

Treatment for missing prices in the U.S. CPI due to the COVID-19 pandemic

Robert Cage

Assistant Commissioner
Consumer Price Index

Mark Bowman

Economist
Consumer Price Index

prepared for
UNECE Webinar:
Producing the CPI Under Lockdown
November 4, 2020



Imputation: A bird's eye view

■ Definition

1. Procedure for entering a value where response is *missing* or *unusable*
2. Process of replacing missing data with *substituted* values

■ What is the objective of price imputation?

- ▶ Select best *proxy* for the current month price of the originally selected, previously priced product
- ▶ Choice of 'best proxy' should be tethered to measurement objective of elementary index
- ▶ Choice of 'best proxy' may vary due to the *reason* for non response

■ Common techniques

- ▶ Listwise deletion
- ▶ Single imputation

Hot deck

Cold deck

Class Mean (cell-relative)

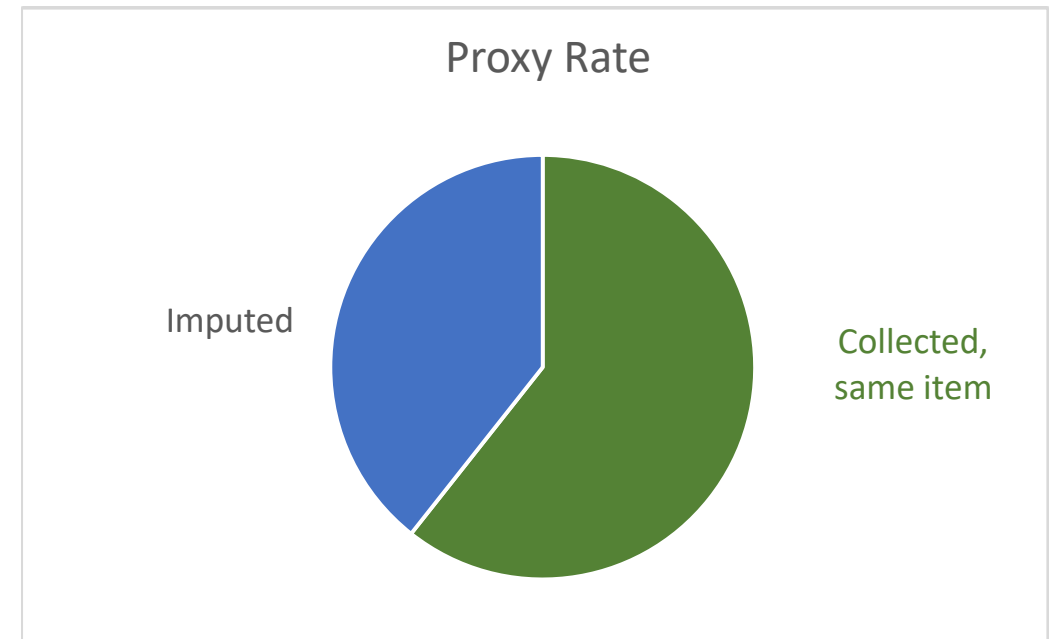
Regression

- ▶ Multiple imputation



Imputation: BLS outcomes pre-COVID

		April to September 2019	
DATA COLLECTION OUTCOME		Number of Quotes	Frequency
PROXY PRICE	Price collected	464,481	64.6%
	Same item as previous month	435,807	60.6%
	Substitute item	28,674	4.0%
	<i>Comparable</i>	16,030	55.9%
	<i>Quality change adjusted</i>	4,404	15.4%
	<i>Non-comparable, cell mean imputed</i>	3,281	11.4%
	<i>Non-comparable, sub class imputed</i>	4,959	17.3%
	Price not collected	254,184	35.4%
	Cell-relative imputed	87,443	12.2%
	Off-cycle, carry forward	166,741	23.2%

