

Finding the Goldilocks data collection frequency for the Consumer Price Index

LUIGI PALUMBO¹ AND TIZIANA LAURETI²

¹ BANK OF ITALY

² UNIVERSITY OF TUSCIA

The views expressed in this presentation are those of the authors only and do not necessarily represent the views of the Bank of Italy or the Eurosystem.

How often is “just right”?

Assumptions:

- Timing of CPI price collection: prices often vary during the period
- Data collection, processing, and storage has a cost

Objectives:

Estimate variability of CPI m-o-m variations at different sampling-in-time frequencies

Propose and empirically validate a framework to determine the optimal data collection frequency



Electricity and gas prices in Italy

- Mandatory publication of unregulated market offers for electricity and gas to guarantee transparency
- Used by ISTAT for official CPI price data collection
- Simplified data collection:
 - 9 cities
 - National consumption profile
 - Electricity consumption: 2700kWh/year
 - Gas consumption: 1400 m3/year
 - Selection of rates for electricity
 - Providers weighted by national market shares
 - Type of contracts (fixed or variable prices) weighted according to official reports

The screenshot displays the 'portale offerte luce e gas' website interface. On the left, there are filters for 'Garanzie', 'Modalità di pagamento', and a checkbox for 'Nuovo contratto'. A 'Confronta' button is visible below these filters. The main content area shows two energy offers:

- ENERGIA POSITIVA SOCIETA COOPERATIVA:** 100% Energia Positiva PRODUTTORE Casa. Annual price: 904,87 € annui. Offered at -102,66 € rispetto a Maggior Tutela. Valid from 31/08/2023 to 31/12/2023.
- Iren Mercato SpA:** IREN STAY LUCE VERDE PREZZO FISSO E-BIKE. Annual price: 906,16 € annui. Offered at -101,37 € rispetto a Maggior Tutela. Valid from 11/12/2023 to 20/01/2024.

Each offer includes a 'VAI AL DETTAGLIO' button and icons representing environmental benefits.

Daily Time-Product-Dummy index

- Weighted Time-Product Dummy index:

$$\ln p_{it} = \sum_{t=1}^T \delta_t D_t + \sum_{i=1}^N \beta_i D_i + \varepsilon_{it}$$

p_{it} : average price of operator i in month t

D_t : dummy equal to 1 if month is equal to t and zero otherwise

D_i : dummy equal to 1 for prices of operator i and zero otherwise.

- Weighted Least Squares using operator i market share as weight for each observation.
- Aggregate daily price level:

$$P_t = e^{\delta_t}$$

Month-on-month CPI variations

- Month-on-month rate of change:

$$\pi_t = \frac{\sum_{j=1}^k C P I_j^t}{\sum_{l=1}^k C P I_l^{t-1}} - 1$$

- Number of potential changes at each sampling frequency:

