University of Basel

Faculty of
Business and Economics

December 2022

Gathering of data relevant for PV investment decisions

WWZ Working Paper 2022/10

Fintan Oeri

A publication of the Center of Business and Economics (WWZ), University of Basel.

Universität Basel
Peter Merian-Weg 6
4052 Basel, Switzerland
wwz.unibas.ch

Corresponding Author:
Fintan Oeri
fintan.oeri@unibas.ch

Gathering of data relevant for PV investment decisions (WWZ Forum projects FV-71 & FV-79)

Fintan Oeri
fintan.oeri@gmail.com

November 1, 2022

Contents

1 Introduction

This document describes the data that was collected in the course of the two WWZ Forum research projects, FV-71 and FV-79, led by Prof. Pascal Gantenbein and Prof. Aya Kachi.[1] Among other outputs, the projects also resulted in an R shiny app[2], which makes some of the collected data on PV installations accessible and illustrates potential applications.[3] The app presents two aggregated datasets. However, both build on a set of individual data sources. A major difficulty any analysis on the PV distribution in Switzerland faces is that

[1]We gratefully acknowledge the generous financial support of the WWZ Forum Förderverein for both research projects.

[2]`https://founibas.shinyapps.io/02_shinyapp/`

[3]For access to the underlying data, please contact Fintan Oeri.

relevant explanatory variables build on distinct observation units (political municipalities, addresses, post codes, operators with unclear areas of activity). Not only does the linking between these change over time (e.g., because of municipalities merging), in a substantial number of cases it is not possible to identify a unique link (e.g., same postcode and town names can refer to multiple political municipalities). In addition, the authors of the different datasets sometimes refer to the same entity using different names and abbreviations or do not provide official codes (e.g., grid operators changing their names over time). The R script "20220822_datapreparation.R" (available on request) outlines a pragmatic approach to handle some of these challenges. The approach taken uses a long data structure (i.e., separate rows for each year, category etc.) with 30+ entries per year and municipality. Postcode/town is not an official spatial unit but governed by the post and most of the data does not feaure that level of granularity. Hence, differences in values between separate postcodes from the same municipality are not reflected. However, to facilitate searching the data, the aggregated entries for each political municipality are duplicated for all corresponding postcode/town names. The following briefly describes all data sources used, i.e.:

- cleaned and processed data on PV installations in Switzerland
- large dataset combining all data together
- data on population size
- data linking observation units
- data on electricity prices
- data on feed-in-tariffs
- data on subsidies
- data on PV potential

2 Combined, processed datasets for shiny app

2.1 All data combined: *df_fullmerge*

This dataset combines all available data assumed to affect the PV investment decision as well as a range of outcome variables, i.e., variables on PV installations. It serves as the foundation for the second tab of the app.

- *Observation unit*: Political municipality - year - operator - product type - electricity consumption profile
- *Time range*: 2015-2023 (2021-2023 incomplete)
- *File*: ~/01_data/08_preppedata/20220822_fullmerge.RData

- *Raw data storage*: ∼/01_data
- *Raw data source*: multiple, see below

2.2 Installed PV: *df_installed*

This dataset builds on the data on PV installations and describes the installed PV in various ways. It provides the foundation for the first tab of the app.

- *Observation unit*: Political municipality - year
- *Time range*: 2015-2021 (2022 incomplete)
- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/01_installed
- *Raw data source*: `https://opendata.swiss/de/dataset/elektrizitatsproduktionsanlagen`

3 Individual, processed datasets to be combined

3.1 Population size: *pop*

This data features the population size of municipalities over time. It serves to calculate per capita equivalents. Since I was not able to obtain data newer than 2020, the missing values for the years 2021-2023 were filled in using the value from 2020.