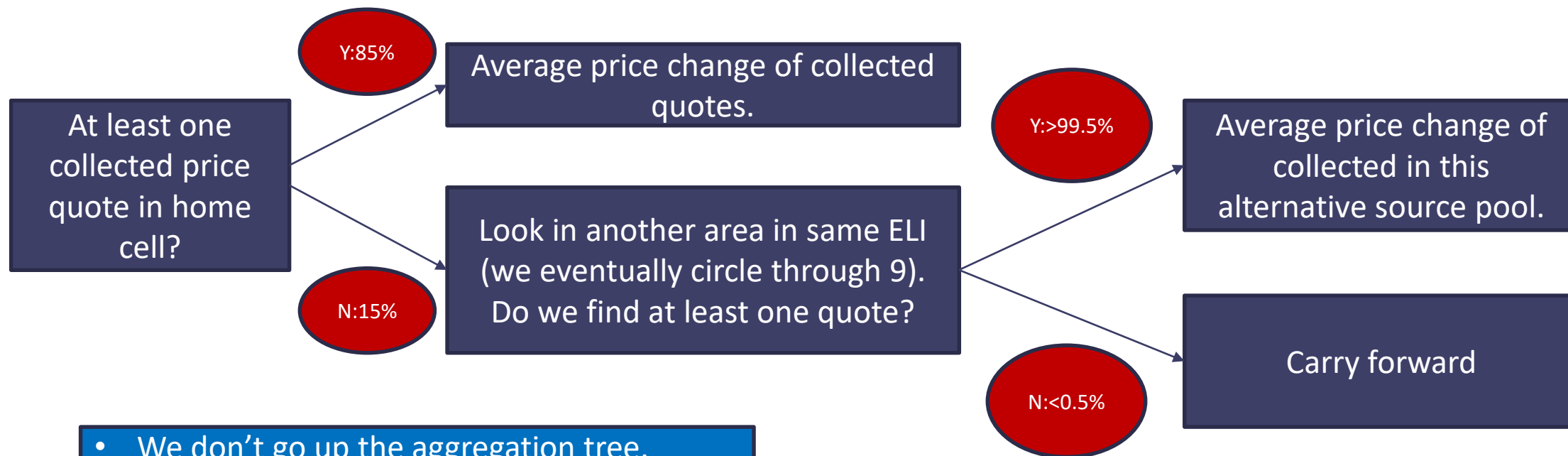


Structure: Methods

- 243 ITEMS by 32 AREAs \approx 7700+ elementary / basic indexes
- Each ITEM has at least one Entry Level Item (ELI)
- Each AREA has at least one Primary Sampling Unit (PSU)
- All ELIs not necessarily sampled in each ITEM
- Imputation of missing prices occurs at the quote level using ELI-PSU