$$C(15, k)^2 = \binom{15}{k}^2 = \left(\frac{15!}{k! (15 - k)!} \right)^2$$

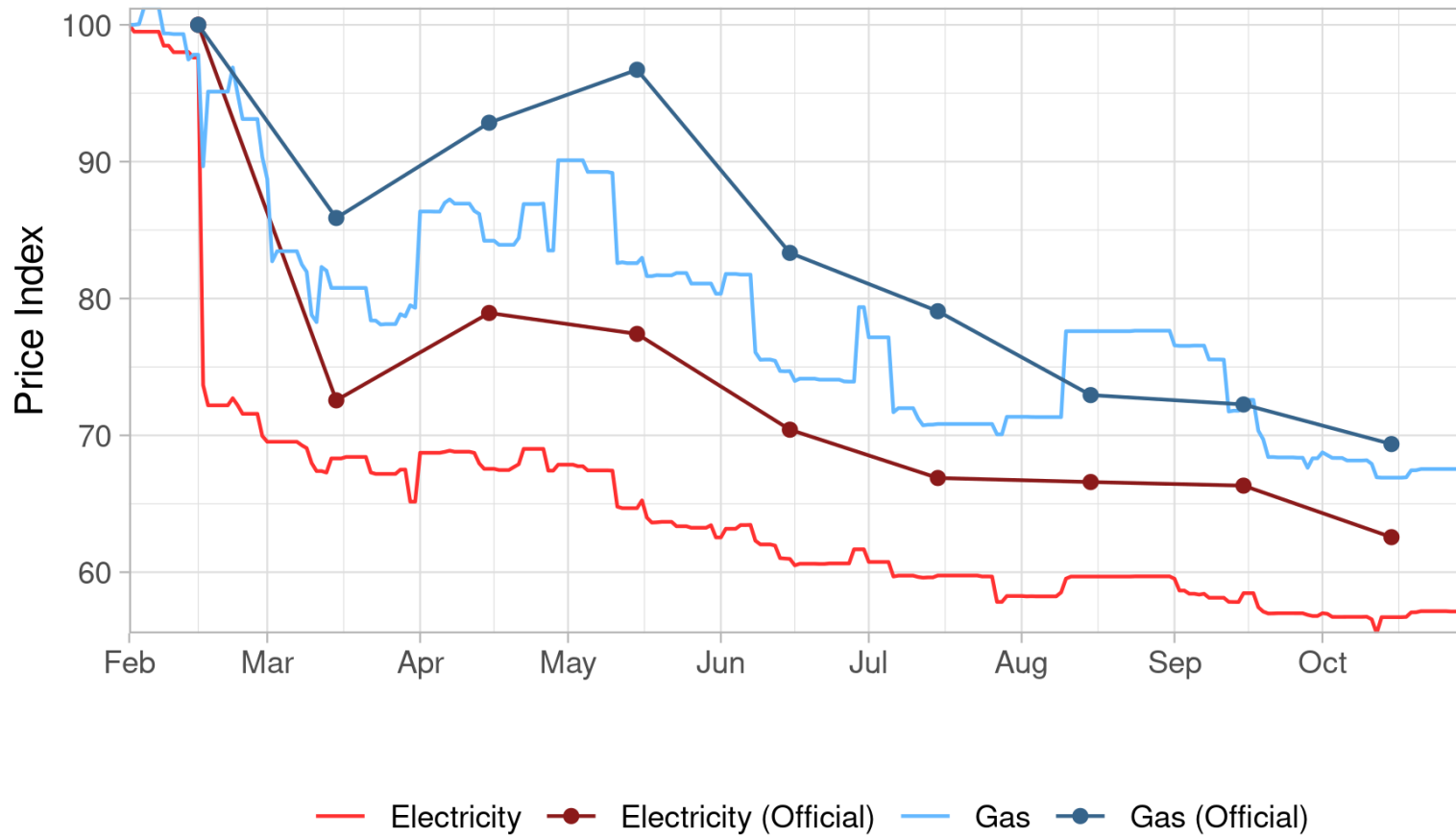
- Very large total number of combinations, over 155M each month for each utility.
Almost 2.5B combinations for this research

Optimization framework

- Uncertainty on CPI measurement can be represented by the StDev of potential m-o-m rates of change
- Costs for data collection, processing, and storage are usually expressed in monetary terms
- **Need to convert one of the two metrics (or both) to have a common unit of measure for optimization**
- Our approach: express the collection cost in minimum reduction of StDev to be obtained through additional data collection efforts

