, A., Ohnsorge, F., and Unsal, F. (2019a). Inflation: Concepts, Evolution, and Correlates. *Policy Research Working Paper*.
- Ha, J., Kose, M. A., and Ohnsorge, F. L. (2019b). Understanding Inflation in Emerging and Developing Economies. *Policy Research Working Paper*.
- Hagenauer, J., Omrani, H., and Helbich, M. (2019). Assessing the performance of 38 machine learning models: the case of land consumption rates in Bavaria, Germany. *International Journal of Geographical Information Science*, 33(7):1399–1419.
- Hastie, T., Qian, J., and Tay, K. (2021). An Introduction to glmnet.
- Joutz, F. L. (1997). Forecasting CPI Food Prices: An Assessment. American Journal of Agricultural Economics, 79(5):1681–1685.
- Khan, M. (1994). Market-based early warning indicators of famine for the pastoral households of the Sahel. World Development, 22(2):189–199.

- Koomen, E., Diogo, V., Dekkers, J., and Rietveld, P. (2015). A utility-based suitability framework for integrated local-scale land-use modelling. *Computers, Environment and Urban Systems*, 50:1–14.
- Koopman, S. J. and Durbin, J. (2000). Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis*, 21(3):281–296.
- Koopman, S. J., Lit, R., and Nguyen, T. M. (2019). Modified efficient importance sampling for partially non-Gaussian state space models. *Statistica Neerlandica*, 73(1):44–62.
- Kuhn, M., Weston, S., Keefer, C., and Coulter, N. (2012). Cubist Models For Regression.
- Lebow, D. E. and Rudd, J. B. (2003). Measurement Error in the Consumer Price Index: Where Do We Stand? *Journal of Economic Literature*, 41(1):159–201.
- Lee, K. and Braithwaite, J. (2020). High-Resolution Poverty Maps in Sub-Saharan Africa. arXiv, 2009.00544.
- Markowitz, H. M. (1991). Foundations of Portfolio Theory. The Journal of Finance, 46(2):469.
- Maxwell, D., Khalif, A., Hailey, P., and Checchi, F. (2020). Determining famine: Multi-dimensional analysis for the twenty-first century. *Food Policy*, 92:101832.
- Modugno, M. (2011). Nowcasting Inflation using High Frequency Data. Working Paper Series.
- Morellos, A., Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., and Mouazen, A. M. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, 152:104–116.
- Murray, J. S. (2018). Multiple Imputation: A Review of Practical and Theoretical Findings. *Statistical Science*, 33(2):142–159.
- Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., and McBratney, A. B. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, midinfrared, and their combined spectra. *Geoderma*, 352:251–267.
- Ouyang, H., Wei, X., and Wu, Q. (2019). Agricultural commodity futures prices prediction via long- and short-term time series network. *Journal of Applied Economics*, 22(1):468–483.
- Quinlan, J. R. (1992). Learning With Continuous Classes. *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence*, pages 343—348.

- Reinsdorf, M., Triplett, J. E., Reinsdorf, M., and Triplett, J. E. (2009). A Review of Reviews: Ninety Years of Professional Thinking About the Consumer Price Index. In *Price Index Concepts and Measurement*, pages 17–83. National Bureau of Economic Research, Inc.
- Rubin, D. B. (1976). Inference and missing data. Biometrika, 63(3):581-592.
- Rubin, D. B. (1996). Multiple Imputation After 18+ Years. Journal of the American Statistical Association, 91(434):473.
- Sbahi, S., Ouazzani, N., Hejjaj, A., and Mandi, L. (2021). Neural network and cubist algorithms to predict fecal coliform content in treated wastewater by multi-soil-layering system for potential reuse. *Journal of Environmental Quality*, 50(1):144–157.
- Seabold, S. and Coppola, A. (2015). Nowcasting prices using Google trends: an application to Central America. *Policy Research Working Paper Series*.
- Seaman, J. and Holt, J. (1980). Markets and Famines in the Third World. *Disasters*, 4(3):283–297.
- Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika*, 81(1):115–131.
- The World Bank (2019). Inflation in Emerging and Developing Economies: Evolution, Drivers, and Policies. The World Bank.
- van Buuren, S. (2012). Flexible Imputation of Missing Data. Flexible Imputation of Missing Data.
- van Buuren, S. and Groothuis-Oudshoorn, K. (2011). Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3):1–67.
- Wan, E. A. and Van Der Merwe, R. (2000). The unscented Kalman filter for nonlinear estimation. In *IEEE 2000 Adaptive Systems for Signal Processing*, Communications, and Control Symposium, AS-SPCC 2000, pages 153–158. Institute of Electrical and Electronics Engineers Inc.
- Wang, D., Andrée, B. P. J., Chamorro, A. F., and Spencer, P. G. (2020). Stochastic modeling of food insecurity. World Bank Policy Research Working Papers.
- Waterlander, W. E., Jiang, Y., Nghiem, N., Eyles, H., Wilson, N., Cleghorn, C., Genç, M., Swinburn, B., Mhurchu, C. N., and Blakely, T. (2019). The effect of food price changes on consumer purchases: a randomised experiment. *The Lancet Public Health*, 4(8):e394–e405.
- Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques. Elsevier Inc.
- World Bank (2021). South Sudan Economic Update, June 2021: Pathways to Sustainable Food Security. Technical report, World Bank, Washington, DC.

TABLES AND FIGURES

Table A1—Summary of raw food price data.

Country	Currency	Markets	Items	Data Coverage	Start date
Afghanistan	AFN	9/40	4	69.27%	Jan-07
Burkina Faso	XOF	63/65	3	47.79%	Jan-07
Burundi	BIF	61/68	7	35.93%	Jan-07
Cameroon	XAF	12/51	11	20.82%	Jan-07
Central Afr. Rep.	XAF	18/40	3	36.89%	Jan-08
Chad	XAF	35/56	3	39.17%	Jan-07
Congo, Rep.	XAF	5/11	7	44.47%	May-10
Congo, Dem. Rep.	CDF	26/83	10	35.43%	Nov-07
Gambia, The	GMD	15/28	11	41.91%	Jan-07
Guinea-Bissau	XOF	3/45	7	56.63%	Feb-16
Haiti	HTG	9/9	6	68.83%	Jan-07
Iraq	IQD	18/18	13	48.02%	May-11
Lao PDR	LAK	17/17	5	46.04%	Feb-12
Lebanon	LBP	26/26	15	61.49%	Mar-12
Liberia	LRD	18/24	3	49.44%	Mar-07
Mali	XOF	77/126	6	58.37%	Jan-07
Mozambique	MZN	25/52	7	63.13%	Jan-07
Myanmar	MMK	36/165	3	43.00%	Apr-07
Niger	XOF	68/79	4	79.60%	Jan-07
Nigeria	NGN	33/35	16	27.59%	May-12
Somalia	SOS	18/28	4	55.49%	Jan-07
South Sudan	SSP	9/20	8	52.35%	Jan-07
Sudan	SDG	14/14	3	58.01%	Jan-07
Syrian Arab Rep.	SYP	36/91	15	56.99%	Aug-11
Yemen, Rep.	YER	24/24	12	45.48%	Nov-08
Average		27/49	7	51.47%	

Note: The first column reports the local currency in which prices are measured, the second column reports the number of markets from which data is used as a fraction of all known market location for which predictions are made, the third columns reports the number of food items for which data is used to construct the food price index, the fourth columns reports the total number of price observations as a share of all $market \times time$ combinations where markets is the first number in the third column, the final column reports when the estimated price index starts.