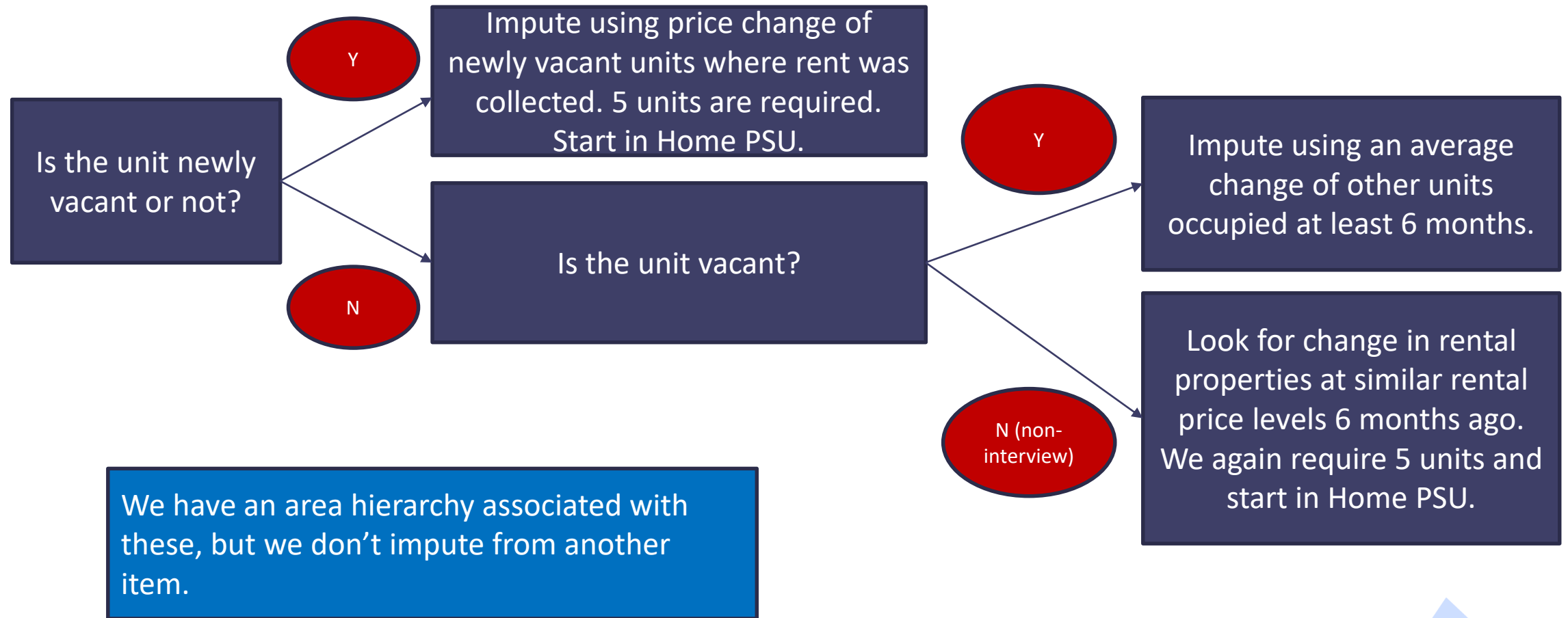
Commodities & Services (C&S) imputation



- We don't go up the aggregation tree.
- Recent internal studies have suggested a stronger correlation within item/across area than across area/within item.
- National pricing

The percentages are approximate for post-pandemic period.

Housing imputation



Other imputations

- Upfront subsampling
- Pricing frequency
- Imputation when an item is replaced especially those with a product life cycle
- Pandemic related:
 - ▶ What if the pandemic didn't happen, what price would we collect/use?
 - ▶ What about items that are available for sale, but not on the website?
 - ▶ Sit-down restaurant “impute” price change using price of the carry out meal



Performance

- We have done a variety of metrics that measure imputation and collection rates before and after data collection restrictions.
- <https://www.bls.gov/covid19/effects-of-covid-19-pandemic-on-consumer-price-index.htm>



Collection modes

Table 1. Collection modes for C&S price survey, percent

Month	2019			2020		
	Personal visit	Telephone	Online	Personal visit	Telephone	Online
March	74	10	16	44	14	42
April	74	10	16	0	18	82
May	74	10	16	0	17	83
June	73	10	17	0	17	83
July	74	10	16	0	17	83
August	73	10	17	0	17	83
September	73	10	17	0	17	83
Avg. Feb 2019-Feb 2020	73	10	17			

Table 2. Collection modes for the housing survey, percent

Month	2019		2020	
	Personal visit	Telephone	Personal visit	Telephone
March	36	64	30	70
April	38	62	0	100
May	39	61	0	100
June	38	62	0	100
July	37	63	0	100
August	37	63	0	100
September	38	62	0	100
Avg. Feb 2019-Feb 2020	38	62		



Response rates

Table 3. Response rates for C&S price survey, percent

Month	2019		2020	
	Collected prices	Uncollected prices	Collected prices	Uncollected prices
March	87	13	76	25
April	86	14	65	35
May	87	13	68	32
June	86	14	70	30
July	86	14	70	30
August	86	14	72	28
September	86	14	73	27
Avg. Feb 2019-Feb 2020	86	14		