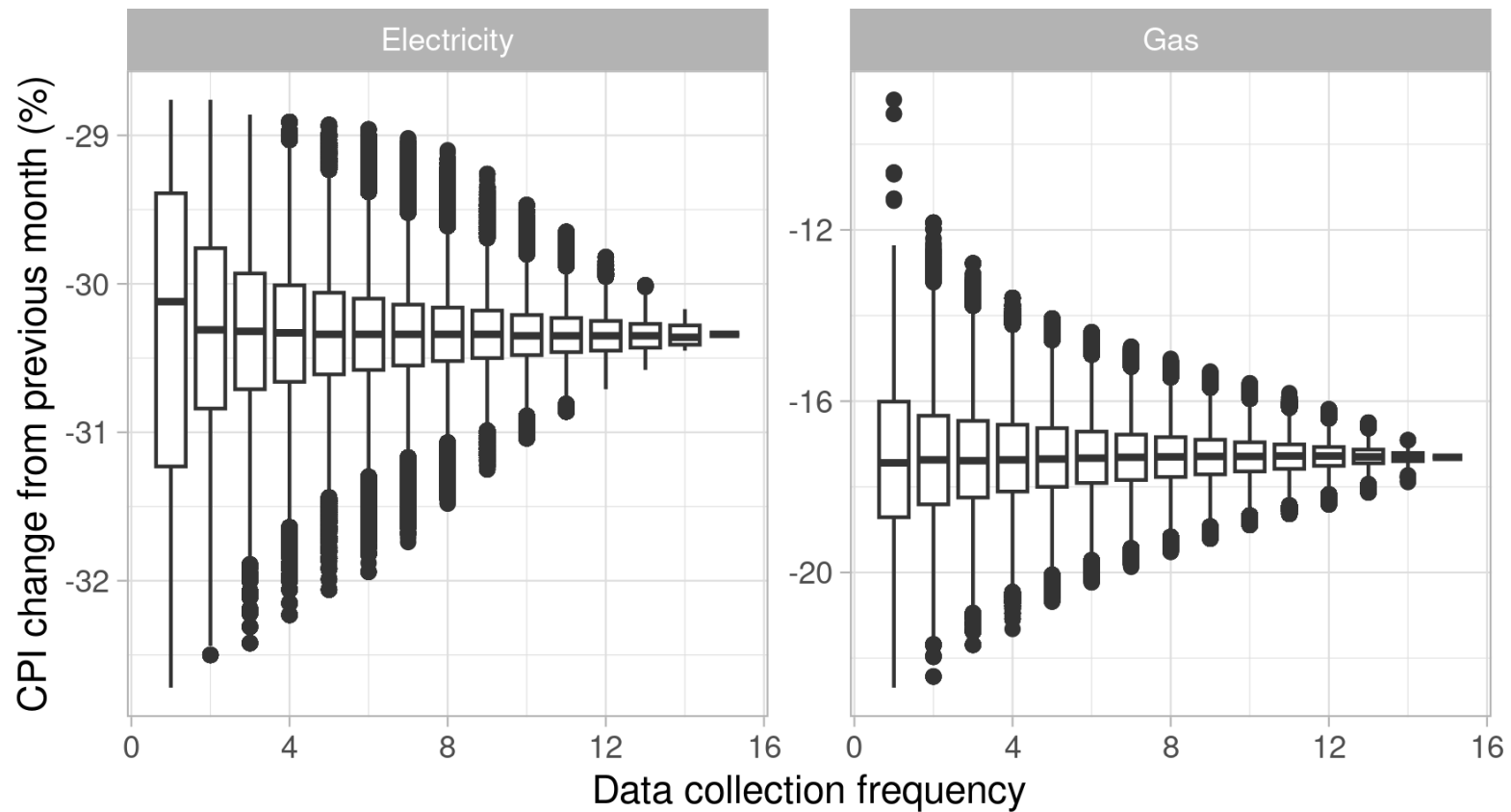
$$\min(UncertaintyCost(k) + CollectionCost(k)), k \in \mathbb{N}^+ : \{1 \leq k \leq 15\}$$