Source: The statistics have been prepared by the author for this paper based on end-of-August (2021) food price data from World Food Program.

Table A2—Summary of food price index components.

```
Index Component
Bread - Retail (1 kg, Index Weight = 1), Rice (Low Quality) - Retail (1 kg, Index Weight = 1),
 Country
 Afghanistan
                                                                                         Wheat - Retail (1 kg, Index Weight = 1), Wheat Flour - Retail (1 kg, Index Weight = 1)
Beans (Niebe) - Retail (1 kg, Index Weight = 1), Maize (White) - Retail (1 kg, Index Weight = 1),
 Burkina Faso
                                                                                        Beans (Niebe) - Retail (1 kg, Index Weight = 1), Malze (White) - Retail (1 kg, Index Weight = 1), Millet - Retail (1 kg, Index Weight = 1) Rice (Low Quality, Local) - Retail (1 kg, Index Weight = 1), Beans - Retail (1 kg, Index Weight = 1), Malze (White) - Retail (1 kg, Index Weight = 1), Bananas - Retail (1 kg, Index Weight = 1), Cassava Flour - Retail (1 kg, Index Weight = 1), Onions - Retail (1 kg, Index Weight = 1), Sweet Potatoes - Retail (1 kg, Index Weight = 1)
 Burundi
                                                                                        Sweet Fotation's retain (1 g, index Weight = 1), Rice (Local) - Wholesale (90 kg, Index Weight = 0.01), Wheat Flour - Retail (1 kg, Index Weight = 1), Beans (Niebe) - Wholesale (90 kg, Index Weight = 0.01), Maize - Wholesale (90 kg, Index Weight = 0.01), Sorghum (Red) - Wholesale (90 kg, Index Weight = 0.01), Bananas - Retail (12 kg, Index Weight = 0.08), Potatoes - Retail (1 kg, Index Weight = 1),
 Cameroon
                                                                                         Cassava (Fresh) - Retail (5 kg, Index Weight = 0.2), Cocoyam (Macabo) - Retail (20 kg, Index Weight = 0.05), Plantains - Retail (1 kg, Index Weight = 1) Oil (Palm) - Retail (1 L, Index Weight = 1), Rice - Retail (1 kg, Index Weight = 1),
 Central Afr. Rep.
                                                                                         \begin{aligned} & \text{Maize - Retail (1 kg, Index Weight} = 1) \\ & \text{Maize (White) - Retail (1 kg, Index Weight} = 1), \\ & \text{Millet - Retail (1 kg, Index Weight} = 1), \end{aligned}
                                                                                         Sorghum (Red) - Retail (1 kg, Index Weight = 1)
Bread - Retail (1 kg, Index Weight = 1), Oil (Palm) - Retail (1 L, Index Weight = 1),
 Congo, Rep.
                                                                                      Bread - Retail (1 kg, Index Weight = 1), Oil (Palm) - Retail (1 L, Index Weight = 1),
Rice (Mixed, Low Quality) - Retail (1 kg, Index Weight = 1), Wheat Flour - Retail (1 kg, Index Weight = 1),
Beans (White) - Retail (1 kg, Index Weight = 1), Groundnuts (Shelled) - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Rice (Local) - Retail (1 kg, Index Weight = 1),
Sugar - Retail (1 kg, Index Weight = 1), Wheat Flour - Retail (1 kg, Index Weight = 1),
Beans - Retail (1 kg, Index Weight = 1), Maize - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Maize Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Maize Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Maize Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Maize Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Maize Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1), Maize Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail (1 kg, Index Weight = 1),
Cassava Flour - Retail
Cong. Dem. Rep.
                                                                                         Cassava (Cossette) - Retail (1 kg, Index Weight = 1), Plantains - Retail (1 kg, Index Weight = 1)
Oil (Vegetable) - Retail (1 L, Index Weight = 1), Rice (Small Grain, Imported) - Retail (1 kg, Index Weight = 1),
 Gambia, The
                                                                                        On (vegetanie) - retail (1 kg, Index Weight = 1), Ruce (Smail Grain, Imported) - retail (1 kg, Index Weight = 1), Beans (Dry) - Retail (1 kg, Index Weight = 1), Groundnuts (Shelled) - Retail (1 kg, Index Weight = 1), Millet - Retail (1 kg, Index Weight = 1), Bananas - Retail (1 kg, Index Weight = 1), Onions - Retail (1 kg, Index Weight = 1), Tomatoes - Retail (1 kg, Index Weight = 1), Garrots - Retail (1 kg, Index Weight = 1), Carrots - Retail (1 kg, Index Weight = 1)
                                                                                        Cartos - Retail (1 kg, Index Weight = 1)
Oil (Vegetable, Imported) - Retail (1 L, Index Weight = 1), Rice (Imported) - Retail (1 kg, Index Weight = 1),
Sugar - Retail (1 kg, Index Weight = 1), Wheat - Retail (1 kg, Index Weight = 1),
Millet - Retail (1 kg, Index Weight = 1), Sorghum - Retail (1 kg, Index Weight = 1),
Fonio - Retail (1 kg, Index Weight = 1)
Guinea-Bissau
                                                                                      Fonio - Retail (1 kg, Index Weight = 1)

Oil (Vegetable, Imported) - Retail (1 Gallon, Index Weight = 0.26), Rice (Tchako) - Retail (1 Marmite, Index Weight = 0.37),
Sugar (White) - Retail (1 Marmite, Index Weight = 0.37), Wheat Flour (Imported) - Retail (1 Marmite, Index Weight = 0.37),
Beans (Black) - Retail (1 Marmite, Index Weight = 0.37), Maize Meal (Local) - Retail (1 Marmite, Index Weight = 0.37),
Bread (Khoboz) - Retail (1 Unit, Index Weight = 1), Oil (Vegetable) - Retail (1 L, Index Weight = 1),
Rice - Retail (1 kg, Index Weight = 1), Sugar - Retail (1 kg, Index Weight = 1),
Wheat Flour - Retail (1 kg, Index Weight = 1), Beans (White) - Retail (1 kg, Index Weight = 1),
Potatoes - Retail (1 kg, Index Weight = 1), Tomatoes - Retail (1 kg, Index Weight = 1),
Eggs - Retail (1 Unit, Index Weight = 12), Milk (Powder) - Retail (1 kg, Index Weight = 1),
Dates - Retail (1 kg, Index Weight = 1), Cheese (Local) - Retail (1 kg, Index Weight = 1),
Lentils - Retail (1 kg, Index Weight = 1)
 Haiti
 Iraq
```

Note: Weights have been rounded to two digits. Note that for index with one item measured in 1 kg and one item measured in 100 kg, the latter has only a weight of 0.01, which has the same effect as scaling the latter unit of measurement to 1 kg and weighting both equally.

Source: The statistics have been prepared by the author for this paper based on price data selected using end-of-August (2021) food price data from World Food Program.