Table 4. Response rates for the housing survey, percent

Month	2019			2020		
	Collected rents	Vacant rents	Uncollected rents	Collected rents	Vacant rents	Uncollected rents
March	67	6	27	63	5	32
April	68	6	26	61	4	35
May	69	6	25	61	4	35
June	68	6	26	62	4	34
July	69	6	25	60	4	36
August	67	6	27	62	4	34
September	65	6	29	60	4	36
Avg. Feb 2019-Feb 2020	68	6	26			



Imputation source

Table 5. Imputation source for C&S price survey, percent

Month	2019			2020		
	Home cell	Different cell	Carry forward	Home cell	Different cell	Carry forward
March	94	6	0	90	10	0
April	92	8	0	84	16	0
May	92	8	0	85	15	0
June	92	8	0	86	14	0
July	92	8	0	86	14	0
August	91	9	0	86	14	0
September	93	7	0	87	13	0

Table 6. Imputation source for the housing survey, percent

Month	2019		2020	
	Noninterview	Vacancy	Noninterview	Vacancy
March	82	18	87	13
April	82	18	90	10
May	82	18	91	9
June	81	19	89	11
July	81	19	89	11
August	83	17	91	9
September	83	17	90	10



Problematic items

- Historical collection issues:

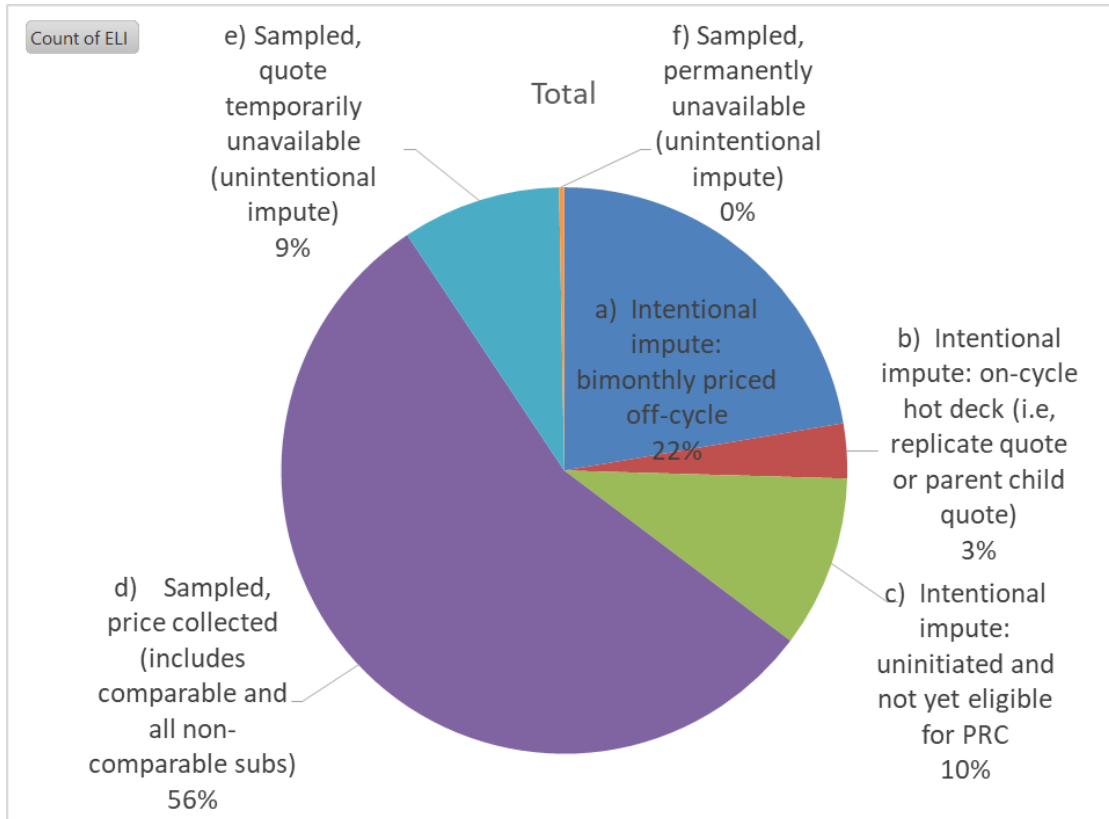
- ▶ lamb; clothing alteration; appliance repair; alternative fuels; motor vehicle maintenance; state vehicle registration; intercity train fare