Daily vs. official CPI

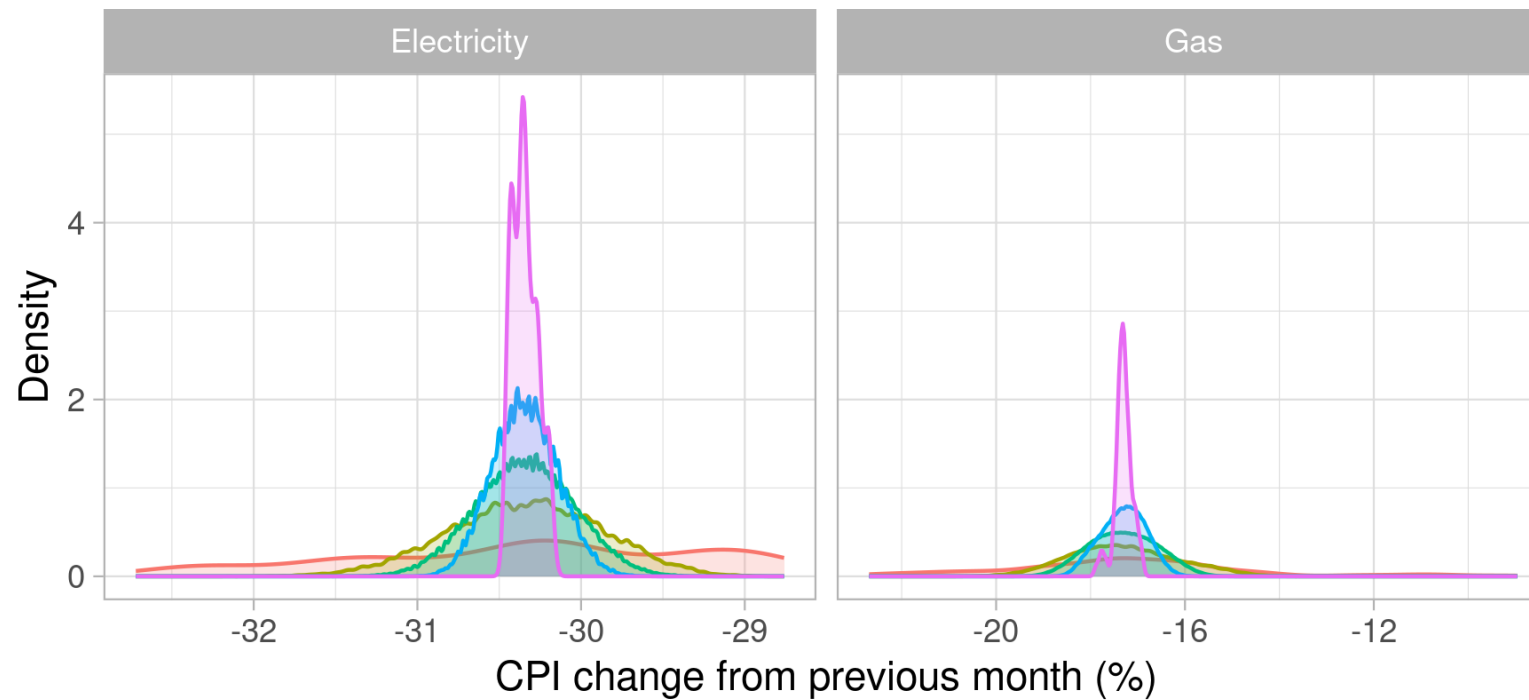


- Substantial intra-month price variation
- Daily CPI is not adjusted for changes in energy price subsidies