- *Observation unit*: Political municipality - year
- *Time range*: 2015-2020 (entries for 2021-2023 use data from 2020)
- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/02_population
- *Raw data source*: `https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/stand-entwicklung/bevoelkerung.assetdetail.18344310.html`

3.2 Linking dataset: *link_all*

This is the empty equivalent to df_fullmerge and it includes all handles by which to merge the different datasets.

- *Observation unit*: Political municipality - year - operator - product type - electricity consumption profile
- *Time range*: 2015-2023

- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/03_spatialunits
- *Raw data source A*: `https://www.bfs.admin.ch/bfs/de/home/grundlagen/agvch/gwr-korrespondenztabelle.html`
- *Raw data source B*: `https://www.elcom.admin.ch/elcom/de/home/themen/strompreise/tarif-rohdaten-verteilnetzbetreiber.html`

3.3 Electricity prices: *electricityprice_allyears*

This data features the electricity prices. Prices for electricity consumption may affect investment decisions in that they have an impact on the amount of money that can be saved through self-consumption of self-produced PV. Following the conventional presentation format, electricity prices are reported for a given set of consumption profiles. The data also reflects the different prices available for different product types (standard vs. cheapest). It does not, however, feature all available tariffs (day/night, summer/winter, green electricity). Information on whether there are additional tariffs available is included, though.

- Observation unit: operator - year - product type - electricity consumption profile
- *Time range*: 2015-2022 (data technically available as of 2009)
- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/04_electricityprices
- *Raw data source*: `https://www.strompreis.elcom.admin.ch/`

3.4 Feed-in-tariffs: *fit_allyears*

This data contains information on feed-in-tariffs, which the owners of PV installations receive in exchange for the electricity they produce. In contrast to the other data sources, this data is not freely available. To access it through its API one needs to apply for a key (free of charge; see link on data source for more info).

- Observation unit: operator - year
- *Time range*: 2015-2023 (data from 2015 & 2016 was provided in an different format)
- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/05_feedintariffs
- *Raw data source*: `https://www.vese.ch/wp-content/uploads/pvtarif/pvtarif2/appPvMapExpert/pvtarif-map-expert-data-de.html`

3.5 Subsidies: *subsidies_subset*

This data contains information on available subsidies and stems from a semi-automated data gathering effort from 2019 and 2020. While more detailed information is available, the broad range of different types of subsidy schemes required to standardise the information to a binary indicator. Moreover, since the available data sources allow to confirm but not rule out potential subsidies for a given municipalitiy, all non "1" are labelled as "NA".

- *Observation unit*: political municipality - year
- *Time range*: 2015-2020
- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/06_subsidies
- *Raw data source*: `https://energiefranken.ch/` & subsidy specific links

3.6 PV potential: *potential_subset*

For each municipality, this dataset lists the potential for PV electricity on roofs and facades. Despite new buildings being developed, the potential is assumed to be constant in all years because I was not able to obtain data on its development over time. The potential is calculated in terms of energy (gigawatt per hour), while the installations are described in terms of their power (flow rate of the energy, kilowatt). Determining to what extent a municipality's PV potential is fulfilled requires the estimation of the aggregated energy its installations produce (conversion of peak kilowatt peak (kwp) to energy). However, an installation's kwp is determined under standardised condition while the same installation produces very different amounts of energy depending on its location (e.g., because of sun exposure and altitude). This makes it difficult to estimate the energy produced from an installation. For the conversion, I assumed that, for every kwp, an installation produces 1000 kwh in one year (this is slightly above the Swiss average of 950 kwh, see `https://www.uvek-gis.admin.ch/BFE/storymaps/EE_Elektrizitaetsproduktionsanlagen/`).

- *Observation unit*: political municipality
- *Time range*: one observation only (2022)
- *File*: ∼/01_data/08_preppedata/20220822_alldata_prepped.RData
- *Raw data storage*: ∼/01_data/07_potential
- *Raw data source*: `https://opendata.swiss/de/dataset/solarenergiepotenziale-der-schweizer-gemeinden`