- Issues related to the pandemic:

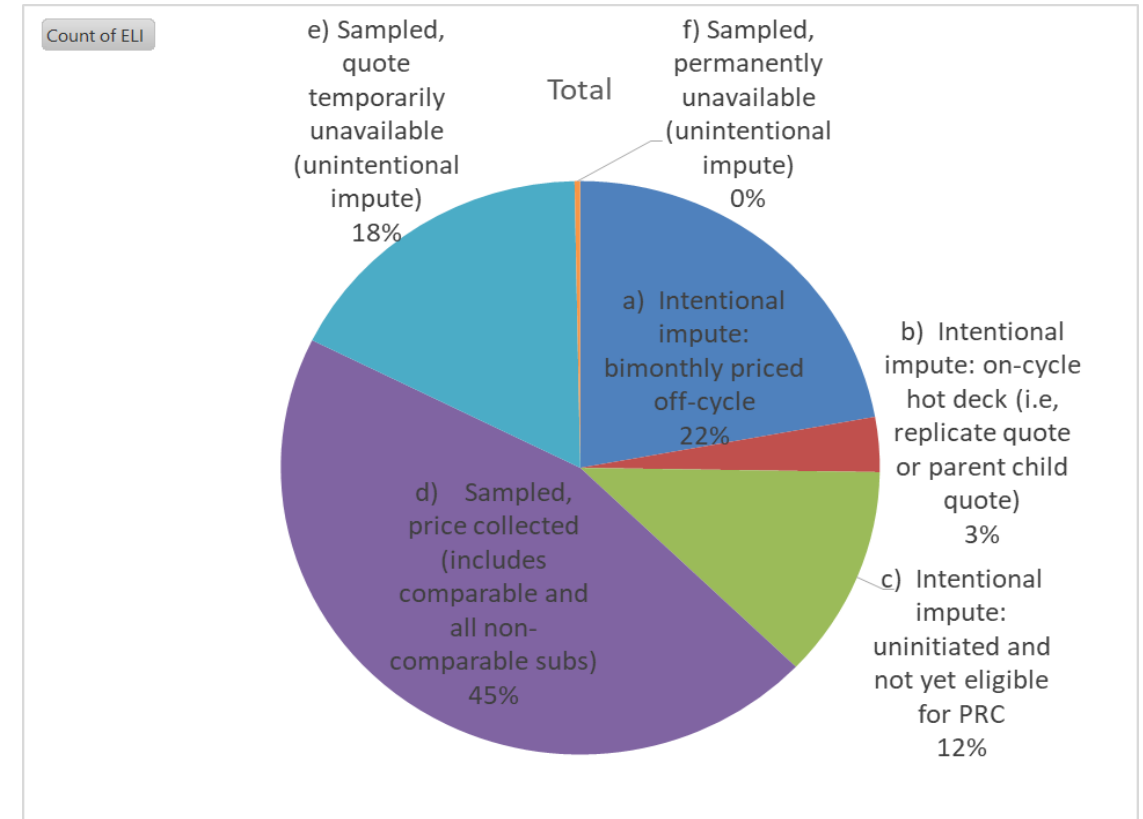
- ▶ school books; food from vending machines & mobile vendors; admission to sporting events

Growth of temporarily missing (C&S)