Boxplot of potential m-o-m CPI variations

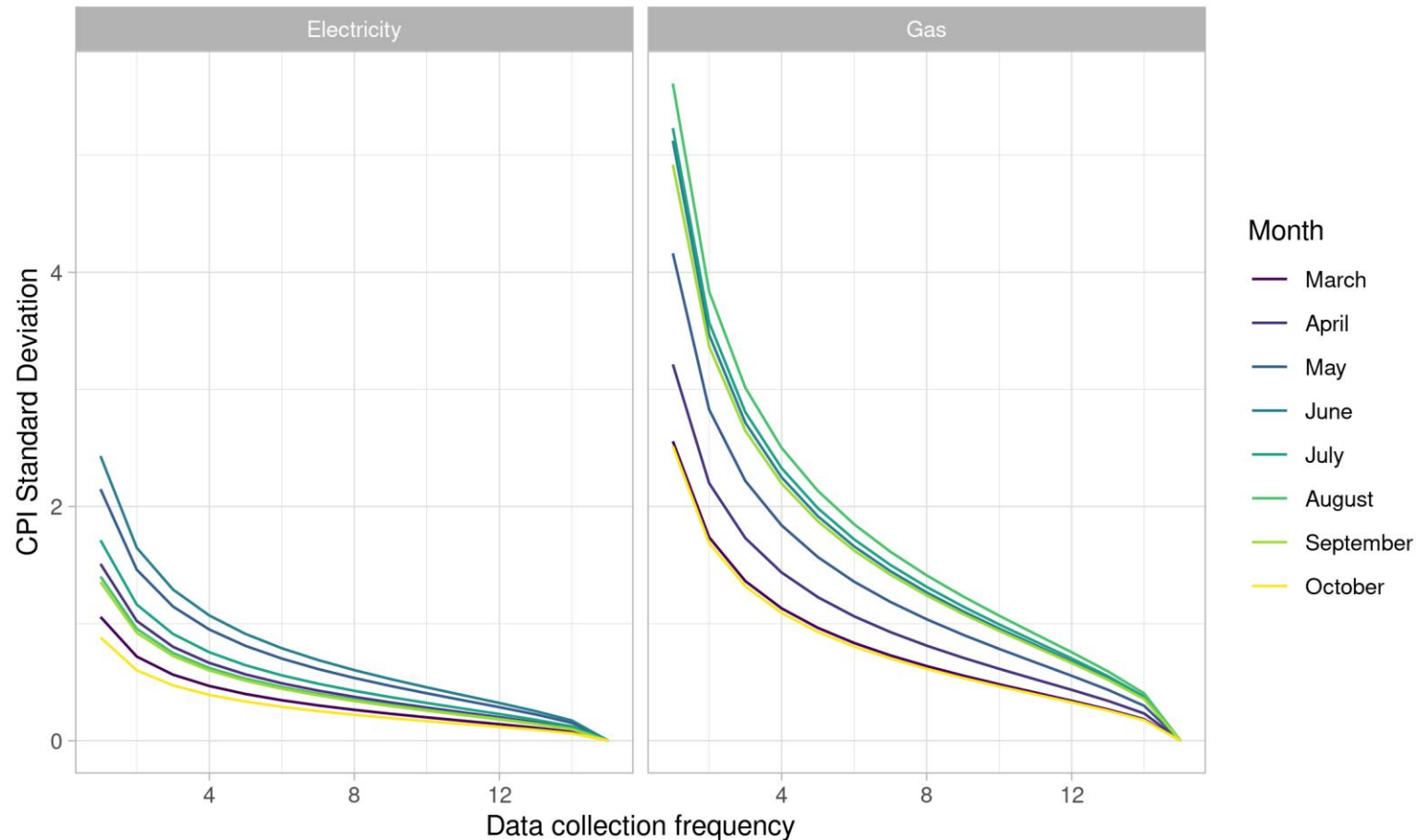


Density of potential m-o-m CPI variations



Data collection frequency ■ 1 ■ 4 ■ 7 ■ 10 ■ 14

StDev of potential m-o-m CPI variations



- Gas prices were much more volatile than electricity ones
- Levels vary across months and utilities, but paths are consistent