Table A3—Summary of food price index components (continued).

```
Country
                                        Index Component
Lao PDR
                                        Oil (Soybean) - Retail (1 L, Index Weight = 1), Rice (Glutinous, Second Quality) - Retail (1 kg, Index Weight = 1),
                                        Sugar (Brown) - Retail (1 kg, Index Weight = 1), Eggs - Retail (1 Unit, Index Weight = 12),
                                        Garlic (Small) - Retail (1 kg, Index Weight = 1)
                                        Bread (Pita) - Retail (1 kg, Index Weight = 1), Oil (Sunflower) - Retail (5 L, Index Weight = 0.2),
Rice (Imported, Egyptian) - Retail (1 kg, Index Weight = 1), Sugar (White) - Retail (1 kg, Index Weight = 1),
Lebanon
                                        Wheat Flour - Retail (1 kg, Index Weight = 1), Beans (White) - Retail (1 kg, Index Weight = 1), Eggs - Retail (30 pcs, Index Weight = 0.33), Milk (Powder) - Retail (900 G, Index Weight = 1.11),
                                        Cabbage - Retail (1 kg, Index Weight = 1), Cucumbers (Greenhouse) - Retail (1 kg, Index Weight
                                        Tomatoes (Paste) - Retail (1.3 kg, Index Weight = 0.77), Bulgur (Brown) - Retail (1 kg, Index Weight = 1), Cheese (Picon) - Retail (160 G, Index Weight = 6.25), Chickpeas - Retail (1 kg, Index Weight = 1),
                                        Lentils (Red) - Retail (1 kg, Index Weight = 1)
Oil (Palm) - Retail (1 Gallon, Index Weight = 0.26), Rice (Imported) - Retail (50 kg, Index Weight = 0.02),
Liberia
                                        Cassava (Fresh) - Retail (50 kg, Index Weight = 0.02)
                                       Cassava (Friend) - Retail (1 kg, Index Weight = 1), Beans (Niebe) - Retail (1 kg, Index Weight = 1),
Groundnuts (Shelled) - Retail (1 kg, Index Weight = 1), Maize - Retail (1 kg, Index Weight = 1),
Millet - Retail (1 kg, Index Weight = 1), Sorghum - Retail (1 kg, Index Weight = 1),
Oil (Vegetable, Local) - Retail (1 L, Index Weight = 1), Rice (Imported) - Retail (1 kg, Index Weight = 1),
Mali
Mozambique
                                        Sugar (Brown, Local) - Retail (1 kg, Index Weight = 1), Wheat Flour (Local) - Retail (1 kg, Index Weight = 1), Groundnuts (Small, Shelled) - Retail (1 kg, Index Weight = 1), Maize (White) - Retail (1 kg, Index Weight = 1), Maize (Maize Meal (White, Without Bran) - Retail (1 kg, Index Weight = 1)
                                        Oil (Palm) - Retail (1 L, Index Weight = 1), Pulses - Retail (1 kg, Index Weight = 1),
Rice (Low Quality) - Retail (1 kg, Index Weight = 1)
Myanmai
Niger
                                        Rice (Imported) - Retail (1 kg, Index Weight = 1), Maize - Retail (1 kg, Index Weight = 1),
                                       Millet - Retail (1 kg, Index Weight = 1), Sorghum - Retail (1 kg, Index Weight = 1)

Bread - Retail (1 Unit, Index Weight = 1), Oil (Palm) - Retail (750 ML, Index Weight = 1.33),

Rice (Imported) - Wholesale (50 kg, Index Weight = 0.02), Groundnuts (Shelled) - Wholesale (100 kg, Index Weight = 0.01),

Maize (White) - Wholesale (100 kg, Index Weight = 0.01), Millet - Wholesale (100 kg, Index Weight = 0.01),
Nigeria
                                        Native (White) - Wholesale (100 kg, Index Weight = 0.01), Similer - Wholesale (100 kg, Index Weight = 0.01), Sorghum (White) - Wholesale (100 kg, Index Weight = 0.01), Bananas - Retail (1.3 kg, Index Weight = 0.77), Maize Flour - Retail (1.3 kg, Index Weight = 0.77), Cassava Meal (Gari, Yellow) - Wholesale (100 kg, Index Weight = 0.01), Cowpeas (White) - Wholesale (100 kg, Index Weight = 0.01), Eggs - Retail (30 pcs, Index Weight = 0.33),
                                        Milk - Retail (20 G, Index Weight = 50), Oranges - Retail (400 G, Index Weight = 2.5),
Watermelons - Retail (2.1 kg, Index Weight = 0.48), Gari (White) - Wholesale (100 kg, Index Weight = 0.01)
                                        Oil (Vegetable, Imported) - Retail (1 L, Index Weight = 1), Rice (Imported) - Retail (1 kg, Index Weight = 1),
Somalia
                                        Maize (White) - Retail (1 kg, Index Weight = 1), Sorghum (Red) - Retail (1 kg, Index Weight = 1) Oil (Vegetable) - Retail (1 kg, Index Weight = 1), Wheat Flour - Retail (1 kg, Index Weight = 1),
South Sudan
                                        Beans (Red) - Retail (1 kg, Index Weight = 1), Groundnuts (Shelled) - Retail (1 kg, Index Weight = 1), Maize (White) - Retail (3.5 kg, Index Weight = 0.29), Millet (White) - Retail (3.5 kg, Index Weight = 0.29),
                                        Sorghum (White, Imported) - Retail (3.5 kg, Index Weight = 0.29), Sesame - Retail (3.5 kg, Index Weight = 0.29)
                                        Wheat - Wholesale (90 kg, Index Weight = 0.01), Millet - Retail (3.5 kg, Index Weight = 0.29), Sorghum - Retail (3 kg, Index Weight = 0.33)
Sudan
                                       Soriginii - Retail (1.1 kg, Index Weight = 0.03)

Bread (Bakery) - Retail (1.1 kg, Index Weight = 0.91), Oil - Retail (1 L, Index Weight = 1),
Rice - Retail (1 kg, Index Weight = 1), Sugar - Retail (1 kg, Index Weight = 1),
Wheat Flour - Retail (1 kg, Index Weight = 1), Beans (White) - Retail (1 kg, Index Weight = 1),
Tomatoes - Retail (1 kg, Index Weight = 1), Eggs - Retail (30 pcs, Index Weight = 0.33),
Dates - Retail (1 kg, Index Weight = 1), Yogurt - Retail (1 kg, Index Weight = 1),
Syrian Arab Rep.
                                        Bulgur - Retail (1 kg, Index Weight = 1), Cheese - Retail (1 kg, Index Weight = 1), Chickpeas (Yellow) - Retail - Retail (1 kg, Index Weight = 1), Lentils - Retail (1 kg, Index Weight = 1),
                                       Oil (Vegetable) - Retail (1 Packet, Index Weight = 2), Parsley - Retail (1 Packet, Index Weight = 1), Rice (Imported) - Retail (1 kg, Index Weight = 1), Sugar - Retail (1 kg, Index Weight = 1), Wheat - Retail (1 kg, Index Weight = 1),
Yemen, Rep.
                                       Wheat Flour - Retail (1 kg, Index Weight = 1), Beans (Kidney Red) - Retail (1 kg, Index Weight = 1), Onions - Retail (1 kg, Index Weight = 1), Potatoes - Retail (1 kg, Index Weight = 1),
                                        Omntoes - Retail (1 kg, Index Weight = 1), Eggs - Retail (1 kg, Index Weight = 12),
Peas (Yellow, Split) - Retail (1 kg, Index Weight = 1), Lentils - Retail (1 kg, Index Weight = 1)
```

Note: Weights have been rounded to two digits. Note that for index with one item measured in 1 kg and one item measured in 100 kg, the latter has only a weight of 0.01, which has the same effect as scaling the latter unit of measurement to 1 kg and weighting both equally.