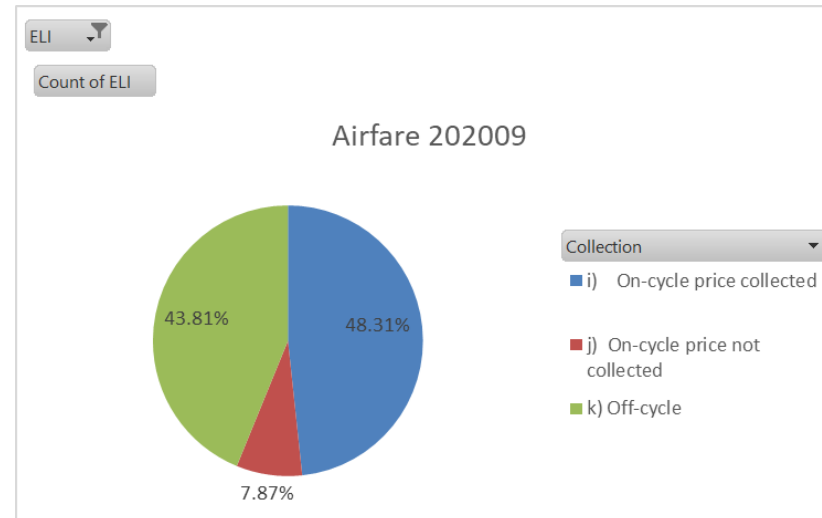
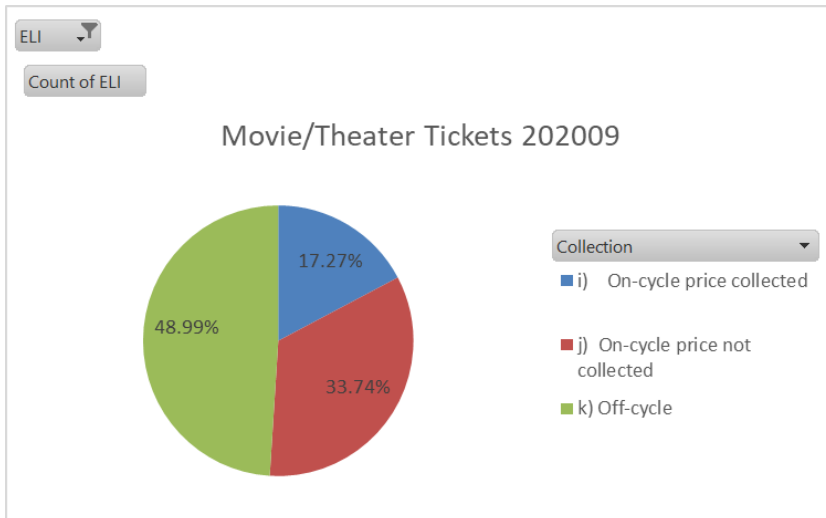
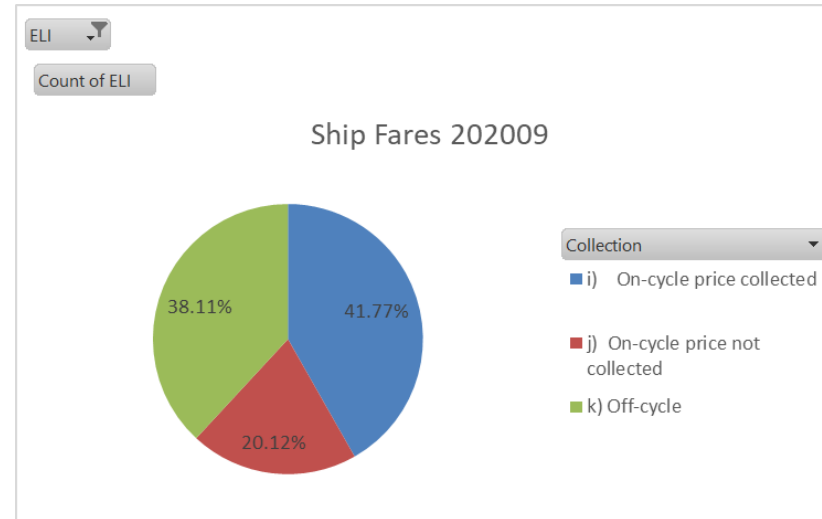
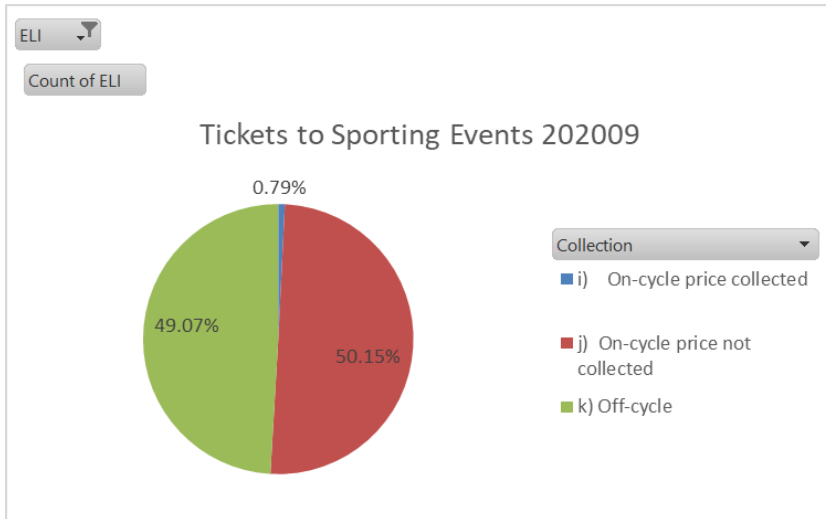
September 2019