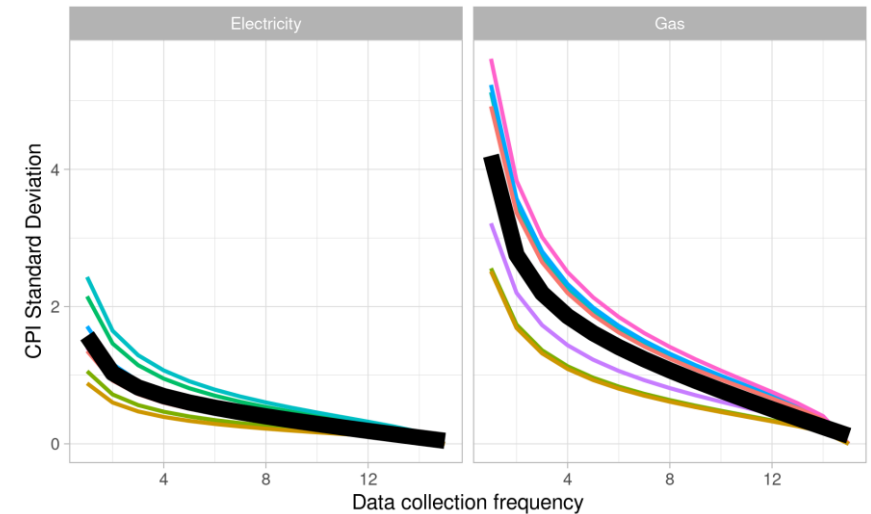
Modeling uncertainty

$$UncertaintyCost(k) = \beta_0 + \beta_1 k + \frac{\beta_2}{k} + \eta$$

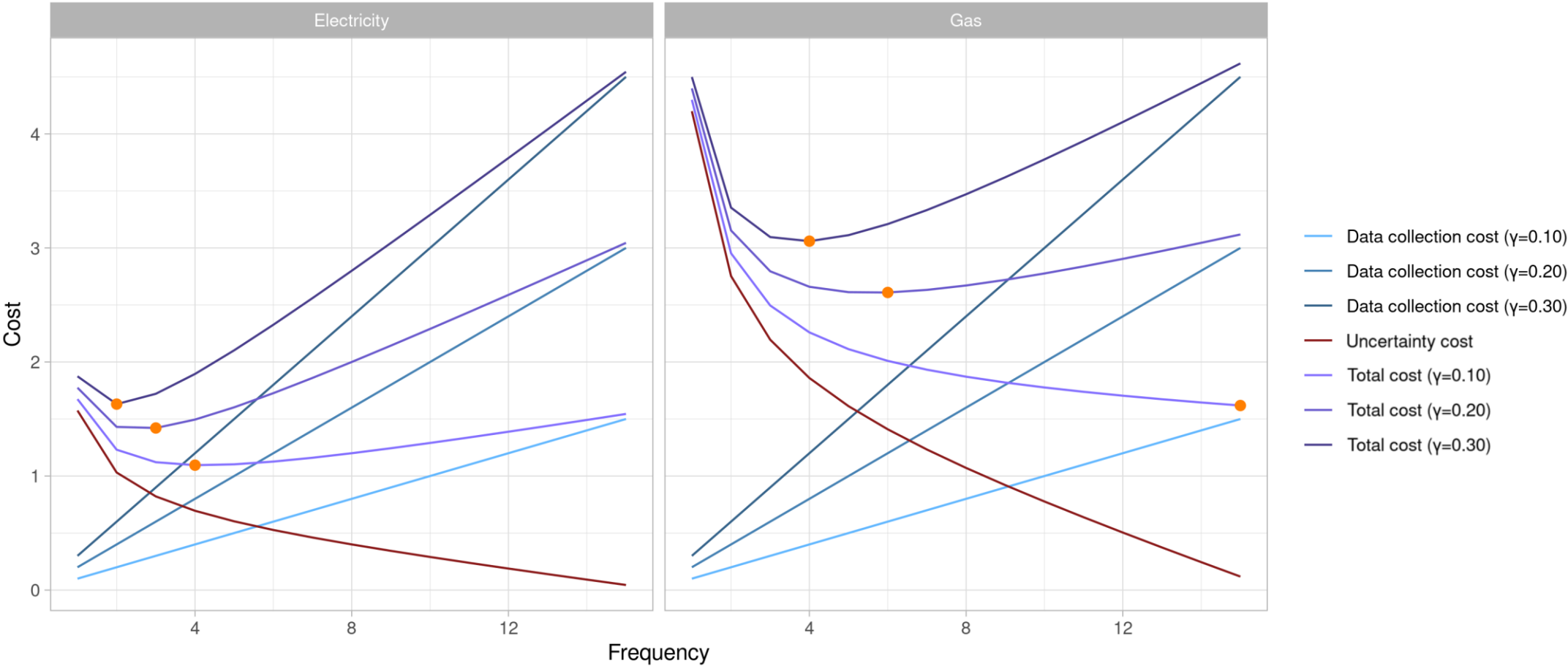
	Standard Deviation					
	Electricity			Gas		
	(1)	(2)	(3)	(4)	(5)	(6)
β_0	0.147*** (0.029)	1.167*** (0.049)	0.614*** (0.075)	0.394*** (0.073)	3.117*** (0.124)	1.649*** (0.182)
β_1		-0.084*** (0.005)	-0.042*** (0.006)		-0.224*** (0.014)	-0.114*** (0.016)
β_2	1.581*** (0.090)		1.002*** (0.117)	4.218*** (0.225)		2.664*** (0.283)
Adjusted-R2	0.720	0.668	0.794	0.746	0.694	0.824

