C14/23/026 The impact of temporary VAT reductions on grocery retail

A Data Management Plan created using DMPonline.be

Creators: Kathleen Cleeren, n.n. n.n.

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Principal Investigator: n.n. n.n., Kathleen Cleeren

Data Manager: n.n. n.n., Kathleen Cleeren

Project Administrator: n.n. n.n., Kathleen Cleeren

Grant number / URL: C14/23/026

ID: 205960

Start date: 30-09-2023

End date: 30-09-2027

Project abstract:

We investigate the pass-through of the German temporary VAT reduction policy in 2020 to grocery prices at the introduction and potential asymmetries in price increases after the lift of the temporary VAT cut. We investigate heterogeneity across retailers, product categories and brands, and also study the impact of different pricing strategies. As such, we study the difference in consumer reactions to retailers that include the temporary VAT cut in prices on the shelves versus retailers that deduct the VAT reduction at the counter. In addition, we investigate how retailers' response to the temporary VAT cut interacts with the omnipresent left-digit pricing bias strategy (i.e., the preference for .99 prices over .00 prices).

Last modified: 26-03-2024

C14/23/026 The impact of temporary VAT reductions on grocery retail

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		*Indicate: **N**(ew data) or **E**(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
IRI retail scanner data	Dataset with sales information on all grocery products sold per week per store for Germany and the Netherlands	E	D	N	.csv	587 GB	
Anonymous household scanner data	Dataset for a panel of (anonymous) households which contains information on all grocery purchases made during several years for the Netherlands and Germany.	E	D	N	.csv	45 GB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

The IRI dataset is delivered by IRI a market research company (currently named Circana https://www.circana.com) through the agreement with AiMark (a institute that bridges academics and practitioners Aimark (goaimark.com)).
The household scanner dataset is delivered by Aimark.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

The researchers signed an NDA of 3rd party (AiMark) with regard to the anonymous household scanner data & the retail scanner data. This prevents us from disseminating these data.

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

- Codebook with description of variables is provided with dataset
- Explanatory comments in code script to describe variables that are newly created

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Other (specify below)

Stored on enterprise server storage network at the Faculty of Economics and Business, KU Leuven. This server storage is in-house identified and inventoried, and encrypted at rest.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

Incremental and timely multiple backups of the data are taken by ICT support and stored cross data center in the other of the two data centers

concerned. At any point in time, different time snapshots and full backups of data exist that allow for recovery up to the latest seven days, the latest 13 weeks and the latest three months. Storage servers are virtualized servers, that can be restored on multiple virtual server hosts inside the university data centers involved.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Access to data, as well as reading and editing rights, are restricted to a limited number of authenticated persons (i.e. the PhD student and promotors), based on security groups managed and audited through a web interface by the research lead itself. Access to the web interface itself is shielded through a Shibboleth authentication process. Access from other devices or access without proper access authorization is prevented by a combination of system and storage access control, encryption and firewalling of the systems involved. Access within the KU Leuven firewalls is through machine certificates and encryption restricted to a subset of university managed devices.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

The cost of data storage has been determined to be 9,303 euros and has been paid using internal funds (IMP/20/003 – project “The Impact of COVID-19 on the supermarket grocery sector”). However, we will prolong this storage using the funds of this C1 project.

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- Other (specify below)

Network drive at Faculty of Economics and Business, KU Leuven

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The data storage cost per TB has been determined to be 443 euros per year. These costs have been paid until the end of 2028. We will prolong this using means from the C1 funds.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- No (closed access)

If access is restricted, please specify who will be able to access the data and under what conditions.

Only the supervisors (Kathleen Cleeren and Marco Kotschedoff) and the PhD student will be able to access the data. The latter will only be able to access the data as long as he/she is under contract at KU Leuven.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- Yes, privacy aspects
- Yes, intellectual property rights

The researchers signed an NDA of 3rd party (AiMark) with regard to the anonymous household scanner data & retail scanner data.

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other (specify below)

The data cannot be made publicly available.
The data are already available on the protected network drives of the Faculty but only accessible to the researchers.

When will the data be made available?

- Other (specify below)

The data cannot be made available given the NDAs.

Which data usage licenses are you going to provide?

If none, please explain why.

- Other (specify below)

The data cannot be reused given the signed NDAs.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- No

What are the expected costs for data sharing? How will these costs be covered?

The data cannot be shared given the signed NDAs

Responsibilities

Who will manage data documentation and metadata during the research project?

The promotors (Kathleen Cleeren and Marco Kotschedoff)

Who will manage data storage and backup during the research project?

Data storage: Server and storage equipment is managed through operations and system management software and monitored 24/7 by ICT support staff with notification and escalation procedures in place.
Data backup: Incremental and timely multiple backups of the data are taken by ICT support

Who will manage data preservation and sharing?

The promotors will preserve the data. Both of them have a permanent contract at KU Leuven and guarantee the commitment to the project.
The data cannot be shared given the signed NDAs.

Who will update and implement this DMP?

The promotors (Kathleen Cleeren and Marco Kotschedoff)