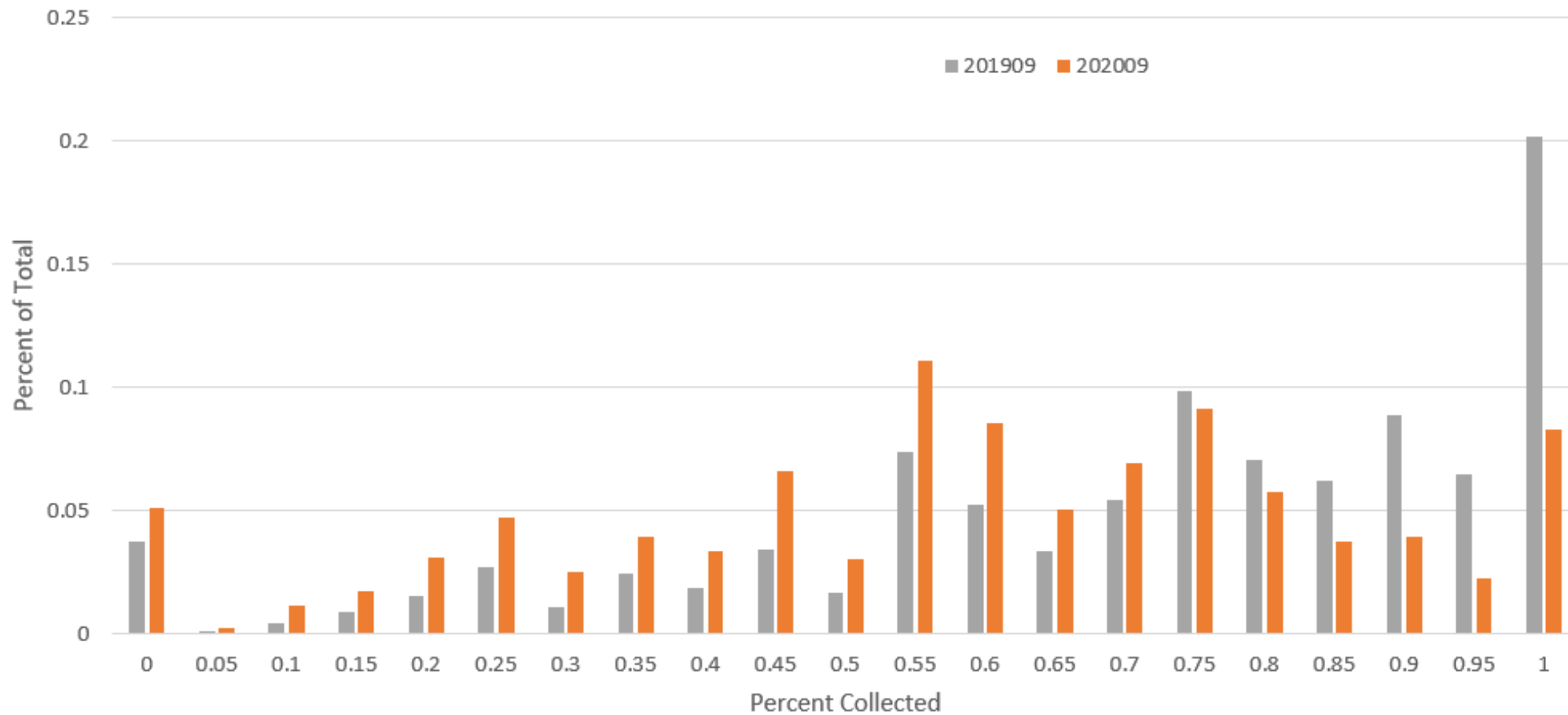
September 2020



Entry level items impacted by the pandemic



Weighted average of quotes



Imputation: BLS outcomes during COVID

DATA COLLECTION OUTCOME		April to September 2019		April to September 2020	
		Number of Quotes	Frequency	Number of Quotes	Frequency
PROXY PRICE	Price collected	464,481	64.6%	383,238	51.8%
	Same item as previous month	435,807	60.6%	358,242	48.4%
	Substitute item	28,674	4.0%	24,996	3.4%
	<i>Comparable</i>	16,030	55.9%	14,019	56.1%
	<i>Quality change adjusted</i>	4,404	15.4%	5,121	20.5%
	<i>Non-comparable, cell mean imputed</i>	3,281	11.4%	2,030	8.1%
	<i>Non-comparable, sub class imputed</i>	4,959	17.3%	3,826	15.3%
	Price not collected	254,184	35.4%	356,276	48.2%
	Cell-relative imputed	87,443	12.2%	177,317	24.0%
	Off-cycle, carry forward	166,741	23.2%	178,959	24.2%



Responses to questions / Future plans

- Imputation methods for products that are still available in the market, but where it is not possible to collect prices for the full set of sampled products
 - ▶ We impute by the average price movements of available products following current procedures.
- Or products where there are no transactions because of the lockdown
 - ▶ If we can find an offer price online we will use this, but this is already an imputation. Consider tickets to sporting events.

Responses to questions (cont.)

- Products with strong seasonal price variation
 - ▶ We have a “class-mean” imputation that estimates the price of non-comparable replacements using the price change of comparable replacements, e.g. electronics, new vehicles, apparel.
 - ▶ For items that are only available part of the year we select two samples corresponding to the items that are expected to be available in each part of the year.

