



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Département fédéral de l'intérieur DFI

**Office fédéral de la statistique OFS**  
Economie

---

## **Using electronic data for the rental price index: child's play?**

---

## Table of contents

1	Terms and abbreviations .....	3
2	Rental price index compilation in brief .....	3
2.1	Rental collection .....	4
3	Introducing automatic data transmission for the rental price index .....	5
3.1	Automatic data transmission in a nutshell.....	5
3.2	Project plan .....	6
3.2.1	Initialisation phase 2018-2022 .....	6
3.2.2	Transition and introduction phase 2023 .....	7
3.2.3	Next steps 2024-2029 .....	8
4	Practical aspects of automatic data transmission.....	8
4.1	Request and response mechanism .....	10
4.1.1	The EGID-EWID request.....	10
4.1.2	The EGID-tenant request .....	10
4.2	Return codes .....	12
4.3	Why some dwellings do not match.....	13
4.4	Why some dwellings mismatch .....	14
4.5	Data interpretation.....	15
4.5.1	Interpretation rule for ADDRESS_KIND.....	15
4.5.2	Interpretation rule for BUILD_YEAR_NB .....	16
4.5.3	Interpretation rule for OWNER_TYPE_NB .....	17
5	Results .....	18
5.1	Matching rates.....	18
5.1.1	Procedure for dwellings that do not match.....	19
5.2	Increase of efficiency in data management .....	19
5.2.1	Time saved by real estate management companies .....	20
5.2.2	Time saved by private households responding to the screening phase .....	20
5.2.3	Time saved by the FSO .....	20
5.3	Improvement of data quality for selected variables.....	21
6	Key learnings .....	21
6.1	Automatic data transmission is not child's play but worth the effort .....	21
6.2	Discovering potential errors and improving overall quality control.....	22
6.3	Insights into the real estate management software .....	22
6.4	Different data collection techniques may lead to different results.....	23
7	Appendices .....	24
7.1	Appendix 1 mpeRequest_FindUnit_EgidEwid .....	24
7.2	Appendix 2 mpeRequest_FindUnit_EgidTenant.....	25
7.3	Appendix 3 mpeResponse .....	26
7.4	Appendix 4 rental index compilation in more detail.....	29

## 1 Terms and abbreviations

Terms and abbreviations	Definition
<b>EGID</b>	Federal building identifier (Eidgenössischer Gebäudeidentifikator)
<b>EWID</b>	Federal dwelling identifier (Eidgenössischer Wohnungsidentifikator)
<b>Follow-up survey</b>	Third phase of the rental price survey conducted with the landlord for up to 7 quarters in a row.
<b>FSO</b>	Swiss Federal Statistical Office
<b>Garaio REM AG</b>	Software company developing one of the leading real estate management softwares used by landlords.
<b>Garaio REM</b>	Real estate management software developed by Garaio REM AG.
<b>GWR</b>	Federal Register of Buildings and Dwellings (Eidgenössisches Gebäude- und Wohnungsregister)
<b>Initial survey</b>	Second phase of the rental price survey conducted with the landlord, following the first phase (screening).
<b>mpe EGID-EWID request</b>	Automatic data transmission initial request based on the federal building and dwelling identifiers.
<b>mpe EGID-tenant request</b>	Automatic data transmission follow-up request based on the federal building identifier, floor, room count and tenant's name.
<b>mpe-response</b>	Automatic data transmission response.
<b>mpe-service</b>	Automatic data transmission interface developed by Garaio REM AG for the rental survey.
<b>PRESTA3</b>	Statistics software used by the FSO to manage and calculate the consumer price index data.
<b>Screening</b>	First phase of the rental price survey conducted with private households inhabiting the selected dwellings, preceding the initial survey conducted with the landlord.
<b>sedex</b>	Secure data exchange service provided by the Federal Statistical Office.

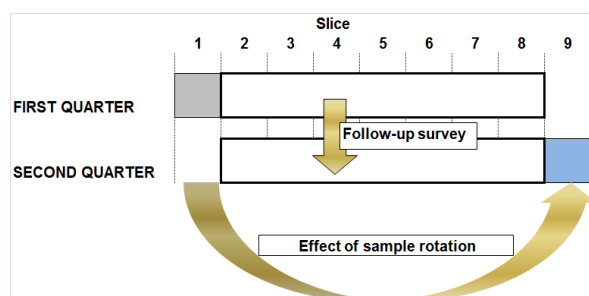
## 2 Rental price index compilation in brief

With a total weighting of over 18%, the rental price index is the most important sub-index in the Swiss consumer price index basket. It is broken down into a rental price index for tenants-occupied housing (15%) and a rental price index for owner-occupied housing (3%)<sup>1</sup>.

The rental price index is based on a rotating panel of approximately 10 000 dwellings, of which 1/8 is replenished quarterly (Figure 1) using a stratified random sample. The stratification combines the building's age and number of rooms (Figure 2).

<sup>1</sup> An "enhanced" rental equivalence method is used. Movements in rent for tenant-occupied properties are imputed to owner-occupied dwellings (dwellings with a special rental status such as subsidised housing, cooperative housing or dwellings with lower rents through family ties are not used for the imputation). The specific expenditure structure of owner-occupiers is used to aggregate the rent sub-indices. The structure of expenditure is indeed significantly different from that of tenants (there are more large units than small ones – for example dwellings with 6 or more rooms older than 20 years have a weight of 13.3% compared with 1.857% for tenants – see Figure 2).

**Figure 1 Rental price index sample rotation**



**Figure 2 Ex-post stratification and weights**

	0-5 years	6-10 years	11- 20years	> 20 years	
Number of rooms	1	0.110%	0.067%	0.100%	3.489%
	2	1.538%	0.586%	0.716%	12.513%
	3	3.285%	1.810%	1.985%	28.031%
	4	3.303%	2.497%	3.141%	24.411%
	5	0.821%	0.751%	1.214%	7.208%
	6	0.140%	0.145%	0.282%	1.857%

Source: FSO – Consumer Price Index, 2021

© FSO 2022

For a more detailed description of the formulas used to calculate the rent price index see appendix 7.4.

## 2.1 Rental collection

The rental index is the only sub-index within the Consumer Price Index based on a random sample. The new housing units replacing one-eighth of the existing sample each quarter are selected on the basis of a dwelling survey framework designed specifically for the rental index (which is a combination of the Federal Register of Buildings and Dwellings and the person and household survey sampling frame). Once the new dwellings in the sample are selected (about 6000), an initial questionnaire (screening) is sent to the tenant in order to get the name of the landlord<sup>2</sup>. Following the screening, an initial survey is carried out in contact with the landlord, who is requested to fill out a detailed form, indicating in addition to the rent all the characteristics of the dwelling: lease terms (duration, changes of tenants, type of lease), structural details about the dwelling (type of building, number of rooms, surface area, floor, year of construction,) and information on the owner type. Then a follow-up survey is conducted quarterly to compile any changes relative to the information collected in the previous period and the grounds for rent adjustments.

About 70% of the landlords are real estate management companies. Larger real estate managers sometimes have to fill out dozens or even hundreds of online or paper questionnaires per quarter, which generates a large workload. For this reason, in 2020, the FSO and a provider of a property management software launched a joint project for the automatic transmission of data between the FSO and professional landlords. This project aimed to both modernise collection techniques and reduce the burden on professional landlords. Since May 2023, part of the data required for the rental price index has been transmitted electronically, as XML files via a secure asynchronous data

<sup>2</sup> Landlords are preferred over tenants insofar as they possess precise information about the dwelling that they let, are generally better equipped administratively and are used to participating in the survey regarding the rental index. Moreover, development of rents occurring as tenants change can be observed.

exchange, from the first major property management company.

### 3 Introducing automatic data transmission for the rental price index

#### 3.1 Automatic data transmission in a nutshell

In Switzerland many real estate management companies manage their real estate portfolio and rental agreements with a software. There is a limited number of software companies which develop real estate management software. This means that many real estate management companies use the same software to manage their rental agreements. An automatic data transmission was therefore developed in collaboration with the leading software development company on the Swiss market: Garaio REM AG. Working with the software company instead of a single real estate management company will allow the FSO to ultimately access the rental data of all Garaio REM AG clients using the same technical interface. Also, the real estate management companies with the largest number of questionnaires in the rental index sample use the Garaio REM software.

The interface to extract the data (mpe-service) was developed by Garaio REM AG. It allows the FSO to connect with their software to access the landlord database. The FSO provided its secure data exchange platform for the data transmission ([www.sedex.ch](http://www.sedex.ch)). While the data transmission mechanism is simple, its development was a complex task including many hurdles, which will be presented later in this paper. In summary, the data transmission is triggered by a data transmission request sent by the FSO for each dwelling included in the rental survey. After a few seconds a response is sent back to the FSO, with or without rental data, depending on whether the dwelling has been found in the client database of the software company. The request and response mechanisms are explained in detail in chapter 4.1.

#### How automatic data transmission differs from data collection through questionnaires

One big difference between these two data collection methods is related to the interpretation of data. For questionnaires, the interpretation is done by the respondent, while for automatic data transmission the interpretation is done by the data recipient. The real estate managers subjectively interpret the questions. For automatic data transmission, the interpretation is centralised and set up beforehand in the price statistics software. Hence the interpretation process is more consistent, but more data is needed. While 121 variables are sent via automatic data transmission, only 55 unique variables are included in the three questionnaires.