Source: The statistics have been prepared by the author for this paper based on price data selected using end-of-August (2021) food price data from World Food Program.

Table A4—Summary of estimation results.

Country	Avg. inflation	Max. draw-down	Avg. volatility	CV-score
Afghanistan	5.76%	-44.02%	8.69%	0.87
Burkina Faso	4.08%	-37.74%	15.25%	0.77
Burundi	3.90%	-27.74%	13.26%	0.77
Cameroon	1.25%	-20.40%	7.01%	0.95
Central Afr. Rep.	1.07%	-29.92%	19.36%	0.75
Chad	4.17%	-45.92%	19.02%	0.69
Congo, Rep.	1.56%	-24.17%	11.76%	0.87
Congo, Dem. Rep.	7.63%	-16.67%	8.64%	0.87
Gambia, The	4.92%	-15.31%	10.52%	0.71
Guinea-Bissau	1.62%	-18.56%	14.51%	0.81
Haiti	7.60%	-34.82%	11.98%	0.82
Iraq	-1.07%	-25.58%	5.12%	0.90
Lao PDR	0.83%	-3.44%	1.56%	0.93
Lebanon	20.35%	-21.96%	14.07%	0.92
Liberia	11.67%	-17.95%	10.57%	0.93
Mali	2.46%	-23.97%	8.63%	0.84
Mozambique	9.10%	-29.33%	9.40%	0.84
Myanmar	1.82%	-38.76%	11.83%	0.74
Niger	4.03%	-23.71%	9.84%	0.87
Nigeria	7.82%	-20.89%	7.44%	0.92
Somalia	5.54%	-48.67%	12.95%	0.83
South Sudan	43.03%	-29.55%	25.11%	0.87
Sudan	47.16%	-30.24%	18.93%	0.87
Syrian Arab Rep.	35.25%	-23.49%	17.01%	0.88
Yemen, Rep.	9.55%	-23.84%	14.13%	0.78

Note: The first three columns respectively report average annualized inflation, maximum draw-down, and average annual realized volatility in percentages. The final column reports the cross-validated confidence score that ranges from 0 to 1 for the final food price index using the calculations from the paper. Additional cross-validation statistics can be found on the World Bank Data Catalog page where a live version of the data base is maintained.

Source: The statistics have been prepared by the author for this paper based on end-of-August (2021)

Source: The statistics have been prepared by the author for this paper based on end-of-August (2021) food price data from World Food Program.

Table A5—Linear decomposition of CV-score.

	$Dependent\ variable:\ CV\text{-}score$
	OLS
log Number of markets	-0.034
	(0.028)
log Number of food items	0.021
	(0.032)
Number of markets per food item	0.003
	(0.005)
Data completeness	-0.019
	(0.092)
Inflation rate	0.450**
	(0.138)
Volatility	-1.305**
	(0.361)
Max. draw-down	0.024
	(0.138)
Constant	1.018***
	(0.098)
Observations	25
\mathbb{R}^2	0.627
Adjusted R ²	0.474
Residual Std. Error	0.052 (df = 17)
F Statistic	3.792 (df = 7; 17)

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Simple linear regression estimates that decompose the cross-validation performance score into key characteristics of the imputation problem. Percentage covariates are modeled as numeric digits of the same unit as the dependent variable, e.g. a monthly volatility of 2% enters the regression as a value of 0.02. The simple results highlight that prediction performance is not significantly related to data dimensions of the problem, instead high volatility and high inflation are determinants of imputation accuracy. Jointly, these two variables explain almost half of the in-sample variation. The signs of the coefficients suggest that when volatility increases, prediction performance deteriorates and data collection becomes more important. When the inflation trend is strong relative to volatility, then the price trend is clearer in the data and imputation accuracy increases, suggesting that robust inflation tracking is possible even when little ground truth data is available.

Source: The statistics have been prepared by the author for this paper based on end-of-August (2021) food price data from World Food Program.

32 $ANDR \acute{E} E$ DECEMBER - 2021

Table A6—Realized Annualized Food Price Inflation.

Country	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021*
Afghanistan	38.04	55.36	-29.02	2	4.44	18.49	0.95	-1.14	1.31	-0.14	1.29	1.56	1.17	14.09	2.8
Burkina Faso	27.63	-11.96	13.69	-1.57	16.35	-5.15	-8	-0.68	-1.02	1.39	14.46	-9.27	-12.64	21.61	47.93
Burundi	7.31	1.31	15.57	3.53	-6.19	23.57	-2.75	-9.7	21.06	23.54	-3.83	-11.18	15.5	1.13	-16.2
Cameroon	3.12	-0.51	2.1	1.58	1.82	4.82	-2.3	-2.62	-1.9	-6.23	5.15	-3.24	0.05	13.91	7.62
Central Afr. Rep.	2.02	11.31	-3.27	3.63	-0.39	-4.79	-1.41	-12.37	38.04	-21.28	-6.38	5.8	2.24	4.61	14.03
Chad	-11.1	62.25	-7.59	-26.54	57.55	-5.1	7.08	1.41	0.09	-27.3	18.13	-12.55	0.85	25.46	31.07
Congo, Rep.				4.08	-5.65	9.85	-5.97	14.92	12.94	-14.66	14.78	-14.77	-1.13	6.35	1.36
Congo, Dem. Rep.	10.64	21.35	7.83	-1.64	28.96	2.14	-8.04	-0.03	-2.52	12.94	24.37	3.82	11.56	8.26	0.6
Gambia, The	1.85	16.01	-8.69	2.28	7.44	5.16	2.04	3.08	1.2	18.81	-11.58	20.73	0.65	24.28	-5.77
Guinea-Bissau									-3.33	4.83	-12.24	14.76	-2.75	-5.84	28.79
Haiti	2.24	28.17	-17.21	6.06	3.6	11.58	-11.57	1.71	30.81	3.15	-0.5	5.46	45.86	1.19	25.98
Iraq					-0.35	-0.17	1.34	-6.99	0.28	-2.93	-7.45	-8.21	1.84	8.07	7.08
Lao PDR	0.22	0.31	0.39	0.68	0.58	0.16	1.6	2.11	-0.49	-1.02	0.71	0.6	5.88	0.73	-0.23
Lebanon						7.32	1.12	-1.96	-14.64	2.13	4.74	0.43	18.01	131.39	237.14
Liberia	20.5	26.59	3	9.86	18.7	12.72	-8.2	11.18	-7.25	19.8	6.11	33.26	10.31	15.92	9.83
Mali	6.54	5.1	1.11	-2.43	19.5	-4.07	-5.42	-1.75	2.06	1.13	2.76	2.76	-5.3	4.51	21.2
Mozambique	36.47	25.8	-9.16	23.36	6.11	-0.14	-3.3	-3.5	19.9	74.51	-26.05	-0.85	11.3	12.88	0.55
Myanmar	6.67	-8.33	-19.68	12.4	-5.37	-8.49	4.57	-6.89	16.43	13.26	-2.25	2.31	4.81	3.25	45.03
Niger	3	33.49	-0.12	-7.91	17.46	2.71	1.22	-10.76	-1.91	3.55	4.58	-5.93	-1.47	10.15	31.23
Nigeria					24.31	2.84	4.28	-2.01	1.6	37.98	2.24	-7.63	-2.21	24.84	27.67
Somalia	46.67	54.55	-18.55	12.62	2.83	-24.26	1.18	6.74	-5.3	9.58	-1.49	-1.4	5.52	11.35	9.6
South Sudan	-2.23	21.42	18.47	15.26	69.16	2.54	-9.32	15.87	120.68	451.01	63.75	33.22	31.43	63.63	20.17
Sudan	23.83	57.15	28.14	1.79	38.1	29.91	31.92	23.1	3.71	13.96	71.59	119.75	62.1	271.37	72.13
Syrian Arab Rep.					13.32	11.14	80.45	0.27	66.91	19.73	-11.96	-12.1	36.18	245.09	74.74
Yemen, Rep.		8.77	10.33	5.13	19.71	-13.76	2.91	1.09	10.88	6.11	13.3	30.1	4.37	31.81	10.96
FCS	11.4	19.52	-1.71	2.71	13.04	2.66	1.91	0.55	9.83	13.63	4.82	5.14	8.63	27.46	22.65