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

- Approximately hyperbolic cost function
- Parameter estimates consistent with expectations and across utilities



Finding the Goldilocks frequency



Conclusions

- Variability of CPI m-o-m changes can be substantial at low data collection frequencies, affecting aggregated CPI figures
- Increasing data collection frequency yields diminishing returns in terms of CPI uncertainty reduction
- It is possible to determine an optimal data collection frequency leveraging historical price volatility and a cost function for data collection
- Optimality should be periodically reassessed, since both price variability and data collection costs may vary

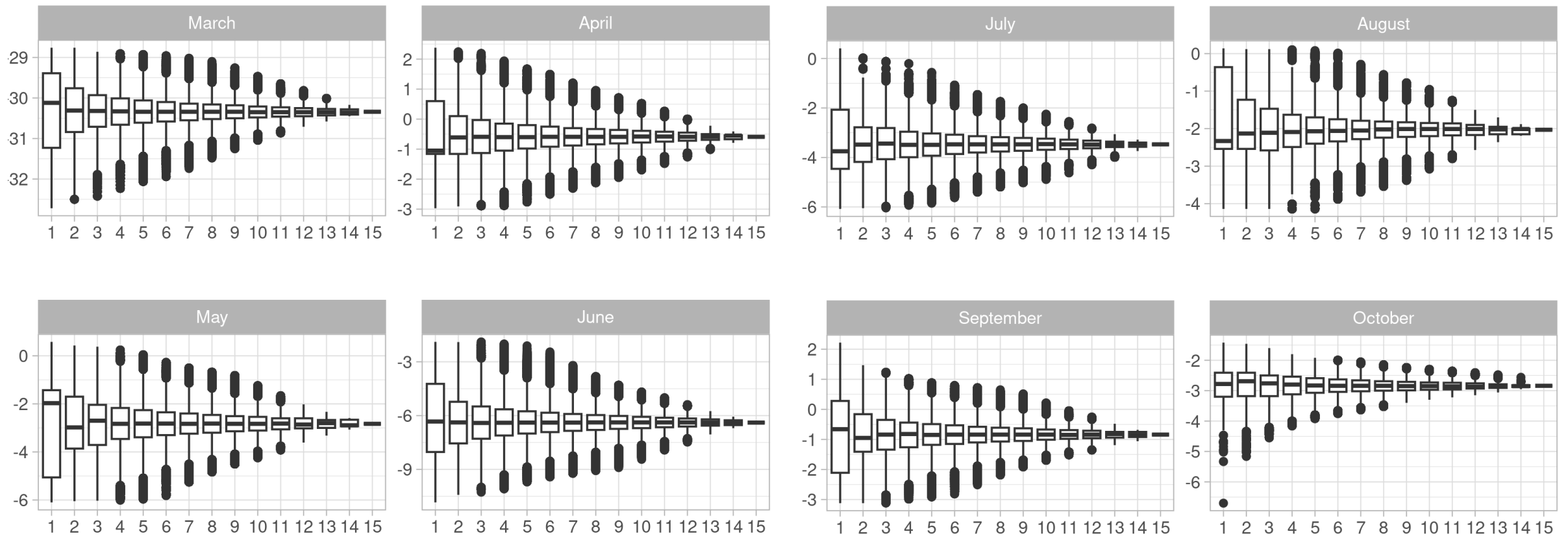
Thank you!

Q&A

LUIGI PALUMBO – LUIGI.PALUMBO@BANCADITALIA.IT

TIZIANA LAURETI – LAURETI@UNITUS.IT

Electricity: Boxplot of potential m-o-m CPI variations



Gas:

Boxplot of potential m-o-m CPI variations