Let's take a look at how the owner type is determined via questionnaire:

*Please indicate the owner type of the managed dwelling.*

- One or more private individuals (*OWNER\_TYPE\_NB = 0*)
- Public authorities (*OWNER\_TYPE\_NB = 1*)
- Housing cooperative (*OWNER\_TYPE\_NB = 2*)
- Construction company or real estate company (*OWNER\_TYPE\_NB = 3*)
- Other AG / GmbH / cooperative (*OWNER\_TYPE\_NB = 4*)
- Other type of owner (please add): (*OWNER\_TYPE\_NB = 5*)
- Owner not known (*OWNER\_TYPE\_NB = 6*)

This owner type categorisation does not exist in Garaio REM's database. To obtain the same data, more variables and interpretation rules are needed. The results with data transmission turn out to be more consistent compared to the questionnaire. In fact, the real estate manager's answer cannot be verified, whereas the data obtained through automatic data transmission can. (See chapter 4.5.3 for more details on which variables were selected and which interpretation rules were applied to obtain the data for this question).

## 3.2 Project plan

### 3.2.1 Initialisation phase 2018-2022

The project was launched during summer 2018. An initial meeting with Garaio REM AG was set up to find out whether the rental price survey was suitable for automatic data transmission. Since the software development was to be half financed by the FSO and half by the clients of Garaio REM AG, the first step was to present the idea to the largest Garaio REM AG clients who would in turn benefit the most from automatic data transmission (see chapter 5.2.1 for an overview of time savings for real estate management companies). As soon as the financial aspects were agreed upon, a procurement process was started by the FSO. The contract with Garaio REM AG was established in June 2020.

#### Focus of the initialisation phase

These were the most important questions addressed during the initialisation phase:

(1) determine the importance of each variable currently collected by questionnaire for the compilation of the rent price index.

All variables were classified according to their importance for the rental price survey:

- A) required variable for calculation and hedonic model without secondary source
- B) required variable for calculation and hedonic model with secondary source
- C) useful variable for checking data before index calculation (plausibilisation)
- D) potentially useful variables for the update of the hedonic rental price model

(2) identify which data currently provided through the paper or online questionnaires are available in the Garaio REM database. While some data is available in the in the appropriate format and can be extracted directly without transformation, most variables need to be interpreted before they can be used in the rental price index. Finally, some data is not available through automatic transmission.

(3) evaluate the quality of the data saved in the Garaio REM database by their clients (see chapter 6.3).

(4) develop interpretation rules for data in order to import the necessary information into the statistics software PRESTA3 (see chapter 4.5).

From 20 June 2020 a coordination meeting was organised every three weeks between Garaio REM AG and the FSO project team. The collaboration was very successful and lead to a review of the survey process for the automatic data transmission. No fewer than 60 meetings took place before the new data collection technique was introduced with the first Garaio REM client. The development of the automatic data transmission interface took more than two years. A major software update will take place in 2025/2026 which will allow to include even more Garaio REM clients to the automatic data transmission project.

### **Reviewed survey process**

For the FSO project team, the most straight forward solution for the automatic data transmission was to integrate this new data collection technique into the current rental survey process which consists of three distinct phases:

- a screening phase targeted at the private households occupying the selected dwellings to identify the permanently rented dwellings and their landlord,
- an initial survey with the landlord,
- 7 follow-up surveys with the landlord to compile any changes.

Each phase has a different questionnaire. Some variables are the same for the three phases, but many of them are distinct. The project team's initial idea was to set up three different data requests, which could be executed during each phase.

After several meetings and brain storming with the Garaio REM AG software developers a new approach emerged, which made the automatic data transmission process simpler. Instead of developing a response for each survey phase with distinct variables, one response including all rental survey variables was developed. The responses can be requested anytime during the data collection process.

### **3.2.2 Transition and introduction phase 2023**

Between January and March 2023, the transition phase started in collaboration with the largest real estate company in the rental survey sample: "Company X"<sup>3</sup>. During the transition phase, Company X transmitted data automatically and filled out the usual questionnaires in parallel.

#### **Focus of the transition and introduction phase**

These were the most important questions addressed during the transition phase:

- (1) ensure that different data collection techniques deliver the same data / results.
- (2) ensure that the interpretation rules for the automatic data transmission had been set up correctly.
- (3) improve the dwelling matching rate.

In order to guarantee an accurate calculation of the rental price index, it is important to ensure that during follow-up surveys, data is delivered for the same sampled dwellings, regardless of the data collection technique. Therefore, an important task of the FSO team was to check whether the data delivered through automatic data transmission was consistent with the data previously delivered through questionnaires. During the transition phase a total of 10 mismatched dwellings were identified out of 213. These were treated in an appropriate way in order to cancel the effect of rental price changes on the rent price index that are due to mismatches. The reasons why some dwellings mismatched are described in more detail in chapter 4.4.

The automatic data transmission was officially introduced in April 2023. All imported data was thoroughly checked for 3 consecutive survey quarters. The end of quarter quality control mechanism was expanded to detect errors related to automatic data transmission and the interpretation rules were continuously optimised.

The resources and effort required to develop the 2020-2022 interface were far greater than originally planned, both on the part of the commissioned software provider Garaio REM AG and the FSO. For this reason, the data delivery could not yet be extended to all clients right away, as originally planned.

---

<sup>3</sup> Anonymised landlord for data protection.

### 3.2.3 Next steps 2024-2029

In order to extend the automatic data transmission to all Garaio REM AG clients, a new contract will be signed between the FSO and Garaio REM AG. The contract includes the data preparation and delivery of approx. 1400 additional dwellings. These will be delivered step by step from 2025, amounting to around 200-300 dwellings per year. The FSO will have to check the data quality of each client and may need to make adjustments to the interpretation rules.

## 4 Practical aspects of automatic data transmission

For the automatic data transmission, all relevant data is sent through a request and response mechanism (see chapter 4.1). The data requests and responses are transmitted in XML format through the FSO secure data exchange platform called sedex (two sample XML-requests and an XML-response are listed in the appendices in chapter 7) to a data transmission interface developed by Garaio REM AG, called mpe-service. The mpe-service sends back a response for each request via sedex within a few seconds. The response file usually includes a return code (see chapter 4.2) and the required data, if applicable.

For each dwelling included in the rental survey an individual request is created automatically through the price statistics software PRESTA3. The requests are sent at the beginning of each survey phase by clicking on the “send request” step of the rental price survey workflow (steps 11, 19 and 25 of Figure 3 below). The requests are made before sending out the paper and online questionnaires, in order to be able to include unmatched dwellings afterwards. Every quarter, approximately 5-8% of dwellings which are surveyed through automatic data transmission remain unmatched (see chapter 4.3).

### **How automatic data transmission allows the FSO to skip the screening step and reduce non-response**

Automatic data transmission can also be used for the first phase of the rental price survey conducted with private households. As described in chapter 2.1, the rental price index is based on a rotating panel of dwellings, of which 1/8 is replenished quarterly (step 1 screening). Automatic data transmission allows the FSO to skip the screening for the new dwellings that are added to the sample every quarter, by simply sending requests for all new dwellings (approx. 6000<sup>4</sup> per quarter) at the beginning of the survey phase. If there are dwellings managed by the landlords who participate in automatic data transmission, they will be identified. This also reduces non-response from households. Whereas prior to automatic data transmission the average number of new dwellings entering the sample for Company X was 40, after the introduction of automatic data transmission this number rose to 70. The response rates for the initial survey and the follow-up survey are almost 100% for Company X. A similar effect is expected for the other Garaio REM AG clients.

If necessary, individual requests for one or more dwellings can be sent anytime during the quarter through the price statistics software interface by entering the dwelling survey number in the mask and a TenantSearchDate (see Figure 3 below).