Note: Figures are annualized month-on-month price changes in percentages. Monthly price data is maintained at the World Bank Data Catalog page associated with the paper. FCS is a geometric average of country rates. The 2021^* figures are based on end-of-August data. Source: Statistics have been prepared by the author for this paper.

TABLE A7—REALIZED ANNUALIZED FOOD PRICE VOLATILITY.

Country	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021*
Afghanistan	7.94	20.49	11.42	9.59	4.84	6.09	2.39	2.37	2.65	1.98	2.04	2.58	3.23	5.82	3.19
Burkina Faso	13.92	23.34	23.37	13.05	10.97	19.97	15.63	15.99	11.97	12.17	12.2	17.68	8.08	10.19	5.67
Burundi	9.63	10.22	10.03	11.8	17.04	14.66	12.77	8.7	22.07	17.99	8.18	12.22	15.08	9.26	7.38
Cameroon	2.44	1.91	1.68	3.55	5.09	12.06	11.45	6.1	6.92	12.87	11.03	8.99	6.44	7.62	4.34
Central Afr. Rep.	34.06	26.11	17.97	23.4	19.73	14.48	26.54	16.88	25.05	9.95	19.78	12.06	14.28	12.02	17.3
Chad	14.16	24.52	26.96	23.26	14.74	21.57	19.76	19.09	15.07	15.07	20.04	17.04	11.1	15.32	9.53
Congo, Rep.				2.79	9.75	10.34	17.2	5.97	19.66	13	12.23	14.23	16.59	9.01	3.36
Congo, Dem. Rep.	2.32	8.23	15.22	9.79	4.26	3.33	5.14	5.22	3.86	7.83	14.4	8.19	7.47	9.41	5.55
Gambia, The	10.81	7.43	7.48	4.12	5.44	4.14	8.76	4.51	7.84	13.54	13.98	13.19	16.68	14.97	6.57
Guinea-Bissau									10.59	12.58	13.52	13.07	11.35	19.18	15.73
Haiti	9.45	21.15	9.77	20.62	10.8	10.22	6.59	3.32	6.27	7.03	3.28	2.48	10.88	20.32	5.79
Iraq					3.18	3.88	3.3	6.13	4.75	3.35	6.4	4.54	5.7	7.01	5.87
Lao PDR	0.82	0.92	0.81	0.94	0.92	1.33	1.09	2.2	0.78	0.95	0.4	1.65	2.69	3.12	1.98
Lebanon						3.71	3.39	3.08	4.86	4.54	6.42	1.79	8.5	13.48	5.29
Liberia	9.73	10.64	19.55	13.73	7.72	5.77	10.02	12.51	8.08	8.58	10.69	16.01	9.29	9.29	7.32
Mali	7.55	12.67	10.76	6.78	4.56	10.78	5.23	6.1	9.07	5.69	6.46	10.83	6.2	6.46	2.68
Mozambique	9.41	5.38	10.92	7.47	5.97	5.35	4.38	4.99	9.88	12.14	7.89	4.07	6.75	7.22	8.62
Myanmar	5.16	17.45	24.57	15.26	15.03	15.15	8.47	9.02	11.94	5.55	6.58	8.09	5.51	6.7	9.06
Niger	4.96	15.12	10.96	11.75	6.13	11.91	8.6	4.58	4.92	11.06	13.73	8.68	3.75	13.97	5.71
Nigeria					13.27	6.58	4.91	6.31	9.18	7.44	8.44	4.1	5.74	5.96	4.63
Somalia	8.66	32.58	12.66	10.19	15.94	10.86	7.03	7.35	7.15	9.25	7.5	4.76	2.35	10.41	4.12
South Sudan	10.19	6.08	26.43	11.19	25.72	29.82	16.24	28.53	26.45	27.4	24.64	34.47	14.02	12.66	28.83
Sudan	14.5	23.13	11.24	16.94	11.55	22.91	12.48	23.65	11.93	11.13	20.34	19.55	16.41	14.93	11.1
Syrian Arab Rep.					7.26	9.8	22.19	15.21	7.39	13.55	4.96	8.89	14.47	11.03	21.09
Yemen, Rep.		3.77	18.45	15.28	23.15	11.87	7.84	11.28	21.82	15.38	7.41	17.58	7.78	9.51	5.85
FCS	9.55	13.92	13.99	11.4	10.38	10.9	9.87	9.34	10.6	10.27	10.34	10.44	9.12	10.52	8.1

Note: Figures are annualized standard deviations of month-on-month price changes in percentages. Monthly price data is maintained at the World Bank Data Catalog page associated with the paper. FCS is a geometric average of country rates. The 2021* figures are based on end-of-August data. Source: Statistics have been prepared by the author for this paper.



Figure A1. Estimated food prices, inflation and intra-month volatilities in Yemen.

Note: Price of food baskets in local currency, local-market average, January 2015 =1. The top charts shows Open, High, Low and Close price estimates of the total food basket price. The food basket consists of a 24 market-average of retail prices of Cooking Oil (Vegetable, 1 L), Rice (Imported, 1 kg), Sugar (1 kg), Wheat (1 kg), Wheat Flour (1 kg), Beans (Kidney Red, 1 kg), Onions (1 kg), Potatoes (1 kg), Tomatoes (1 kg), Eggs (12 Units), Peas (Yellow, Split, 1 kg) and Lentils (1 kg). The bottom chart shows monthly food price inflation as a year-on-year percentage increase in Close prices. Source: Figure prepared by the author for this paper. A Live version of all graphs, including graphs at the subnational level, are maintained in the on line data repository.