Responses to questions (cont.)

- Updates of the sample basket/expenditure weights
 - ▶ The rotation of the sample of stores is typically for 1/8 of the sample every six months resulting in a full refresh every 4 years for most items.
 - Weights are fixed when the outlet rotates in.
 - ▶ We have modified data collection procedures to allow for replacement of some unavailable items with collection via phone or web.
 - ▶ Upper level weights every 2 years for CPI-U and monthly with a lag for C-CPI-U.

Responses to questions (cont.)

- A decision tree that can be adapted and used by statistical offices as part of contingency planning in cases of lockdown and calculation methods when the lockdown is lifted.
 - ▶ When in-person collection resumes and the sampled unique item is available again we will reflect the price change from the string imputed prices to the then current price.

Imputation in the expenditure weights

- Two techniques used to minimize the variance associated with each elementary item–area base-period expenditure estimate.
 - ▶ Data are pooled over an extended period in order to build the expenditure estimates on an adequate sample size.
 - ▶ Second, elementary item–area expenditures are averaged, or composite-estimated, with item-regional expenditures. This has the effect of lowering the variance of each elementary item–area expenditure at the cost of biasing it toward the expenditure patterns observed in the larger geographical area.

Contact Information

Robert Cage

Assistant Commissioner
Consumer Price Index

www.bls.gov/cpi

202-691-6959

cage.rob@bls.gov

Mark Bowman

Economist
Consumer Price Index Program

www.bls.gov/cpi

202-691-7000

bowman.mark@bls.gov