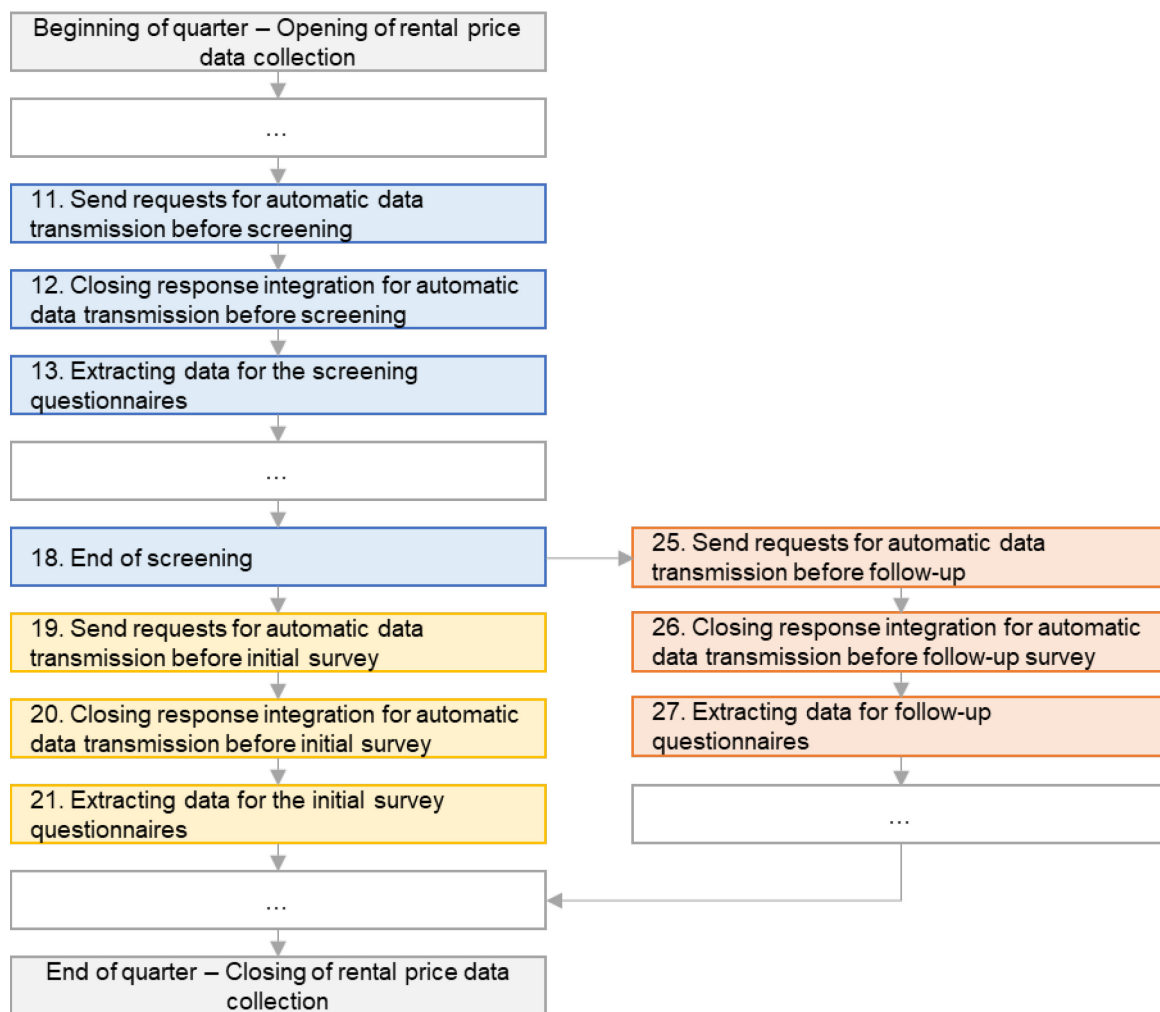
---

<sup>4</sup> To reach approximately 1/8 of the total sample of 10 000 dwellings every quarter, approximately 6000 new dwellings need to be drawn. This number is high because the rate of eligible dwellings is low (approximately 20-23%). On the one hand, approximately 40% of dwellings need to be excluded because they are owner-occupied or correspond to other categories that are not eligible. On the other hand, the non-response rate among private households is over 20%.



**Figure 3 Interface in PRESTA3 for sending requests**

**Figure 4 Integrating the automatic data transmission into the rental survey workflow in PRESTA3**



## 4.1 Request and response mechanism

### 4.1.1 The EGID-EWID request

The requests for the rent survey are triggered by the FSO. The basis for the requests is the sample of the rent survey. To identify the dwellings managed by Garaio REM AG clients, the FSO sends an initial request (mpeRequest\_FindUnit\_EGIDEWID, see appendix 7.1 for the current XML structure) with the following variables:

- survey date (referenceDate),
- federal building identifier (EGID),
- federal dwelling identifier (EWID).

### 4.1.2 The EGID-tenant request

If a building can be identified in the first step using the EGID, but no specific dwelling can be matched, the FSO sends a follow-up request (mpeRequest\_FindUnit\_EGIDTenant, see appendix 7.2 for the current XML structure) with the following variables:

- survey date (referenceDate),
- federal building identifier (EGID),
- federal dwelling identifier (EWID)<sup>5</sup>,
- tenant name (officialName),
- first name of tenant (callName),
- floor\*,
- number of rooms (roomCount)\*,
- tenant search date.

\*If the dwelling can still not be identified, the mpe-service processes the same request again, but without floor and number of rooms. This allows the interface to find a tenant even if floor and room count don't match.

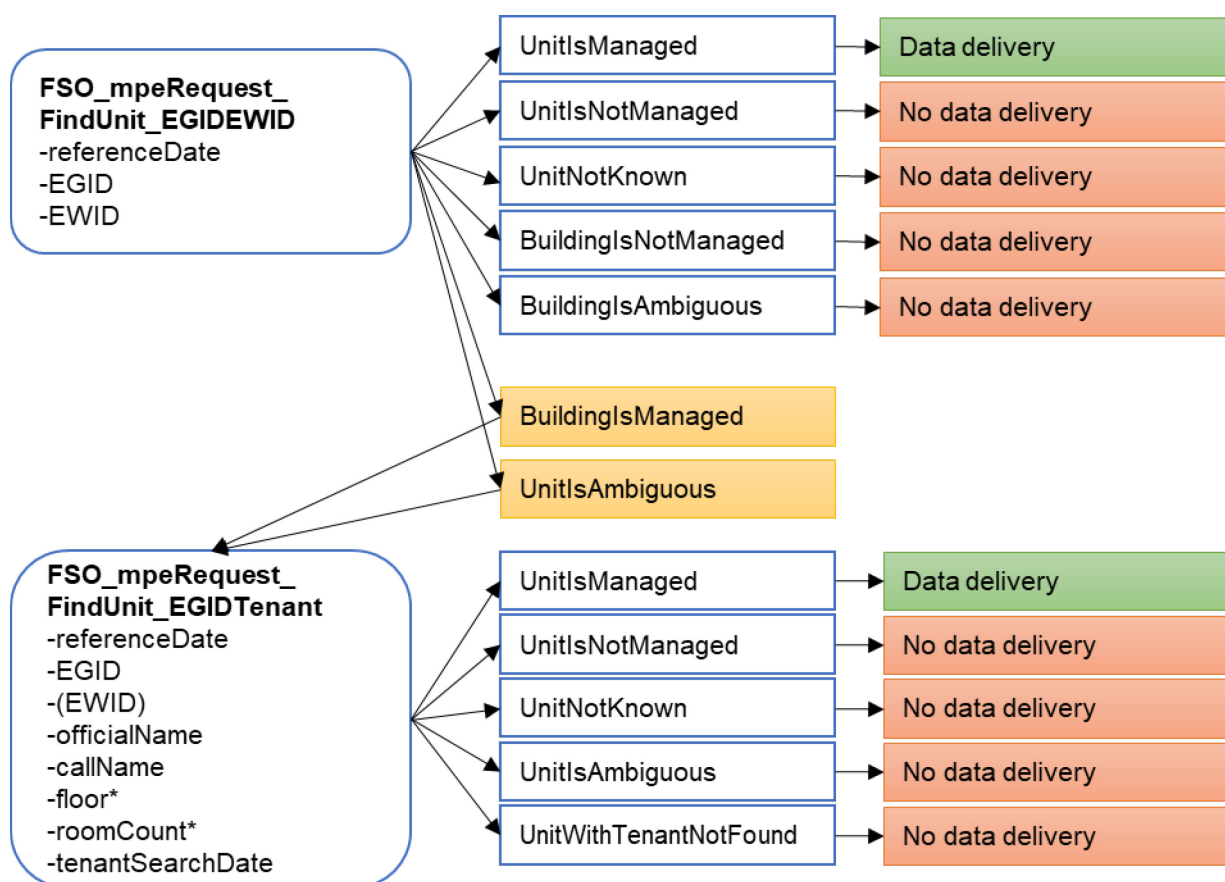
The follow-up request is necessary because the dwelling identifiers (EGID and EWID) are not fully covered in the landlord database. Currently, approximately 35% of matched dwellings match through the second type of request which involves the tenant's name.

When a dwelling is matched unequivocally (UnitsManaged), Garaio REM provides the FSO with all useful data for the successful implementation of the rental price survey. See appendix 7.3 for the current XML structure of delivered data.

---

<sup>5</sup> EWID is sent but not used by the mpe-service.

Figure 5 Request and response mechanism



### Reference date and tenant search date

#### Survey reference date

The rental price index is calculated based on a reference date. Due to the sample size and the step-by-step process, the rental price survey requires a lead time of around 2-3 months. The reference date is therefore always in the future.

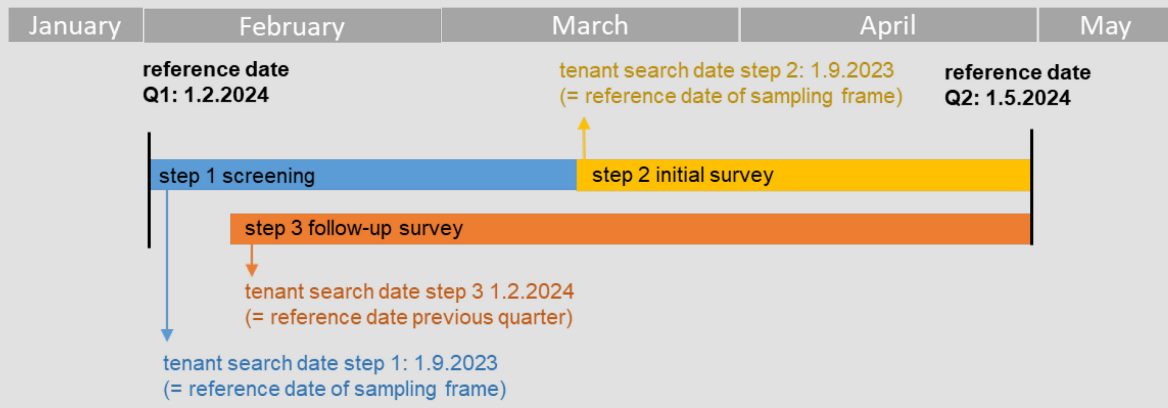
#### Tenant search date

A tenant search date was introduced as an additional parameter for the EGID-tenant request, in order to avoid matching-problems related to the tenant's name used to identify the dwelling and changes related to it. The survey reference date is not always appropriate to match dwellings based on tenants' names:

- **For survey step 1 and 2**, there is a 7-month interval between the date of the data request and the reference date of the tenant's name used to identify the dwelling. Due to the survey process it is not possible to get more up-to-date tenant names from the person and household sampling frame. According to Tenant search date, for the Q2 survey with reference date 1.5.2024, the mpe-service identifies the dwelling based on the tenant search date 1.9.2023 (which corresponds to the reference date of the sampling frame). The data returned however, is based on the survey reference date 1.5.2024.
- **For survey step 3**, there is a 3-month interval between the reference date of the last quarter and the date of the data request of the current quarter. According to Figure 6, for the Q2 survey with reference date 1.5.2024, the mpe-service identifies the dwelling based on the tenant search date 1.2.2024 (which corresponds to the reference date of the previous

quarter). The data returned however, is based on the survey reference date 1.5.2024.