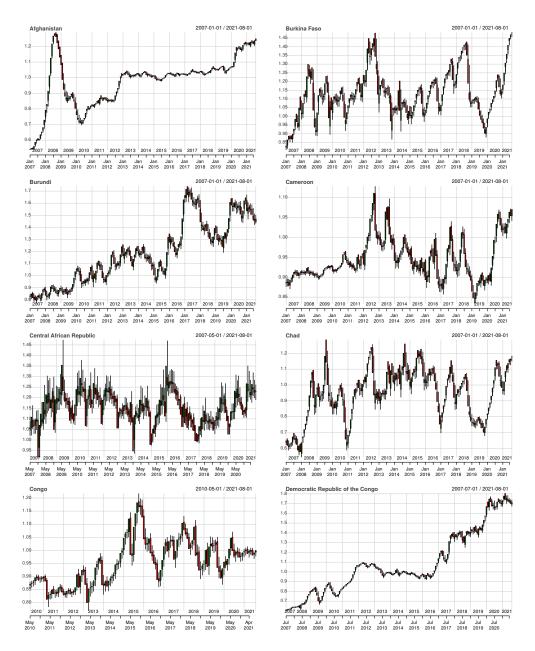


FIGURE A2. ESTIMATED FOOD PRICES AND INTRA-MONTH VOLATILITIES IN AFGHANISTAN, BURKINA FASO, BURUNDI, CAMEROON, CENTRAL AFRICAN REPUBLIC, CHAD, THE REPUBLIC OF THE CONGO AND THE DEMOCRATIC REPUBLIC OF CONGO.

Note: Price of food baskets in local currency, local-market average, January 2015 =1.

Source: Figure prepared by the author for this paper. A Live version of all graphs, including graphs at the subnational level, are maintained in the online data repository.

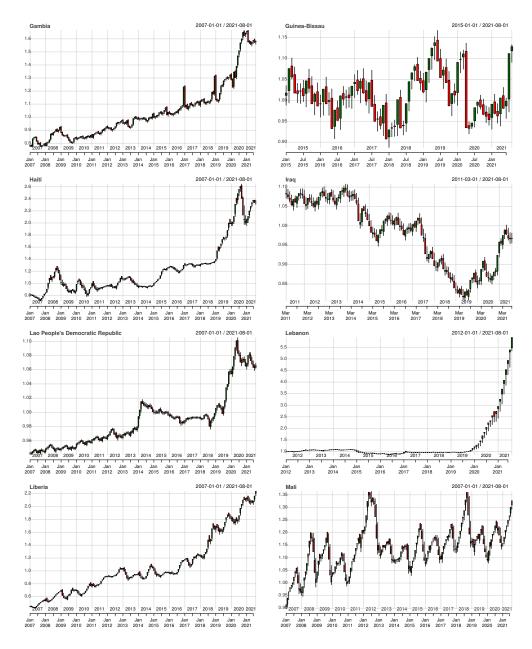


FIGURE A3. ESTIMATED FOOD PRICES AND INTRA-MONTH VOLATILITIES IN THE GAMBIA, GUINEA-BISSAU, HAITI, IRAQ, LAO PEOPLE'S DEMOCRATIC REPUBLIC, AND LEBANON.

Note: Price of food baskets in local currency, local-market average, January 2015 =1.

Source: Figure prepared by the author for this paper. A Live version of all graphs, including graphs at the subnational level, are maintained in the on line data repository.

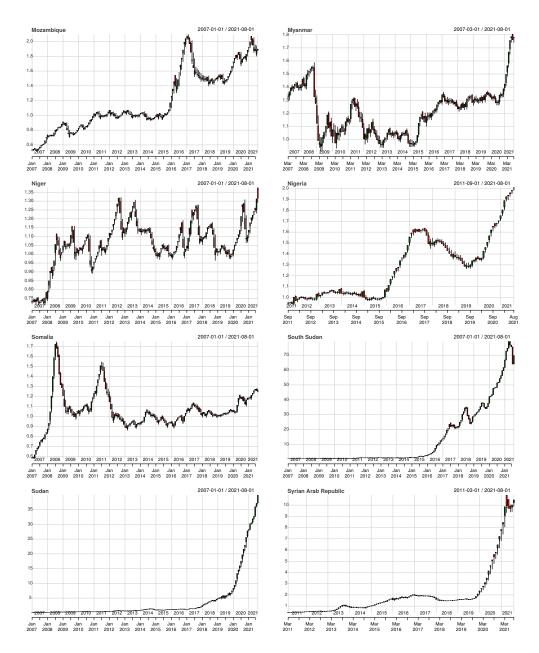


FIGURE A4. ESTIMATED FOOD PRICES AND INTRA-MONTH VOLATILITIES IN MOZAMBIQUE, MYANMAR, NIGER, NIGERIA, SOMALIA, SUDAN, SOUTH SUDAN AND THE SYRIAN ARAB REPUBLIC.

Note: Price of food baskets in local currency, local-market average, January 2015 =1.

Source: Figures prepared by the author for this paper. A Live version of all graphs, including graphs at the subnational level, are maintained in the on line data repository.

MATHEMATICAL APPENDIX

B1. Prediction strategy

The imputation strategy used here is based on an adaption of the Multiple Imputations by Chained Equations (see, van Buuren and Groothuis-Oudshoorn (2011); van Buuren (2012); Murray (2018)). In particular, the original approach has been adapted to allow specify the stochastic properties of the initialization and keep track of cross-validation throughout the imputation process. The adapted code base has been made available, see the links provided in the footnotes of the introduction of the main paper. The code base has also been adapted to allow parallelization of computation. Apart from these practical modifications, the sketch of the algorithm is relatively standard. The algorithm produces multiple likely data sets, which are then pooled into a single imputation. The number of imputations is M=5, and the m-th imputed price data set is $P^{(m)}$ where $m \in (1, \ldots, M)$. Let $P^{-d} = (P^1, \ldots, P^{d-1}, P^{d+1}, P^D)$ denote the collection of the d-1 variables in P except P^d . Note that for a given P^d , the set P^{-d} may itself be incomplete, $(P^1, \ldots, P^{d-1}, P^{d+1}, P^D)$ may be correlated with one another, the relationship between P^d and P^{-d} could be complex, and P^d may also depend on P^d of the another, the relationship between P^d and P^{-d} could be complex, and P^d may also depend on P^d observed random sample from the P^d variate multivariate distribution P^d and that the multivariate distribution of P^d is completely specified by the unknown parameter vector P^d . Thus, the objective is to obtain estimates for P^d . The algorithm estimates a posterior distribution of P^d by iteratively sampling from the conditional distributions

(B1)
$$\begin{aligned} \mathcal{P}(P_1|P_{-1}, X_h; \boldsymbol{\theta}_1) \\ \vdots \\ \mathcal{P}(P_d|P_{-d}, X_h; \boldsymbol{\theta}_d) \end{aligned}$$

The parameters $\theta_1, \ldots, \theta_d$ are specific to the respective conditional densities. Starting from a draw from the marginal distributions, in the current application modeled using univariate time series methods, the *i*-th iteration of chained equations is a Gibbs sampler that successively draws

(B2)
$$\theta_{1}^{*(i)} \sim \mathcal{P}\left(\theta_{1} | P_{1,\text{obs}}, P_{2}^{(i-1)}, \dots, P_{D}^{(i-1)}, X\right) \\ P_{1}^{*(i)} \sim \mathcal{P}\left(P_{1} | P_{1,\text{obs}}, P_{2}^{(i-1)}, \dots, P_{D}^{(i-1)}, X, \boldsymbol{\theta}_{1}^{*(i)}\right) \\ \vdots \\ \boldsymbol{\theta}_{D}^{*(i)} \sim \mathcal{P}\left(\boldsymbol{\theta}_{D} | P_{D,\text{obs}}, P_{1}^{(i)}, \dots, P_{D-1}^{(i)}, X\right) \\ P_{D}^{*(i)} \sim \mathcal{P}\left(P_{D} | P_{D,\text{obs}}, P_{1}^{(i)}, \dots, P_{D}^{(i)}, X, \boldsymbol{\theta}_{D}^{*(i)}\right)$$

where $P_d^{(i)} = (P_{d,obs}, P_d^{*(i)})$ is the *i*-th imputation of price d at iteration i. The imputations of the previous iteration $P_d^{*(i-1)}$ enter the next imputation $P_d^{*(i)}$ through the other price variables $P_{-d}^{*(i)}$. At each iteration, $P_d^{*(i-1)}$ can also be used to generate synthetic cases by adding a random draw of conditional expectations for missing entries to the dependent side of next regression on the same price variable. The process is iterated 8 times, a stopping criterion can be guided by keeping track of prediction validation criteria across i. The imputations and their updates here are done by drawing the conditional means from the posterior distribution of penalized linear regression or cubist regressions. Note that due to the randomness in the cubist algorithm itself, the randomness of the initialization, and the randomness in the synthetic data, there remains substantial randomness in successive iterations that allows the sequence to visit a large variety of likely prediction models for the missing values. For instance, since the models used for prediction can extrapolate beyond the range of training data values to generate synthetic cases with different properties than observed data, and because there are additional sources of randomness such as related to parameter tuning and random components of the cubist model, each iteration may weaken, amplify, or change local correlations between the columns in the data in the next imputation iteration. Hence, while a pre-determined linear regression specification approach artificially amplifies the relations between the columns of the data by reinforcing its own learning pattern throughout the iterations, the stochastic method allows the sequence to find prediction models for the missing values

ANDR'EE DECEMBER - 2021

beyond what seemed likely from initially observed data alone.

B2. Ensemble

Let Q be the quantity of interest, simple linear difference calculation using fixed-weight averages of P. The ultimate goal of the multiple imputation strategy will be to obtain an estimate \hat{Q}

(B3)
$$\mathbb{E}(\hat{Q}|\mathbf{P}) = Q,$$

 ${\bf P}$ denoting the *true* price index population. Since ${\bf P}$ is unknown, Q is unknown. The amount of uncertainty in the estimate \hat{Q} thus depends on what is known about ${\bf P}_{\rm mis}$. Since we can only recreate it with uncertainty based on information in ${\bf P}_{\rm obs}$, the idea is to summarize a distribution of Q under varying estimates of ${\bf P}_{\rm mis}$. In other words, the possible functions Q given what has been observed in ${\bf P}_{\rm obs}$ have a posterior distribution $\mathcal{P}(Q|{\bf P}_{\rm obs})$ which in turn can be decomposed into two parts

(B4)
$$\mathcal{P}(Q|\mathbf{P}_{\text{obs}}) = \int \mathcal{P}(Q|\mathbf{P}_{\text{obs}}, \mathbf{P}_{\text{mis}}) \mathcal{P}(\mathbf{P}_{\text{mis}}|\mathbf{P}_{\text{obs}}) d\mathbf{P}_{\text{mis}}$$

In this, $\mathcal{P}(Q|\mathbf{P}_{\text{obs}}, \mathbf{P}_{\text{mis}})$ is the posterior distribution of inflation in the hypothetically complete price data and $\mathcal{P}(\mathbf{P}_{\text{mis}}|\mathbf{P}_{\text{obs}})$ is the posterior distribution of the missing price data given the observed price data. Suppose that $\mathcal{P}(\mathbf{P}_{\text{mis}}|\mathbf{P}_{\text{obs}})$ is used to draw various likely price data sets for \mathbf{P}_{mis} , denoted as $\dot{\mathbf{P}}_{\text{mis}}$. Then, associated inflation Q can be calculated from $(\dot{\mathbf{P}}_{\text{mis}},\mathbf{P}_{\text{obs}})$. By repeating this process multiple times, one can obtain the posterior distribution for Q and equation B4 shows that Q than equals the expectation over all draws:

(B5)
$$\mathcal{P}(Q|\mathbf{P}_{\text{obs}}) = \mathbb{E}(\mathbb{E}([Q|\mathbf{P}_{\text{obs}}, \mathbf{P}_{\text{mis}}]|\mathbf{P}_{\text{obs}})),$$

which suggests that when $\hat{Q}^{(m)}$ is the estimated model using the *m*-th imputation, then the combined model using all the imputations is equal to the ensemble estimate

(B6)
$$\hat{Q} = \frac{1}{M} \sum_{m=1}^{M} \hat{Q}^{(m)} = Q \left(\frac{1}{M} \sum_{m=1}^{M} \hat{\mathbf{P}}^{(m)}; \cdot \right),$$

where the second equality is due to the linearity in the simple linear difference formulation of Q.

B3. Performance against the unobserved prediction objective

Let p_{it}^d be a possible observed price quote for food item $d \in (1,\ldots,D)$ that may be observed at location $i \in (1,\ldots,N)$ and time $t \in (1,\ldots,T)$. Let $\mathbf{p}^d \ni p_{it}^d$ be the possibly incomplete vector of prices for food item d generated by stacking all $N \times T$ entries. $P = \mathbf{p}^1,\ldots,\mathbf{p}^D$ is the $NT \times D$ matrix that collects all the possible price points and consists of observed and missing parts, $P_{\text{obs}} = (\mathbf{p}_{\text{obs}}^1,\ldots,\mathbf{p}_{\text{obs}}^D)$ and $P_{\text{mis}} = \mathbf{p}_{\text{mis}}^1,\ldots,\mathbf{p}_{\text{mis}}^D$. Suppose the true prices are generated by some process that is only partially observed and with error. In particular, for every individual price signal $d \in (1,\ldots,D)$, the focus is on a T-period sequence $\{\pi_t^d\}_{t=1}^T$ that is a subset of the realized path of the stochastic sequence $\pi^d := \{\pi_t^d\}_{t \in \mathbb{Z}}$. Suppose that $\{\pi_t\}_{t=1}^T$ is unobserved, but that there is an observed sequence

(B7)
$$\mathbf{p}_{\text{obs}}^d := \{ \mathbf{p}_{t \text{ obs}}^d = \mathcal{M}^d(\boldsymbol{\pi}_t^d) \}.$$