**Figure 6 Tenant search date**



## 4.2 Return codes

In order to reflect specific cases and clarify the steps that need to be taken as a consequence, several return codes were developed. The only return code that will deliver all required data for the rent survey is “UnitIsManaged”. This response means that a dwelling matches unequivocally in the Garaio REM landlord database. Another important code that is used is “UnitIsNotManaged” which doesn’t mean that the dwelling was not found, but indicates a possible change of landlord. In order to find out who the new landlord is, the previous landlord needs to be contacted. As of today, there is no automated solution possible to receive the new landlord’s address through the automatic data transmission process.

**Table 1 Overview and interpretation of return codes**

Return code	Interpretation
<b>BuildingsManaged</b>	Dwelling was not found at referenceDate but a building (EGID) that is under management was found. This means that no matching EWID was found in the landlord’s database and the dwelling must be searched for using additional search parameters (floor, number of rooms, tenant name).
<b>BuildingsNotManaged</b>	Dwelling was not found at referenceDate but a building (EGID) was found. The building is not under management at the specified reference date. This code generally indicates a change of landlord.
<b>BuildingsAmbiguous</b>	Dwelling was not found at referenceDate but several buildings (EGID) were found with the specified EGID. This response can occur due to a constraint in an older version of the software for specific cases, where the same EGID is entered for several managed buildings in the same area. This problem will be solved with the major software update that will be implemented in 2025/2026.
<b>UnitIsManaged</b>	Dwelling was found at referenceDate. The dwelling is under management at referenceDate. It is only with this code that all required survey data is sent to the FSO.
<b>UnitIsNotManaged</b>	Dwelling was found at referenceDate. The dwelling is not under management at referenceDate. This generally indicates a change of landlord.

<b>UnitNotKnown</b>	Dwelling was not found at referenceDate. The building (EGID) was not found.
<b>UnitIsAmbiguous</b>	Dwelling could not be found unequivocally based on the request data at referenceDate. Possible reasons are: - The search matches several dwellings, because of erroneous data in landlord database. - The tenant mentioned in the person element rents several dwellings that match the query.
<b>UnitWithTenantNotFound</b>	This return code is used for EGID with Tenant request when no Tenant was found at tenantSearchDate for specific Building by EGID

<b>Error codes</b>	<b>Interpretation</b>
<b>MinimalRequirementNotMet</b>	Used endpoint does not match minimal requirement
<b>NetworkError</b>	Network error occurred
<b>UnexpectedError</b>	Unexpected error occurred
<b>RequestIsErroneous</b>	Not xsd valid
<b>XmlIsMalformed</b>	Not a valid XML format

### 4.3 Why some dwellings do not match

During the transition phase of automatic data transmission, the matching rates were slightly lower than today because of some issues that were resolved in collaboration with the real estate Company X and Garaio REM AG.

There are currently two types of cases left for which dwellings don't match:

#### 1. A dwelling that doesn't currently match and never matched before

Some dwellings won't match, even if they are currently under management of a Garaio REM client. At survey stage 1 and 2 the landlord has been identified by the FSO. Most of the time the problem is related to data quality issues with the used building identifier EGID. In fact, only the EGID is used to identify a building. The building address is not used. EGID is used because it is a unique identifier, which is easy to search for (numeric value between 1-900 000 000). An address, in contrast, is more difficult to match. The federal building identifier is the official identifier used in the GWR, it is backed by law and the landlords record it in the Garaio REM software. The data coverage tends to depend on whether the use of EGID and EWID on rental agreements are mandatory by law in a Canton (region). It is not yet mandatory in most Cantons.

#### 2. A dwelling that doesn't currently match but has matched before

Furthermore, some dwellings don't match anymore in the current quarter, although they matched in the previous quarter. There are two possible reasons:

1. The dwelling was previously matched using the EGID-tenant request and a tenant change took place which was recorded by the landlord after the last request date. As the new tenant's name was not yet sent to the FSO with the previous data transmission, the old tenant's name used by the FSO to identify the dwelling does not match the new tenant's name anymore at the new reference date. In such cases the automatic data transmission interface usually sends a "UnitWithTenantNotKnown" response. The request can then be repeated using an older tenant search date, which will update the tenant's name and deliver the data for the current reference date.
2. The landlord changed since the last request date and the new landlord is not a Garaio REM

AG client. This is typically communicated to the FSO through the return code “UnitIsNotManaged” which means that the dwelling was found but is not under management at the indicated reference date.

The real estate management companies are constantly working on improving data quality, as the EGID / EWID identifiers are also used for other data transmission projects to authorities or electricity companies (automatic relocation reports). It is therefore expected that the number of unmatched questionnaires will decrease over time.

#### 4.4 Why some dwellings mismatch

During the transition phase a total of 10 mismatches were identified out of 213 sampled dwellings. The mismatches were identified based on differences in number of rooms, floor, surface, tenant names, beginning of tenant agreement and tenant agreement reference. Each case was analysed with the help of Garaio REM AG and Company X. There were two explanations for mismatched dwellings during the transition phase:

1. In the first case, the automatic data transmission sends data for a different dwelling, because the federal building and dwelling registry (GWR)<sup>6</sup> and Company X do not attribute the federal dwelling identifier (EWID) to the same dwelling in the building. Therefore, the characteristics of the dwelling no longer match, even though they were correctly provided by questionnaire.
2. In the second case, the automatic data transmission sends data for the correct dwelling. However, the dwelling characteristics do not match because the questionnaire provided data for a different dwelling.

Before January 2023 the federal building and dwelling identifiers were not printed on the rental survey questionnaires and therefore not yet used to identify a sampled dwelling. The sampled dwellings were identified mainly through the tenant’s name, number of rooms and floor.

Overall, mismatched dwellings are problematic only for the transition phase from one data collection technique to another and for a follow-up survey. Since the quality control mechanism was expanded with the introduction of automatic data transmission, changes during follow-up surveys can now be identified and appropriate corrections are made. If the data provided for a newly sampled dwelling does not match the GWR from the beginning, the sample will no longer be completely randomised in a sense, but the rental prices are still measured accurately.

Finally, some matching issues cannot be solved within the current software architecture:

- The first issue is related to the EGID and the data architecture of the Garaio REM software. The EGID can be entered at different levels in the database (property, building or dwelling). The mpe search algorithm searches exclusively at the property level which is usually reported on the building and house levels. However, there are cases where a property includes several buildings with several EGIDs that cannot all be entered at the property level. It is currently impossible to match such dwellings, if the property EGID doesn’t match the building EGID. This issue can be solved with future software updates.
- The second issue is related to external management mandates. For some dwellings Company X has external management mandates, which means that the data is not directly stored in the Garaio REM database and makes it impossible to access the data. This issue can be solved when more Garaio REM clients will be included in the automatic data transmission.

---

<sup>6</sup> The federal dwelling identifier is attributed by the local town authorities and then sent to the landlords. When there are several dwellings with the same number of rooms on the same floor, it is sometimes challenging to identify the correct dwelling if the tenants cannot be identified.

## 4.5 Data interpretation

As mentioned in chapter 3.2.1, one of the main focus points of the initialisation phase was to develop interpretation rules to correctly import the data into the statistics software. Only a few variables can be imported directly without any treatment (such as, tenant agreement reference number, tenant last name or tenant name). The majority of variables need more or less complex interpretation rules. Hereafter some examples of challenging variables and their interpretation rules will be described in more detail.

### 4.5.1 Interpretation rule for ADDRESS\_KIND

One of the most difficult questions to ascertain through automatic data transmission was whether the selected dwellings were eligible for the rental survey. This question is usually answered by the household during the screening phase. As only the net rents of permanently for residential purposes rented dwellings are considered, the contacted households need to confirm that

- (1) they do not own the occupied dwelling (*APARTMENT\_OWNER\_FLG*),
- (2) the dwelling is their principal residence<sup>7</sup> (*ADDRESS\_KIND*), and
- (3) they rent an unfurnished dwelling<sup>8</sup> (*ADDRESS\_TYPE*).