In this equation, \mathcal{M}^d is a function that describes how price data of commodity d is measured. It may produce measurement error in the form of additive outliers as well as data gaps. The important distinction is thus that true prices $\boldsymbol{\pi}^d$ are assumed to exist but it is only possible to partially observe $\mathbf{p}_{\text{obs}}^d$ by surveying, inadvertently introducing errors and missing entries. The sequence $\mathbf{p}^d = (\mathbf{p}_{\text{obs}}^d, \mathbf{p}_{\text{mis}}^d)$ can thus be split in missing values $\mathbf{p}_{\text{mis}}^d$ and possibly contaminated observations $\mathbf{p}_{\text{obs}}^d$. The two-fold aim is to proxy $\boldsymbol{\pi}_{\text{obs}}^d$ with an estimate $\hat{\mathbf{p}}_{\text{obs}}^d$ by filtering $\mathbf{p}_{\text{obs}}^d$ from outliers, and proxy $\boldsymbol{\pi}_{\text{mis}}^d$ by estimating

 $\hat{\mathbf{p}}_{\mathrm{mis}}^d$ by filling in entries for $\mathbf{p}_{\mathrm{mis}}^d$ based on the information contained in $\hat{\mathbf{p}}_{\mathrm{obs}}^d$. Since the true targets are unobserved, a direct criterion is difficult to establish but the objective can be summarized as minimizing the divergence $\|\hat{\mathbf{p}}^{d=1,\dots D} - \boldsymbol{\pi}^{d=1,\dots D}\|$, in turn estimated with an L_1 -norm based metric for the prediction function that generates $\hat{\mathbf{p}}_{\mathrm{mis}}^d$, validated by the out-of-sample predictions it makes for the outlier-filtered $\hat{\mathbf{p}}_{\mathrm{obs}}^d$ that serves as a proxy for $\boldsymbol{\pi}_{\mathrm{obs}}^d$.

B4. Intra-month estimates

The price-level estimates are accompanied by intra-month price range estimates represented as an Open-High-Low-Close time series.

(B8)
$$\begin{bmatrix} \hat{O} \\ \hat{H} \\ \hat{L} \\ \hat{C} \end{bmatrix}_{t} = \begin{bmatrix} \mathbb{E}P_{t} | \mathcal{F}_{t-1} \\ P_{t-1} + \mathbb{E}\Delta_{\alpha>0.50} P_{t} | \mathcal{F}_{t-1} \\ P_{t-1} + \mathbb{E}\Delta_{\alpha<0.50} P_{t} | \mathcal{F}_{t-1} \\ P_{t} \end{bmatrix}$$

where P is the imputed price series and $\mathbb{E}\Delta_{\alpha}$ is the expected change in the α -percentile cases. The combined results can be plotted on a candle chart. The majority of price action can then be assumed to have occurred within the body of the candles, while the wicks indicate the average price of respectively the highest 50% of intra-month prices, and the average price of the lowest %50 of intra-month prices.

The first three quantities in equation B8 are estimated by modeling the time-varying distribution of the month-on-month inflation process as an autoregressive moving average process with fractionally integrated generalized autoregressive conditional heteroskedasticity (ARMA-fiGARCH) following Baillie et al. (1996). This is a time-varying density

(B9)
$$F_t = (\mu_t, \sigma_t, \vartheta)$$

where μ_t is a conditional mean process defined as an ARMA(p,q) process

(B10)
$$\mu_t = c + \sum_{j=1}^p \phi_j \mu_{t-j} + \sum_{j=1}^q \theta_j \varepsilon_{t-1} + \varepsilon_t,$$

and the conditional variance is specified as a fractionally integrated GARCH process of order (p, d, q) and ϑ is a vector of remaining parameters of the distribution. The conditional variance is defined as follows. First, let the standard GARCH(p, q) process be defined as

(B11)
$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2$$

with σ_t^2 as the conditional variance, ω an intercept, and L the back-shift operator with $\alpha(L) = \sum_{j=1}^q \alpha_j L^j$ and $\beta(L) = \sum_{j=1}^p \beta_j L^j$. This model has an ARMA representation of the squared process:

(B12)
$$(1-L)\phi(L)\varepsilon_t^2 = [1-\alpha(L)-\beta(L)]\varepsilon_t^2 = \omega + [1-\beta(L)](\varepsilon_t^2 - \sigma_t^2)$$

with $\phi(L) = \sum_{j}^{\max(p,q)-1} = \phi_j L_j$. The fractionally integrated GARCH is obtained by replacing the back-shift operator (1-L) with a truncated fractional difference operator

(B13)
$$(1-L)^d = \sum_{k=0}^{K=1000 \sim \infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} L^k.$$

Ignoring the approximation error due to the truncation, at d=0 the model equals the standard GARCH in which volatility shocks decay at an exponential rate. Similarly, when d=1, the AR polynomial of the GARCH has a unit root and the model equals the integrated GARCH in which shocks persist forever. When there are level shifts in volatility process, an integrated GARCH usually better describes the data

 $ANDR\'{E}$ DECEMBER - 2021

than the standard GARCH. Shifts in the volatility process may stem for example from price controls. However, the unconditional variance is undefined in this model which is theoretically difficult to conceive. The fractionally integrated GARCH that results under values 0 < d < 1 allows the GARCH process to have hyperbolic memory in the volatility process such that the volatility process shifts gradually. Such long-memory volatility features have been observed widely in both agricultural commodities (Chang et al., 2012) and general inflationary shocks (Baillie et al., 2002).

Finally, the log likelihood requires specifying the remainder parameters in λ . The model is estimated using the Generalized Error Distribution:

(B14)
$$G(z;\vartheta) = \frac{\nu e^{-0.5 \left| \frac{z - \mu}{\lambda} \right|^{\nu}}}{2^{1+\nu^{-1}} \lambda \Gamma(\nu^{-1})}$$

with $\vartheta=(\mu,\lambda,\nu)$ as the parameter vector that define location, scale and shape. The distribution is symmetric and unimodal and so the location parameter defines both the mode, median and mean of the distribution. The distribution generalizes the Normal Distribution, when $\nu=2$, but also allows for higher or lower kurtosis. For example, when ν decreases, the distribution flattens. When $\nu=1$, the distribution follows the Laplace distribution, while it tends to the Uniform distribution when $\nu\to\infty$.

The conditional volatility estimates can be used to calculated Expected Shortfall by integrating under the Value-at-Risk distribution

(B15)
$$ES_{\alpha}(X) = -\frac{1}{\alpha} \int_{0}^{\alpha} VaR_{\gamma}(X)d\gamma,$$

with VaR_a being the $(1-\alpha)$ quantile of the estimated returns distribution. Since the conditional return distribution is time-varying and fully specified by the model in B9, the time-varying Expected Shortfall can be estimated by calculating time-varying $VaR_{\alpha,t} = \mu_t + \hat{\sigma}_t | t - 1G^{-1}(a)$, where G^{-1} is the inverse PDF function of the Generalized Error Distribution. The quantity $\mathbb{E}\Delta_{\alpha}P_t$ is then estimated by the empirical equivalents of $ES_{\alpha,t}$. The algorithms are implemented following (Ghalanos, 2020). The autoregressive order of both the ARMA and GARCH processes are kept at 1, the moving average orders are selected using the AICc allowing up to three lags.