These three criteria turned out to be quite difficult to ascertain through automatic data transmission because no data fields that correspond to those questions exist in the landlord databases. Unfortunately, it cannot be assumed that all objects managed by the real estate company are non-owner-occupied dwellings that are unfurnished and used as principal residences. The database also contains furnished dwellings, secondary residences, owner-occupied dwellings and rental contracts paid by the employer which need to be excluded. This information can only be inferred by exclusion through a complex decision tree using relevant available data.

**To determine whether the dwelling is used as a principal residence the following formula is used:**

```
if dwelling/idxCategory == "INDUS" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 8, 9,
10, 11, 12, 13,
14, 15, 16, 17, 18, 19, 20, 21, 23, 28, 29, 30, 31, 32, 34, 36, 37, 38]:
    ADDRESS_KIND = 0
elif ((dwelling/idxCategory == "HOUSE" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 9,
10, 11, 12, 13]) or
(dwelling/idxCategory == "APPT" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 8, 9, 10,
11])) and
occupancy/tenancyAgreement/isTenancyResidential == True and
occupancy/tenancyAgreement/isSecondaryResidence == True:
    ADDRESS_KIND = 1
elif ((dwelling/idxCategory == "HOUSE" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 9,
10, 11, 12, 13]) or
(dwelling/idxCategory == "APPT" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 8, 9, 10,
11])) and
occupancy/tenancyAgreement/isTenancyResidential == False
    '''Wohnung in WFS 5 und Feld occupancy/tenancyAgreement/usage manuell bearbeiten.
    Eventuell Apartment_NB in (separate) Log-Datei schreiben.'''
elif (dwelling/idxCategory == "INDUS" and dwelling/idxType in [22, 25, 27, 33, 35]) or
(dwelling/idxCategory == GASTRO and dwelling/idxType in [1, 16]):
    ADDRESS_KIND = 2
elif ((dwelling/idxCategory == "HOUSE" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 9,
10, 11, 12, 13]) or
```

<sup>7</sup> business addresses, vacation homes, retirement homes, hotels, hospitals or assisted living facilities for the elderly are excluded.

<sup>8</sup> tenants of a furnished dwelling, subtenants, tenants of a dwelling paid by the employer, occupants of an apartment for free or tenants of a farm are excluded.

```

(dwelling/idxCategory == "APPT" and dwelling/idxType in [1, 2, 3, 4, 5, 6, 7, 8, 9, 10,
11])) and
occupancy/tenancyAgreement/isTenancyResidential == True and
occupancy/tenancyAgreement/isSecondaryResidence == False:
    if occupancy/tenancyAgreement/tenantKind == "LEGAL":
        ADDRESS_KIND = 0
    elif occupancy/tenancyAgreement/tenantKind == "PHYSICAL":
        ADDRESS_KIND = 3 # SI PHYSICAL, alors contrôler que ADDRESS_KIND est différent
de 0 (Geschäftsadresse), sinon WFS 5
    elif dwelling/idxCategory == "AGRI" and dwelling/idxType in [1, 2, 3]:
        ADDRESS_KIND = 3
    elif occupancy/tenancyAgreement/usage in ("appartement secondaire",
"Laden/Wohnung/Lager", "Zweitwohnung"):
        ADDRESS_KIND = 1
    elif occupancy/tenancyAgreement/usage = "Arztpraxis":
        ADDRESS_KIND = 0
    else:
        '''Papierformular oder eSurvey senden. Eventuell Apartment_NB in (separate) Log-
Datei schreiben.'''

```

Where:

Data received from Garaio REM AG	Meaning of data received
dwelling/idxCategory	IDX3.01 is a format used to publish standardised real estate announcements. idxCategory defines the object category ('APPT','HOUSE','INDUS', etc.).
dwelling/idxType	Defines the object type for each object category (for APPT, type 1 =apartment; type 2=duplex, type 3=attic; etc.).
occupancy/tenancyAgreement/ <b>isTenancyResidential</b>	Is an interpreted information. True means there is a residential tenancy agreement related to the object.
occupancy/tenancyAgreement/ <b>isSecondaryResidence</b>	Represents a data field (flag) indicating a vacation home.
occupancy/tenancyAgreement/ <b>tenantKind</b>	Represents a data field indicating whether the tenant is 'PHYSICAL' (physical person) or 'LEGAL' (business)
occupancy/tenancyAgreement/ <b>usage</b>	Represents a non-mandatory and non-standardized text field where information on the usage of the object can be entered.

#### 4.5.2 Interpretation rule for BUILD\_YEAR\_NB

In the Garaio REM software there is more than one field to enter the construction year of a building. It was therefore decided that the mpe-service would transmit the latest year of construction available, which was assumed to be the most accurate one. During the transition period, it was discovered that for 20% of the sampled dwellings there was a difference between the construction year delivered through automatic transmission and the construction year available in the federal building and dwelling registry. For about 10% of these dwellings the difference was more than 10 years. An in-depth analysis of the data recorded for these dwellings in the Garaio REM software showed that sometimes real estate managers record a full-scale renovation as a new construction date and enter these in the same fields. To overcome this issue, it was decided to add more variables to the delivered data in order to avoid importing erroneous data. In addition to the latestYearOfConstruction the mpe-response now delivers every entered construction year and every related remark for more accurate interpretation. If the difference between the construction year from GWR and the delivered construction year is bigger than 5 years, the FSO employee will manually check all delivered construction years and make the necessary corrections in the price statistics software.



**To determine the year of construction of the dwelling the following formula is used:**

```

if MPI-Formular = "FIRST":
  if building/latestYearOfConstruction is not None:
    if GWR_CONSTRUCTION_YY is None
    or building/latestYearOfConstruction - GWR_CONSTRUCTION_YY > 5:
      set WFS = 5
      and add "Contrôler l'année de construction." in COMMENT_BLOB
    else:
      BUILD_YEAR_NB = building/latestYearOfConstruction
  elif GWR_CONSTRUCTION_YY is not None:
    BUILD_YEAR_NB = GWR_CONSTRUCTION_YY
  else:
    BUILD_YEAR_NB = 1950
if BUILD_YEAR_NB > BEGIN_RENT_NB:
  set WFS = 5
  and add "Vérifier l'année de construction, car début de bail plus ancien que l'année
de construction." in COMMENT_BLOB

```

Where:

Data received from Garaio REM AG	Meaning of data received
building/latestYearOfConstruction	Represents the most recent year of construction entered by the landlord

Rental survey data used	Meaning
GWR_CONSTRUCTION_YY	Year of construction of the building from the federal building and dwelling registry used to check consistency of the data delivered by Garaio REM. GWR_CONSTRUCTION_YY is not available for all buildings.
BEGIN_RENT_NB	Represents the beginning of the rental agreement delivered by Garaio REM as occupancy/tenancyAgreement/validFrom
WFS	Is the survey work flow status of the dwelling. In the interpretation process some conditions are determined which set the dwellings status for manual treatment (WFS = 5).

(1950 is used as a default value for missing data in the rental price index).

### 4.5.3 Interpretation rule for OWNER\_TYPE\_NB

The following information is delivered through data transmission to determine the owner type:

- A. if the owner is a physical person, the information “physical person” is delivered;
- B. If there is a unique enterprise identification number UID<sup>9</sup> available in the landlord database, the UID is delivered;
- C. if no UID is available, the name and address of the selected dwelling’s owner is delivered.

This information allows the FSO to interpret the owner type and ensure it is correctly and consistently categorised for all dwellings. In order to categorise the owner, the interpretation rules use information from the General Classification of Economic Activities (NOGA) in addition to the delivered data.

**To determine the type of owner of the dwelling the following formula is used:**

<sup>9</sup> The unique enterprise identification number (UID) is a unique identifier allocated to every enterprise active in Switzerland. It allows the authorities to automatically retrieve an enterprise’s reference data when in contact with the enterprise.

```

if MPI-Formular = "FIRST":
    if owner == physicalPerson:
        OWNER_TYPE_NB = 0
    elif owner == namedOrganisation:
        '''Exportieren der namedOrganisation-Felder in eine Datei, die
        in einem externen System importiert werden kann, um die UIDs zu
        ermitteln. Anschliessend kann die Datei mit den ermittelten UIDs
        in PRESTA 3 importiert werden, um das Feld OWNER_TYPE_NB auszufüllen.'''
        set WFS = 5
        and add "Type de propriétaire à déterminer manuellement ou à l'aide du fichier
        *.csv !" in COMMENT_BLOB
    elif owner == identifiedOrganisation:
        '''benutze die View 'V_PREIS_LEG_UNITS' und ermittle die Felder :
        (LegalformCode, NameLegalUnit, NogaOutputENT) anhand den Felder
        (uidOrganisationIdCategory + uidOrganisationId)'''
        if UIDID not (valid or found):
            set WFS = 5
            and add "Type de propriétaire à déterminer manuellement !" in COMMENT_BLOB
        elif LegalformCode in [17, 20, 21, 22, 23, 24]:
            OWNER_TYPE_NB = 1
        elif LegalformCode == 8 and NameLegalUnit like
            ["Wohngenossenschaft",
            "Wohnbaugenossenschaft",
            "coopérative d'habitation",
            "coopérative de construction"]:
            OWNER_TYPE_NB = 2
        elif NogaOutputENT in [412001, 412002, 412003, 439905, 681000, 682001, 682002,
        683100, 683200]:
            OWNER_TYPE_NB = 3
        elif LegalformCode in [6, 7, 8, 14, 15, 16, 3, 4, 5, 10]
            OWNER_TYPE_NB = 4
        else
            OWNER_TYPE_NB = 5

```

Data received from Garaio REM AG	Meaning of data received
owner	the delivered data will either include a flag indicating a 'PHYSICAL PERSON', or a UID of an 'IDENTIFIED ORGANISATION' or the full address of a 'NAMED ORGANISATION' if no UID is recorded in the landlord's database.
NogaOutputENT, LegalformCode, NameLegalUnit	Are variables from the General Classification of Economic Activities (NOGA) that were attributed to the owner type category.

## 5 Results

### 5.1 Matching rates

The matching rates for the current quarter (May 2024) are very satisfying for Company X with 96% for the initial and follow-up surveys and 89% for the screening phase. The matching rates improved significantly during the transition phase, thanks to the optimization of the matching mechanism (see which measures below Figure 7). For the follow-up survey the matching rate increased from 66% during the first trial in February 2023 to 96% in May 2024.

Figure 7 Matching rates

Matching rates per quarter and phase	5.2024	2.2024	11.2023	8.2023	5.2023 (Mar23)	5.2023 (Feb23)
Screening phase	89%	91%	87%	91%	82%	82%
Initial survey	96%	98%	94%	97%	70%	70%
Follow-up survey	96%	95%	97%	96%	89%	66%

Number of matched dwellings (UnitIsManaged)	5.2024	2.2024	11.2023	8.2023	5.2023 (Mar23)	5.2023 (Feb23)
Screening phase	71	52	68	71	41	41
Initial survey	77	56	72	76	37	37
Follow-up survey	319	310	290	250	212	160
<b>Total automatic data transmission sample</b>	<b>396</b>	<b>366</b>	<b>362</b>	<b>326</b>	<b>249</b>	<b>197</b>

Number of non-matched dwellings	5.2024	2.2024	11.2023	8.2023	5.2023 (Mar23)	5.2023 (Feb23)
Screening phase	9	5	10	7	9	9
Initial survey	3	1	5	2	16	16
Follow-up survey	15	15	10	10	26	83
<b>Total non matched sample</b>	<b>18</b>	<b>16</b>	<b>15</b>	<b>12</b>	<b>51</b>	<b>51</b>

#### Optimisations to improve the matching rate

The following measures helped to improve the matching rate:

- (1) optimising the search algorithm to match tenant names.
- (2) expanding data coverage in landlord's database to include objects with administrative management mandates.
- (3) adding an additional search round with the tenant's name when the result of the mpeRequest\_FindUnit\_EgidEwid is UnitIsAmbiguous.
- (4) adding an additional search parameter "tenant search date" to change the tenant reference date, if necessary.

### 5.1.1 Procedure for dwellings that do not match

Once per quarter a list with unmatched dwellings is sent to a contact person from Company X. The list includes all the variables that were used to identify the dwelling and its address. Company X checks the data, makes corrections (mostly EWID) in their database and sends back the list. Sometimes the FSO will correct some tenant names based on the inputs of Company X. Afterwards all unmatched requests are resent. The corrections made on both sides usually allow some more dwellings to match. All dwellings which still cannot be matched are sent to Company X by online questionnaire. Every quarter the FSO keeps track of the rate of online questionnaires sent to company X. There are usually under 20 dwellings per quarter that still need to be sent by questionnaire.

### 5.2 Increase of efficiency in data management

The rental survey is a major burden for real estate managers. The largest real estate management companies receive hundreds of questionnaires to complete every quarter. The use of electronic data allows them to be significantly relieved of this burden. Moreover, this new data collection technique allows for a more efficient treatment of data by the FSO. The FSO estimated the total time saved per quarter for the real estate companies and for the FSO employees handling the survey.

### 5.2.1 Time saved by real estate management companies

As mentioned in chapter 5.1, 396 questionnaires of Company X are now processed through automatic data transmission: 77 initial surveys and 319 follow-up surveys. The FSO asked some of the largest real estate management companies how much time they spend on average to fill out a questionnaire. The result was approximately 10 minutes for an initial survey and 5 minutes for a follow-up survey. If Company X had filled out their 396 questionnaires manually, its employees would have spent a total of 39 hours, which represents almost one working week, whereas the time required by Company X for electronic data transmission is practically zero. A minimum effort is required per quarter to process the non-matched questionnaires (last quarter there were 18 questionnaires that couldn't be processed automatically).

### 5.2.2 Time saved by private households responding to the screening phase

Automatic data transmission also allows the FSO to contact fewer private households during the screening phase and at the same time to increase the number of new dwellings that enter the sample, thanks to bypassing non-response.

### 5.2.3 Time saved by the FSO

For the rental survey there currently are three data collection techniques: paper questionnaire, online questionnaire and automatic data transmission. Compared to the data collection via online questionnaire there are no time savings, as the data can already be imported digitally. However, compared to data collection via paper questionnaire, there are substantial time savings with data transmission. The FSO estimates time savings to about 62 hours of working time per quarter, as displayed in Table 2 below.

**Table 2 Time savings through automatic data transmission**

Number of forms	Survey step	Data collection technique	Data collection technique ratio	Time savings per questionnaire (min.)	Total time savings (min.)
180	Initial	Paper	42%	10	756
180	Initial	Online	54%	0	0
180	Initial	Data transmission	4%	0	0
1'020	Follow-up	Paper	29%	5	1479
1'020	Follow-up	Online	67%	0	0
1'020	Follow-up	Data transmission	4%	0	0
400	Screening	Paper	75%	5	1500
400	Screening	Online questionnaire	24%	0	0
400	Screening	Data transmission	1%	0	0
<b>Total time savings per quarter (min)</b>					<b>3735</b>
<b>Total time savings per quarter (hours)</b>					<b>62</b>

#### Table 2 explained

The data of approximately 1600 dwellings will be delivered in the future, as all Garaio REM clients will be included (180 initial surveys, 1'020 follow-up surveys and 400 screening surveys). For each phase, questionnaires are distributed by survey technique. For the initial survey, 42% are currently answered on paper, 54% online and 4% via data delivery (Company X). The time required per processed questionnaire is approx. 10 minutes for an initial survey and 5 minutes for a follow-up survey or screening. This results in time saving of 756 minutes per quarter for the processing of the

initial survey, 1479 minutes for the follow-up survey and 1500 minutes for the screening. This represents approx. 62 working hours per quarter which can be used for other types of task.

### 5.3 Improvement of data quality for selected variables

Automatic data transmission increases data quality in two ways:

1. Since data is transmitted without human interaction, there are fewer errors in the collected data (such as typing errors) and the data is always interpreted in the same way (through preset interpretation rules).
2. Since more data needs to be transmitted to correctly interpret the questions asked, the interpretation of data becomes more consistent.

These two advantages regarding data quality can be illustrated with the following examples: the owner type and the rental prices.

As mentioned in chapter 3.1, the rental survey includes a question for which the property manager needs to interpret the owner type of the selected dwelling. On the questionnaire, the appropriate category needs to be checked out of 7 possible answers. Even though some guidance is provided by the suggested categories on the questionnaire, every answer is subject to the personal interpretation of each property manager. Moreover, there is no way the FSO can check whether the answer on the questionnaire is correct. The only provided information is the checked category. In contrast, more information is delivered by automatic data transmission in order for the FSO to correctly interpret the answer (see chapter 4.5.3). This allows the FSO to interpret the owner type and ensure it is correctly and consistently categorised for all dwellings.

The second example is the information delivered regarding rental prices. For rental prices, every now and then some property managers indicate on their questionnaires that the previously indicated rent was wrong. For the FSO this represents a challenge, since it has to be determined how to apply the data correction based on one check on the questionnaire that indicates an error in the previous survey. This is especially difficult to determine when the error was supposedly overlooked many times (e.g., is communicated only after several follow-up surveys). This issue no longer exists with automatic data transmission, since it originates directly from the landlord's database. There are no typing errors and no wrong data is accidentally picked one quarter (except for mismatched dwellings which are controlled at the end of the quarter). In addition to the current rent, the automatic data transmission also sends the previous rent and the dates from which the current and previous rents were valid. This enables the FSO to make sure no dwellings are mismatched by accident and to guarantee the rental index is based on high quality data.

## 6 Key learnings

### 6.1 Automatic data transmission is not child's play but worth the effort

Since the start of the project numerous challenges have been encountered. The entire project ended up taking much more time and resources than planned both for the FSO and Garaio REM AG. Among the challenges encountered, the most important ones were:

1. The implementation took far more resources than planned: from determining the importance of each variable, to identifying how each variable could be extracted from the landlord's software, to evaluating the quality of the landlord's data, to understanding why some dwellings won't match and make appropriate changes, to developing and implementing data interpretation rules for each variable, to developing an interface that allows data to be imported at the FSO, and finally to evaluating if the delivered data was consistent throughout the data collection techniques and transitioning to automatic data delivery. We estimate that

over one thousand working hours have been invested over the past 3 years on the FSO side of the project.

2. Data coverage and quality are not consistent throughout the Garaio REM AG clients and depend on each client's company policy on data management. For instance, there are some variables that are maintained with great care because they are important to the landlord's core business. Data quality is very high for all information regarding rental agreements including tenant's names, addresses or rental prices. However, for some variables that are crucial for the FSO, but less crucial to the landlord's core business (such as the EWID to identify the dwellings or secondary characteristics about the dwelling), the data coverage is not yet optimal.
3. Not all fields in the landlord real estate management software are standardised. Also, clients have some flexibility in the way they wish to use certain data fields. This creates some uncertainty as to what data to expect from other Garaio REM AG clients. Can the developed interpretation rules be applied for all clients or will there be a need for customisation?
4. Software updates can lead to changes in the way data is stored in the software and therefore how data is sent to the FSO.

## 6.2 Discovering potential errors and improving overall quality control

During the transition period, there were some dwellings that mismatched. The data transmitted automatically was not consistent with the data that was collected through questionnaires. There were differences in rents and characteristics of the dwelling such as floor, number of rooms and surface. Moreover, the tenant's name did not match anymore. With the help of Garaio REM AG and the insights into the landlord's data, the reasons behind those errors were discovered. This also led the FSO to expand its quality controls by adding more checks to the entire rental survey sample. The FSO was not aware of some of these errors beforehand and realised that they could also happen with the other data collection techniques.

One example is when a tenant moves within the same building. Since the tenant's name is one of the variables used to identify the selected dwelling, the FSO observed that instead of keeping the same dwelling and changing the tenant's name, one property manager kept the same tenant's name and changed the characteristics of the dwelling instead to include the move into the survey. In order to identify potential changes of dwellings during the survey, the FSO checks all dwellings with changes in rents, number of rooms and floors, which should not happen unless an error was made in the previous survey. This aspect was overlooked before the automatic data transmission project. Luckily these cases are rare, but they happen from time to time.

## 6.3 Insights into the real estate management software

A very valuable aspect of the collaboration with Garaio REM and Company X were the insights into the real estate management software. Before this project the FSO had no clue on how the rental data was managed and how the property managers proceeded to fill out a rental survey. Thanks to these insights, the questionnaire of the initial survey and follow-up survey were updated and improved.

- The EGID and EWID are now displayed on the questionnaire to help identifying the selected dwelling. Before, the only variables used to identify a selected dwelling were the address, the tenant's name, the number of rooms, the type of building and the floor.
- An entire page of questions was deleted in the initial survey questionnaire after a detailed data analysis of the landlord's data by Garaio REM AG showed that the data quality was mediocre for some variables<sup>10</sup>.

---

<sup>10</sup> These variables included: number of bathrooms, number of toilets, exterior spaces of the dwelling,

These variables were included in the rental questionnaire, even though they were not decisive for the calculation of the rental index, because the FSO was evaluating whether to use them to update the hedonic model used for quality adjustment in the rental price index. The quality of that data had already been called into question before the automatic data transmission project because of inconsistencies and a larger proportion of missing data. The insights gained from the real estate management software confirmed this and the idea of using these variables for the new hedonic model was eventually abandoned.

#### **How are questions answered when data is missing in the real estate management software?**

For a selection of variables Garaio REM AG analysed the proportion of missing data for five of the biggest real estate companies. One interesting example was the presence of elevators in the building. Information was recorded for 0.3% to 54% of all registered dwellings in the system, depending on the real estate company. This led the FSO to wonder on what basis a real estate manager, whose company records almost zero information on the presence of elevators in their data, answers this question in the rental survey. The conclusion was that the real estate managers must know some of this information by memory based on his knowledge of the managed dwellings. Such information cannot be extracted from the database for automatic data transmission.

## **6.4 Different data collection techniques may lead to different results**

When introducing new data collection techniques, their characteristics typically differ from each other. Online questionnaires, for example, allow for more flexible questionnaire design than paper questionnaires as display rules can be added depending on the type of respondent, the type of rental agreement, or the type of dwelling. Since the interpretation process of automatic data transmission is done on the data recipient's side through interpretation rules, more information is needed in order to correctly interpret the data. Therefore, the FSO needs to make sure that different data collection techniques do not lead to different results.

The selection of parking spaces for which an index is calculated and published separately are one of these challenges. In the paper and online questionnaires, the FSO asks landlords to indicate all parking spaces rented with the selected dwelling. It is not specified whether these parking spaces should be located in the same building or not. For automatic data transmission the extraction of this data had to be set up in a slightly different way. For a given dwelling, all parking space contracts related to the name of tenant 1 are extracted and transmitted. This could mean that if tenant 1 also rents a parking space at his work place from the same real estate company, it is also transmitted to the FSO even though it's not related to the selected dwelling for the rent survey. During a meeting with Garaio REM AG and Company X, a real estate manager said that they typically pick only the parking spaces located in the same building for the survey, while the other one answered they would also pick a parking space located in the neighbouring building if it's related to the selected dwelling. This illustrates how different data can be collected depending on the data collection process and also on whoever takes part in the collection process.

---

presence of elevators and Minergie-certificates, year of last total renovation.

## 7 Appendices

### 7.1 Appendix 1 mpeRequest\_FindUnit\_EgidEwid

```
<?xml version="1.0"?>
<delivery xmlns="http://www.garaio-rem.ch/schemas/mpe/request"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.garaio-
rem.ch/schemas/mpe/request ../schemas/mpe/request/2/1/mpeRequest.xsd"
  xmlns:header="http://www.ech.ch/xmlns/eCH-0058/5"
  xmlns:post="http://www.ech.ch/xmlns/eCH-0010/7"
  xmlns:pid="http://www.ech.ch/xmlns/eCH-0044/4"
  xmlns:contact="http://www.ech.ch/xmlns/eCH-0046/5"
  version="2.1">
  <deliveryHeader>
    <header:senderId>sedex://4-213246-6</header:senderId>
    <header:messageId>e355ba1d-ae17-4443-874a-
80389a935637</header:messageId>
    <header:messageType>sedex://1170</header:messageType>
    <header:sendingApplication>
      <header:manufacturer>BFS</header:manufacturer>
      <header:product>MPE Client</header:product>
      <header:productVersion>0.3.10</header:productVersion>
    </header:sendingApplication>
    <header:messageDate>2020-09-07T06:27:34Z</header:messageDate>
    <header:eventDate>2020-09-07</header:eventDate>
    <header:action>1</header:action>
    <header:testDeliveryFlag>1</header:testDeliveryFlag>
  </deliveryHeader>
  <request>
    <findUnitEgidEwid>
      <referenceDate>2020-01-01</referenceDate>
      <building>
        <EGID>444353</EGID>
      </building>
      <dwelling>
        <EWID>6</EWID>
      </dwelling>
    </findUnitEgidEwid>
  </request>
</delivery>
```



## 7.2 Appendix 2 mpeRequest\_FindUnit\_EgidTenant

```
<?xml version="1.0"?>
<delivery xmlns="http://www.garaio-rem.ch/schemas/mpe/request"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.garaio-rem.ch/schemas/mpe/request ../schemas/mpe/request/2/1/mpeRequest.xsd"
  xmlns:header="http://www.ech.ch/xmlns/eCH-0058/5"
  xmlns:post="http://www.ech.ch/xmlns/eCH-0010/7"
  xmlns:pid="http://www.ech.ch/xmlns/eCH-0044/4"
  xmlns:contact="http://www.ech.ch/xmlns/eCH-0046/5"
  version="2.1">
  <deliveryHeader>
    <header:senderId>sedex://4-213246-6</header:senderId>
    <header:messageId>5f55b5ea-b70d-4ac5-aa9f-632e8eda2887</header:messageId>
    <header:messageType>sedex://1170</header:messageType>
    <header:sendingApplication>
      <header:manufacturer>BFS</header:manufacturer>
      <header:product>MPE Client</header:product>
      <header:productVersion>0.3.10</header:productVersion>
    </header:sendingApplication>
    <header:messageDate>2020-09-07T06:27:34Z</header:messageDate>
    <header:eventDate>2020-09-07</header:eventDate>
    <header:action>1</header:action>
    <header:testDeliveryFlag>1</header:testDeliveryFlag>
  </deliveryHeader>
  <request>
    <findUnitEgidTenant>
      <referenceDate>2020-01-01</referenceDate>
      <building>
        <EGID>444353</EGID>
      </building>
      <dwelling>
        <!--EWID darf geliefert werden wird aktuell aber nicht ausgewertet-->
        <!--<EWID>6</EWID>-->
        <floor>3</floor>
        <roomCount>3</roomCount>
      </dwelling>
      <persons>
        <person>
          <officialName>Müller</officialName>
          <callName>Maria</callName>
        </person>
        <person>
          <officialName>Meier</officialName>
          <callName>Max</callName>
        </person>
      </persons>
      <tenantSearchDate>2019-09-01</tenantSearchDate>
    </findUnitEgidTenant>
  </request>
</delivery>
```

### 7.3 Appendix 3 mpeResponse

```
<delivery xmlns="http://www.garaio-rem.ch/schemas/mpe/response"
xmlns:header="http://www.ech.ch/xmlns/eCH-0058/5"
xmlns:post="http://www.ech.ch/xmlns/eCH-0010/7"
xmlns:contact="http://www.ech.ch/xmlns/eCH-0046/5"
xmlns:organisation="http://www.ech.ch/xmlns/eCH-0097/6" version="2.3">
  <deliveryHeader>
    <header:senderId>sedex://4-836264-8</header:senderId>
    <header:messageId>94cb34b9-1b06-4377-9267-
5e3cb98f54dd</header:messageId>

<header:referenceMessageId>638452300319763578</header:referenceMessageId>
<header:messageType>sedex://1171</header:messageType>
<header:sendingApplication>
  <header:manufacturer>GARAIO REM AG</header:manufacturer>
  <header:product>Mietpreiserhebung Service</header:product>
  <header:productVersion>0.0.185</header:productVersion>
</header:sendingApplication>
<header:messageDate>2024-03-05T09:15:37Z</header:messageDate>
<header:eventDate>2024-03-05</header:eventDate>
<header:action>1</header:action>
<header:testDeliveryFlag>1</header:testDeliveryFlag>
</deliveryHeader>
<response>
  <findUnitResult>
    <returnCode>UnitIsManaged</returnCode>
    <dwelling>
      <idxCategory>APPT</idxCategory>
      <idxType>1</idxType>
      <rent>
        <currentRent>
          <validFrom>2019-09-16</validFrom>
          <rentalPrice>2170.0</rentalPrice>
          <additionalCosts>320.0</additionalCosts>
          <totalPrice>2490.0</totalPrice>
          <rentComponents>
            <rentComponent>
              <amount>110.0</amount>
              <description>Heiz- / Warmwasserkosten Akonto</description>
            </rentComponent>
            <rentComponent>
              <amount>210.0</amount>
              <description>Nebenkosten Akonto</description>
            </rentComponent>
          </rentComponents>
          <basisData>
            <costIncreaseDate>2019-07-31</costIncreaseDate>
            <mortgageRate>1.5</mortgageRate>
            <mortgageRateDate>2017-06-02</mortgageRateDate>
            <priceIndexBaseYear>2010</priceIndexBaseYear>
            <priceIndexDate>2019-07-31</priceIndexDate>
          </basisData>
        </currentRent>
      </rent>
    </dwelling>
  </findUnitResult>
</response>
</delivery>
```

```

</rent>
<floor>4</floor>
<surface>124.07</surface>
<roomCount>5.0</roomCount>
<subsidy />
</dwelling>
<building>
  <EGID>881251</EGID>
  <latestYearOfConstruction>1962</latestYearOfConstruction>
  <structuralDetails>
    <constructions>
      <detail>
        <year>1962</year>
        <remark />
      </detail>
    </constructions>
  </structuralDetails>
  <addressInfo>
    <post:street>Samplestreet</post:street>
    <post:houseNumber>10</post:houseNumber>
    <post:town>Sampletown</post:town>
    <post:swissZipCode>1999</post:swissZipCode>
    <post:country>
      <post:countryIdISO2>CH</post:countryIdISO2>
      <post:countryNameShort />
    </post:country>
  </addressInfo>
</building>
<property>
  <administrationKind>VOLL</administrationKind>
</property>
<occupancy>
  <tenancyAgreement>
    <reference>2433.08.0801.12</reference>
    <isSecondaryResidence>false</isSecondaryResidence>
    <isTenancyResidential>true</isTenancyResidential>
    <tenants>
      <tenant>
        <officialName>Samplename</officialName>
        <givenName>Samplename</givenName>
        <tenantKind>PHYSICAL</tenantKind>
      </tenant>
      <tenant>
        <officialName> Samplename2 </officialName>
        <givenName> Samplename2 </givenName>
        <tenantKind>PHYSICAL</tenantKind>
      </tenant>
    </tenants>
    <usage>Wohnung</usage>
    <rentKind>ExternalRent</rentKind>
    <validFrom>2019-09-16</validFrom>
    <agreementCategory>Mietvertrag Wohnung</agreementCategory>
    <subsidy>

```

```

        <wegAdditionalReduction1>false</wegAdditionalReduction1>
        <wegAdditionalReduction2>false</wegAdditionalReduction2>
    </subsidy>
</tenancyAgreement>
</occupancy>
<contractor>
    <contact>
        <contact:address>
            <contact:postalAddress>
                <post:organisation>
                    <post:organisationName>Company X</post:organisationName>
                    <post:organisationNameAddOn1>Real Estate
Company</post:organisationNameAddOn1>
                    <post:firstName>Samplename</post:firstName>
                    <post:lastName> Samplename </post:lastName>
                </post:organisation>
                <post:addressInformation>
                    <post:street>Samplestreet</post:street>
                    <post:houseNumber>10</post:houseNumber>
                    <post:postOfficeBoxText>1999</post:postOfficeBoxText>
                    <post:town>Sampletown</post:town>
                    <post:swissZipCode>1999</post:swissZipCode>
                    <post:country>
                        <post:countryIdISO2>CH</post:countryIdISO2>
                        <post:countryNameShort>Schweiz</post:countryNameShort>
                    </post:country>
                </post:addressInformation>
            </contact:postalAddress>
        </contact:address>
        <contact:email>

<contact:emailAddress>samplename.samplename@companyX.ch</contact:emailAddress>
    </contact:email>
    <contact:phone>
        <contact:phoneNumber>0584445566</contact:phoneNumber>
    </contact:phone>
</contact>
</contractor>
<owner>
    <identifiedOrganisation>
        <uid>

<organisation:uidOrganisationIdCategory>CHE</organisation:uidOrganisationIdCategory>

<organisation:uidOrganisationId>380550119</organisation:uidOrganisationId>
    </uid>
    </identifiedOrganisation>
</owner>
</findUnitResult>
</response>
</delivery>

```

## 7.4 Appendix 4 rental index compilation in more detail

The recorded rents are assigned to a specific cell according to age and number of rooms; within each cell, two sub-indices are calculated: one for the panel segment (segment B), i.e., observations in both the current and previous period, and one for the rotating segment (segment AC), i.e., incoming (segment C) and outgoing (segment A) (Figure 3).

**Figure 3: cell subdivision**

	Previous period t-1	Current period t
Dwellings present in t-1 only	A	
Dwellings present in t-1 and in t	B	B
Dwellings present in t only		C

Source: FSO – rental index

© FSO 2022

The segment B sub-indices are defined as the relationship between the average rents of the two periods, with each calculated using a geometrical mean wherein each rent is weighted using its respective survey weight. No quality adjustments are made, since the same dwellings are present in both periods and the quality of dwellings contained therein is deemed to be constant.

$$L_B^t = \frac{\tilde{x}_B^t}{\tilde{x}_B^{t-1}} \cdot 100 = \frac{\left( \prod_{i=1}^{n_B^t} (x_{iB}^t)^{p_i^t} \right)^{\frac{1}{\sum_i p_i^t}}}{\left( \prod_{i=1}^{n_B^{t-1}} (x_{iB}^{t-1})^{p_i^{t-1}} \right)^{\frac{1}{\sum_i p_i^{t-1}}}} \cdot 100$$

where:

$L_B^t$  = Index of segment B in period t, in comparison with the previous period

$\tilde{x}_B^t$  = Geometrical average of rents in cell B in period t

$x_{iB}^t, i = 1, \dots, n_B^t$  = Observations (rents) in period t for cell B

$n_B^t$  = Number of observations in cell B in period t

$p_i^t, i = 1, \dots, N^t$  = Weight of observation i in period t. This is the sampling weight adjusted for non-response

For the segment AC sub-indices, the dwellings that were present in A the previous period are not comparable with those newly present in C, even though they are in same age class and have the same number of rooms. Ex-post stratification is not sufficient to adjust dwellings quality. Other variables such as surface area, floor, region, type of commune, location, etc. can have a considerable influence on rent levels.

The method used for quality adjustment is "hedonic repricing" (or hedonic quality adjustment)<sup>11</sup>, which is often used in the production of official statistics. Fundamentally, this relies on the estimation of an

<sup>11</sup> For more information about the hedonic model, see, [Rental price index: methodological report - New hedonic rental price model | Publikation | Bundesamt für Statistik \(admin.ch\)](#)

explicit quality-adjustment factor  $\hat{g}^t$

$$\hat{g}^t = \frac{\hat{x}_C^t}{\hat{x}_A^{t-1}} = \frac{\left( \prod_{i=1}^{n_C^t} (\hat{x}_{iC}^t)^{p_i^t} \right)^{1/\sum_i p_i^t}}{\left( \prod_{i=1}^{n_A^{t-1}} (\hat{x}_{iA}^{t-1})^{p_i^{t-1}} \right)^{1/\sum_i p_i^{t-1}}}$$

where

$\hat{x}_C^t$  = Geometrical average of estimated rents in segment C in period t

$\hat{x}_A^{t-1}$  = Geometrical average of estimated rents in segment A in period t-1

...which is then applied to the recorded price change between A and C

$$L_C^t = \frac{\tilde{x}_C^t}{\tilde{x}_A^{t-1} \cdot \hat{g}_c^t} \cdot 100 = \frac{\left( \prod_{i=1}^{n_C^t} (x_{iC}^t)^{p_i^t} \right)^{1/\sum_i p_i^t}}{\left( \prod_{i=1}^{n_A^{t-1}} (x_{iA}^{t-1})^{p_i^{t-1}} \right)^{1/\sum_i p_i^{t-1}} \cdot \hat{g}_c^t} \cdot 100$$

where:

$\tilde{x}_C^t$  = Geometrical average of observed rents in segment C in period t

$\tilde{x}_A^{t-1}$  = Geometrical average of observed rents in segment A in period t-1

$L_C^t$  = Segment C index in period t, in comparison with segment A in period t-1

The two sub-indices pro cell are then combined arithmetically on the basis of the number of observations in each segment to determine the cell's overall index

$$L^t = \frac{n_B}{n_B + n_C} \cdot L_B^t + \frac{n_C}{n_B + n_C} \cdot L_C^t$$

where:

$L^t$  = Overall cell index for period t

$n_B$   $n_C$  = Number of observations in segments B and C

The total index is then obtained by arithmetically combining the 24 cells using their respective weights.